Conditional Tests for the Profitability of Technical Analysis in Currency Trading and Its Economic Fundamentals

Ilias Filippou[a], Po-Hsuan Hsu[b], Tren Ma[c], Georgios Sermpinis[d], Mark P. Taylor[a]

[a] *John M. Olin Business School, Washington University in St. Louis, St. Louis, MO 63130-4899, USA*
[b] *College of Technology Management, National Tsing Hua University, Taiwan*
[c] *Nottingham University Business School, University of Nottingham, Nottingham NG8 1BB, UK*
[d] *Adam Smith Business School, University of Glasgow, Glasgow G12 8QQ, UK*

Abstract

We develop a multivariate functional false discovery rate ($mfFDR$) method that accounts for multiple informative covariates –and the new information they carry– to examine the conditional performance of predictive models with better power. Our simulations suggest that this advantage remains with (i) data dependence, (ii) estimated errors in covariates, and (iii) correlated covariates. We apply the proposed method to control luck in detecting conditionally profitable technical trading rules using 30 developed and emerging market currencies. Our method selects more profitable rules than prior methods; more importantly, these rules offer better out-of-sample performance. Further analyses suggest that such profitability exists due to the barriers to arbitragers as it decreases with computational power, capital account openness, and illiquidity.

Keywords: Multiple testing, Multivariate functional false discovery rate, Informative covariates, Technical analysis, Foreign exchange markets.

Email addresses: iliasfilippou@wustl.edu (Ilias Filippou), pohsuanhsu@mx.nthu.edu.tw (Po-Hsuan Hsu), Tren.Ma@nottingham.ac.uk (Tren Ma), georgios.sermpinis@glasgow.ac.uk (Georgios Sermpinis), mark.p.taylor@wustl.edu (Mark P. Taylor)

1. Introduction

Technical analysis and trading rules are widely used in foreign exchange (FX) trading (Allen and Taylor, 1990; Taylor and Allen, 1992; Menkhoff and Taylor, 2007). Meese and Rogoff (1983) and Chinn and Meese (1995) document that major exchange rates follow a random walk. In contrast, Levich and Thomas (1993) and Neely *et al.* (2009) argue that the profitability of technical trading rules existed in the 1970s and 1980s but declined in the early 1990s, and Neely *et al.* (1997) and Neely (2002) provide evidence showing that such profitability cannot be attributed to systematic risk or government interventions.

However, given that (i) technical trading rules are not theoretically motivated, and that (ii) there exists an enormous number of such rules, the profitability of technical analysis in FX markets is likely subject to data snooping issues.[1] Among all tools to mitigate data snooping bias in social sciences, the multiple testing framework based on White (2000), and Romano and Wolf (2005) offers computationally feasible solutions and has thus been widely applied to research questions involving a large number of predictive models.[2] One common feature of the methodologies in this framework is that the rejection criterion *only* depends on information created by predictive models and does not account for external information.

Using the aforementioned multiple testing methods, some recent studies show that technical trading rules' profitability has declined since the 1990s (Qi and Wu, 2006; Hsu *et al.*, 2016). Nevertheless, such a declining pattern is based on null hypotheses being unconditional zero, which corresponds to investors *not* acquiring external information in assessing technical rules' profitability. Despite its prevalence and convenience, such unconditional testing may not be ideal because currency traders are assessing and picking trading rules based on frequently updated external information. Thus, the value of technical analysis for FX traders may not be appropriately estimated in an unconditional testing framework as has been done in the literature.

In this paper, we propose a new methodology that accounts for more information

[1]Data-snooping or p-hacking is a statistical bias that appears when a dataset is used more than once, for inference and model selection. It can lead to results that seem statistically significant but are due to luck and misuse of data analysis.

[2]Aiming to control for data-snooping in assessing trading rules, numerous multiple testing procedures have been proposed such as the contributions of Sullivan *et al.* (1999, 2001), White (2000), Hansen (2005), Barras *et al.* (2010), Bajgrowicz and Scaillet (2012) and Hsu *et al.* (2010) among others.

sources – which can be captured by "informative covariates" – in forming the rejection criterion, which enhances the statistical power given the same false discovery rate (FDR) defined by Benjamini and Hochberg (1995) (BH henceforth). We name our method as multivariate functional FDR ($mfFDR$). Conceptually speaking, embedding informative covariates (and new information they carry) in multiple testing enables us to form conditional null hypotheses, in which predictive models' profits are zeros conditional on updated information. Such a conditional setting is more consistent with the rational expectation hypothesis and better captures market participants' time-varying standards.

Using simulations, we show that our $mfFDR$ method performs well in controlling for the FDR under various settings. Its performance in terms of power is impressive and beats that of prior methods (i.e., the FDR approach of Storey (2002) and the $fFDR$ of Chen *et al.* (2021) that allows *only one* informative covariate) by more than 44%.[3] In addition, the proposed procedure performs well under weak signal-to-noise data, i.e., the data where the true alternative hypothesis tests have high $p-$values due to high noise. It is also robust under dependent data and when informative covariates are correlated or contain estimation errors.

We implement the $mfFDR$ method to detect conditionally profitable trading rules in a set of more than 21,000 technical trading rules. Using daily data over a maximum of 50 years for 30 U.S. dollar exchange rates, we use our $mfFDR$ method to select rules that outperform zero in a rolling 12-month window. The informative covariates we consider include (i) the auto-correlation of a trading rule's excess return and (ii) the estimates of the alpha, beta, and R-square from a regression of the trading rule's excess return on the excess return of a passive buy-and-hold strategy or a currency market factor.

We then collect outperforming rules to construct a monthly portfolio for a currency or a basket of currencies and track the performance of these portfolios with and without transaction costs. We find that it gains positive profits in 28 over 30 considered currencies for an extended period from 1973 to 2020 after transaction costs, and 16 among them are statistically significant.

To examine the advantage of the $mfFDR$ method, the number of rules selected by

[3]As benchmarks, prior methods include the FDR approach of Storey (2002) and the functional FDR ($fFDR$) in Chen *et al.* (2021) (CRS henceforth) that allows *only one* informative covariate. These two methods have been applied to stock index and mutual fund performance by Barras *et al.* (2010), Bajgrowicz and Scaillet (2012), and Hsu *et al.* (forthcoming).

the $mfFDR$ is larger than those based on prior methods (the FDR and the $fFDR$). More importantly, we find that the out-of-sample (OOS) performance of the $mfFDR$-based portfolio is better than those based on prior methods in two aspects. First, the $mfFDR$-based portfolio with the use of four covariates beats all $fFDR$-based with the use of every single covariate. Second, using a linear combination of the four covariates, such as the first principal (PC1), does not improve the performance of the $fFDR$-based approach compared to using individual underlying covariates. Finally, when we construct a larger portfolio, which is based on a pool of all trading rules applied on all 30 currencies (about 635,850 trading rules = 21,195 $\times$ 30), we find that it gains a Sharpe ratio of 1.06 and 0.95 before and after transaction costs, respectively.

These empirical results highlight the value of directly incorporating more information in multiple testing and offer the following insights. First, considering more covariates in the $mfFDR$ enhances the OOS performance of detected out-performing trading rules. Second, the $mfFDR$ outperforms the $fFDR$ with linear combinations of multiple covariates, suggesting that the $mfFDR$ might effectively extract non-linear information among covariates. Third, prior methods based on the unconditional null hypothesis, such as the FDR, may underestimate technical trading rules' true predictive ability and profits because their performance is not evaluated with comprehensive information sets.

Our second set of empirical analyses aims to present the evolution of the profitability of technical analysis. We split our sample into five decades (1973-1980, 1981-1990, 1991-2000, 2001-2010, and 2011-2020) and examine the performance of technical trading rules using the $mfFDR$ in each decade. We still find a decent number of profitable technical trading rules in the most recent three decades.

We further investigate the performance by rule categories. We show that there is time series variation in the performance of different categories of rules. In particular, the moving average category is the most profitable group over time, both in-sample and OOS.

In our fourth set of analyses, we explore the source of the $mfFDR$-based portfolio. Thereby, we study the evolution of selected rules by the $mfFDR$ in each currency in terms of both the OOS return and the weight of funds distributed into each currency (represented by the proportion of the number of out-performing rules selected in the currency over these numbers in whole currency sample). We find that the profitability of the selected rules differs across currencies and over time periods. In considering each period, there are always some currencies where the selected rules are profitable, and more importantly, the $mfFDR$ correctly allocates more funds into those profitable currencies. These two factors make up the profitability of the $mfFDR$-based portfolio over each of the time periods.

Finally, we attempt to analyze the economic fundamentals behind the profitability of technical rules in FX market. As such, we use the OOS returns of the $mfFDR$-based portfolio and use regressions to understand their determinants. We find that the OOS returns decreases with country-specific computational power, capital account openness, FX market illiquidity, and equity market development.We argue that all these factors are related to the existence and density of arbitragers who speed up price adjustment toward new equilibrium. As technical trading rules mainly detect the trends (or changes in trends) (Menkhoff and Taylor, 2007), they may exploit the delayed or slow changes in equilibrium FX rates. Hence, our empirical evidence suggests that the profitability of technical analysis exists due to the barriers to arbitragers.

Our study contributes to the finance and econometrics literature as follows. From a methodological perspective, we introduce the use of updated external information in testing for data snooping biases.[4] Under the rational expectation hypothesis, the performance of predictive models (and associated null hypotheses) should reflect researchers' and industry practitioners' time-varying expectations. The $mfFDR$ methodology we propose allows researchers to utilize comprehensive information from multiple covariates to test conditional null hypotheses that appear to be more realistic to both academia and industry. By designing and implementing suitable Monte Carlo simulations, we illustrate that our $mfFDR$ approach actually controls for FDR under weak signal-to-noise data, dependent data, or correlated covariates (even those with estimation errors). In addition, our simulations suggest that the $mfFDR$ method has higher power than prior methods (FDR and $fFDR$) that do not update for sufficient information.

From an empirical perspective, we perform the most comprehensive study of technical

[4]Some prior literature highlights the data snooping issues, which leads to the development of some methodologies to guard against such biases. While several multiple testing procedures have been proposed in the past (White, 2000; Hansen, 2005; Barras *et al.*, 2010; Bajgrowicz and Scaillet, 2012; Hsu *et al.*, 2010), they only consider unconditional null hypotheses and use information from predictive models' performance metrics.

trading rules in FX markets to timely assess the predictability of such rules and provide further insights on FX traders' "obstinate passion" in technical analysis (Menkhoff and Taylor, 2007). Our analyses are based on constructing 635,850 trading rules, using FX data of all 30 currencies for a long period (some are as long as almost five decades), and implementing large-scale multiple testing for the profitability of technical analysis. The conditional profitability of some technical rules based on the $mfFDR$ method explains why these simple rules are still popularly used by traders (even though they have been rejected by many prior studies using unconditional tests). More importantly, we implement comprehensive analyses on the economic fundamentals of such profitability and link it to the barriers to arbitragers who accelerate the price adjustments toward new equilibrium.

The rest of the paper is organized as follows. The next section develops the $mfFDR$ framework and designs simulations to show its performance. Section 3 presents descriptions of the data and trading rule universe, and Section 4 is devoted to the trading rule's performance measure. In Section 5, we present the empirical results where the $mfFDR$-based portfolios are constructed on individual currencies and a basket of 30 currencies. Finally, Section 7 concludes this paper.

2. The use of covariates in FDR framework

In this section, we introduce the $mfFDR$ that estimates the false discovery rate as a function of more than one informative covariate. Our approach develops the frameworks of CRS and Hsu *et al.* (forthcoming) on the $fFDR$ with a single covariate. First, we present the setting of our method and its implementation in the context of data snooping and FX trading. Second, we conduct simulations demonstrating the value of multiple informative covariates in controlling the false discovery rate and the superior power of $mfFDR$ compared to the related existing approaches. Third, we validate the performance of our method under certain dependence structures of data which are typical in finance. We confirm that our approach retains its power and control of the false discovery rate when the statistics are dependent and when informative covariates are correlated and contain estimation errors - all these features are common in most financial data and topics.

2.1. The multivariate functional false discovery rate (mfFDR)

Suppose we have n trading rules, each producing an excess return. Assume that there are covariates that convey information about the performance of each trading rule, which is measured by a metric ϕ. As detailed in the following sections, in this study, we use the Sharpe ratio (SR) as our main performance metric. To assess the performance of each strategy, conditional on the realization of the covariates, we conduct a hypothesis test

$$H_0 : \phi = 0, \qquad H_1 : \phi \neq 0. \tag{1}$$

The n trading rules produce n conditional tests as in test (1). Our aim is to detect a maximal number of strategies having significant none-zero ϕ while controlling for FDR, the expected proportion of false discoveries among the hypothesis tests called significant as introduced in BH.

For the convenience of notation, let us consider the single test (1). To formulate the assumptions, we assume that there are d covariates, represented by random variables $Z^1, \ldots, Z^d$, conveying information about the probability of a hypothesis being true null as well as the distribution of the $p-$value of the false null hypothesis. Let us denote by P the random variable representing the $p-$value of the test and $\boldsymbol{Z} = (Z^1, \ldots, Z^d)$. Also, let h be the true status of the hypothesis, that is, $h = 0$ if the null hypothesis is true and $h = 1$ if the null hypothesis is false. To indicate a particular test corresponding to a fund i, we add the subscript i to all mentioned notations, i.e., P_i, h_i and $\boldsymbol{Z}_i$.

To control the FDR at a level $\tau \in (0, 1)$, the conventional decision rule is a null hypothesis i is rejected if and only if $p_i \leq \hat{\Theta}$ where $\hat{\Theta}$ is a common threshold for all hypotheses, determined via a data-driven manner depending on particular procedures, and p_i is observed value of P_i.[5] In contrast, in this study, the threshold depends on realized values of covariates $\boldsymbol{Z}$ and varies across hypotheses, i.e., the rejection condition becomes $p_i \leq \hat{\Theta}(\boldsymbol{z}_i)$ and will be expressed in an implicit form as presented below. As such, our rejection rule will be conditional on the information we have via the realization of the covariates $\boldsymbol{Z}$ that reflect extra information about each hypothesis. This approach

[5]For instance, the procedure of BH ranks the hypotheses based on $p-$values from smallest to highest. Denote by $\tilde{p}_j$s the $p-$values after ranking, the BH procedure seeks for $j^* = \max\{j | \tilde{p}_j \leq j \times \tau/N\}$ where $N = n$, then sets $\Theta = \tilde{p}_{j^*}$. Storey (2002) uses the same procedure with $N =$ (number of $p_i > \lambda$)$/(1 - \lambda)$ for some $\lambda \in (0, 1)$.

is different from conventional ones where the information about $\boldsymbol{Z}$ is ignored, and the hypotheses are treated the same based on a fixed threshold of $p-$value regardless of the differences in the value of $\boldsymbol{Z}$.

For each j ($j = 1, \ldots, d$), we transform the observed value of the covariate Z_i^j to a form so that Z_i^j is uniformly distributed in the interval $[0, 1]$, via using $z_i^j = r_i^j/n$ where r_i^j is the rank of the observed value of Z_i^j in the set of observed values of $Z_1^j, \ldots, Z_n^j$, $j = 1, \ldots, d$. Next, Z^js are the ones in the transformed forms. We assume that, conditional on realization $\boldsymbol{z}$ of $\boldsymbol{Z}$, the hypothesis is a priori true null with probability $\pi_0(\boldsymbol{z})$, i.e. $(h|\boldsymbol{Z} = \boldsymbol{z}) \sim Bernoulli(1 - \pi_0(\boldsymbol{z}))$.

To formulate the theoretical framework, we require $\boldsymbol{Z}$ and P to satisfy: i) Z^js are mutually independent across $j = 1, \ldots, d$; ii) Conditional on $\boldsymbol{Z} = \boldsymbol{z}$, when the null hypothesis is true, P is uniformly distributed on the interval $[0, 1]$ and when the null hypothesis is false, P has a distribution determined by some density function $f_{alt}(p|\boldsymbol{z})$. To develop the theoretical framework, we assume that the aforementioned n tests are independent replications of the test (1), i.e., the triples $(h_i, P_i, \mathbf{Z}_i)$ of tests $i = 1, \ldots, n$ are independent and identically distributed as the triple $(h, P, \boldsymbol{Z})$. In the next sections, we show that our method can also be applied to scenarios when there is a dependence among test statistics and covariates.

The gist of our method is as follows. Given a target τ of the FDR, we do not decide to reject the null hypotheses based on their $p-$values solely. We instead reject a null hypothesis by a rule based on both the $p-$value and the covariates. Thereby, we define each hypothesis with an observed $(p, \boldsymbol{z})$, a posterior probability of being null denoted by $r(p, \boldsymbol{z})$. More specifically,

$$r(p, \boldsymbol{z}) = \mathbb{P}(h = 0|(P, \boldsymbol{Z}) = (p, \boldsymbol{z})) \tag{2}$$

which can be developed further as

$$r(p, \boldsymbol{z}) = \frac{\mathbb{P}(h = 0|\boldsymbol{Z} = \boldsymbol{z})}{\mathbb{P}((P, \boldsymbol{Z}) = (p, \boldsymbol{z}))} = \frac{\pi_0(\boldsymbol{z})}{f(p, \boldsymbol{z})} \tag{3}$$

where $f(p, \boldsymbol{z})$ is the joint density function of the $p-$value and covariates.[6]

[6]The equation (3) is obtained by using the fact that $\mathbb{P}(h = 0, P = p, \boldsymbol{Z} = \boldsymbol{z}) = \mathbb{P}(P = p|(h = 0, \boldsymbol{Z} = \boldsymbol{z})).\mathbb{P}(h = 0|\boldsymbol{Z} = \boldsymbol{z}).\mathbb{P}(\boldsymbol{Z} = \boldsymbol{z}) = \mathbb{P}(h = 0|\boldsymbol{Z} = \boldsymbol{z})$ where the first and last factors equal one as they are

Empirically, $r(p, \boldsymbol{z})$ is estimated by $\hat{r}(p, \boldsymbol{z}) = \hat{\pi}_0(\boldsymbol{z})/\hat{f}(p, \boldsymbol{z})$ where $\hat{\pi}_0(\boldsymbol{z})$ and $\hat{f}(p, \boldsymbol{z})$ are estimators of $\pi_0(\boldsymbol{z})$ and $f(p, \boldsymbol{z})$, respectively. For the sake of space, we refer to the details of the estimation procedures in Appendix A.

If there are any significant hypotheses, the hypothesis with the smallest $\hat{r}(p, \boldsymbol{z})$ will be selected first, then the second smallest one, and so on. Each time a hypothesis is added to the significant set, the FDR is raised. We stop the procedure when the FDR target is reached. Consequently, the rejection region has a form $\Gamma(\theta) = \{(p, \boldsymbol{z}) | r(p, \boldsymbol{z}) \leq \theta\}$ where $\theta \in (0, 1)$ satisfies

$$\int_{\Gamma(\theta)} r(p, \boldsymbol{z}) dp d\boldsymbol{z} \leq \tau. \tag{4}$$

The left side of the condition (4) is the FDR corresponding to the rejection region $\Gamma(\theta)$. Hence, we choose a maximal threshold $\theta = \theta^*$ such as the condition (4) holds. Thus, a hypothesis is significant if and only if its observed $(p, \boldsymbol{z})$ belongs to the set $\Gamma(\theta^*)$. This rejection rule implicitly stands for the mentioned condition $p_i \leq \hat{\Theta}(\boldsymbol{z}_i)$.[7]

In implementation, as $r(p, \boldsymbol{z})$ is estimated by $\hat{r}(p, \boldsymbol{z})$, the FDR can be estimated by

$$\widehat{FDR}_\theta = \frac{1}{\#S_\theta} \sum_{k \in S} \hat{r}(p_k, \boldsymbol{z}_k), \tag{5}$$

where $S_\theta = \{k | \hat{r}(p_k, \boldsymbol{z}_k) \leq \theta, k \in \{1, \ldots, n\}\}$ and $\#S_\theta$ returns the number of elements in the set S_θ.[8]

It is clear that if the null of a hypothesis i is rejected, then so are all hypotheses having $\hat{r}(p, \boldsymbol{z}) \leq \hat{r}(p_i, \boldsymbol{z}_j)$. To make the rejection rule simple and efficient, we use the "functional $q-$value" introduced in CRS and Storey (2003). Intuitively, for each hypothesis i with observed $(p, \boldsymbol{z}) = (p_i, \boldsymbol{z}_i)$, we determine its $q-$value equal to the estimate of the FDR when we reject the null of all hypotheses k having $\hat{r}(p_k, \boldsymbol{z}_k) \leq \hat{r}(p_i, \boldsymbol{z}_i)$, i.e., the $\widehat{FDR}_\theta$ when $\theta = \hat{r}(p_i, \boldsymbol{z}_i)$. Thus, at the given target τ of the FDR, a null hypothesis is rejected if and only if its corresponding $q-$value $\leq \tau$.

density functions of uniform distributions.

[7]If we additionally assume that the $f(p, z)$ is a decreasing function of p, then for each z the $r(p, z)$ is an increasing function of p. The null hypothesis i is rejected if and only if $\hat{r}(p_i, z_i) \leq \theta^*$ which is equivalent to $p_i \leq \hat{\Theta}(\boldsymbol{z}_i)$ for some $\hat{\Theta}(\boldsymbol{z}_i)$.

[8]The $\#S_\theta$ is the number of discoveries given θ while the numerator is the expected number of false discoveries. This estimation is proposed by Newton *et al.* (2004) and Storey *et al.* (2005) and subsequently adopted in CRS.

We name the proposed procedure, where the $\boldsymbol{z}$ is a vector of more than one covariate, as the $mfFDR$. When $d = 1$, the $mfFDR$ is the $fFDR$ of CRS. Thus, we present an intuitive illustration of the value of informative covariates in the $mFDR$ framework, and we compare the performance of the proposed method to others in terms of FDR control and power.

2.2. Simulation studies

We consider the simplest case where we have two informative covariates $\boldsymbol{Z} = (U, V)$. We simulate $n = 10,000$ hypotheses where the proportion of null hypotheses is approximately 0.8.[9] Suppose that the two covariates convey information about the hypotheses. We will demonstrate that by using the $mfFDR$, which utilizes both covariates as inputs, we obtain a higher power in detecting false null hypotheses than the $fFDR$, which uses only one of the two covariates.

The data-generating process in our $mfFDR$ simulations is as follows. In each iteration, we independently draw the elements of the covariates $U = (u_1, \ldots, u_n)$ and $V = (v_1, \ldots, v_n)$ from the uniform distribution $U(0, 1)$. For each hypothesis i we draw its null status h_i from a Bernoulli distribution with a probability of being null $\mathbb{P}[h_i = 0|(u, v) = (u_i, v_i)] = \pi_0(u_i, v_i)$, i.e. $h_i \sim Bernoulli(1 - \pi_0(u_i, v_i))$, where $\pi_0(u, v)$ has one of the two following forms

- $\pi_0(u, v) = \sin[\pi(u + v)/2]$, i.e. a sine function;
- $\pi_0(u, v) = 1 - (u^4 + v^4)/2$, i.e. a monotonic function (concerning each covariate).

We obtain two bundles of (U, V, H) that correspond to the two forms of the $\pi_0(u, v)$, where $H = (h_1, \ldots, h_n)$.

Next, for each triple (U, V, H), we generate $p-$values for tests such that for each true null hypothesis, i.e. the one with $h_i = 0$, its $p-$value is drawn from the uniform distribution $U[0, 1]$, whereas the $p-$value of a false null hypothesis follows the distribution with a density function $f_{alt}(p|(u, v))$. We specify $f_{alt}(p|(u, v))$ by using a Beta distribution $Beta(\alpha, \beta)$, i.e. $f_{alt}(p|(u, v)) \propto p^{\alpha-1}(1-p)^{\beta-1}$ where the α and β are positive real parameters determining the shape of the distribution. Similar to CRS, we set the β as a function

[9] In Section IB of the Internet Appendix, we show that the results are robust under varying of the number of tests used.

of the covariates, specifically as $\beta = 3 + 1.5(u + v)$. Aiming to study different methods' performance in various circumstances, we consider three cases of $\alpha \in \{0.5, 1, 1.5\}$ which, for convenience in presenting, we name as strong, weak and very weak signal (relative to noise), respectively (thus, we have three specifications of a Beta distribution). In the strong signal case, the false null hypotheses are more easily distinguished from the true null ones than those in the weak and very weak signal cases.[10]

The task is to detect the false null hypotheses from the simulated sample with control of the FDR at given targets by using only the $p-$value and the covariates. In our first experiment, we illustrate the role of informative covariates in detecting false null hypotheses. Then we benchmark our procedure, the $mfFDR$ with both U and V as covariates, against the $fFDR$ with only U as a covariate, the FDR of Storey (2002) which we notate as standard FDR ($StdFDR$), and the FDR of BH. By comparing the obtained results with the null status of the hypotheses (i.e., the H), we calculate a false discovery proportion, which is the ratio of the number of true null hypotheses falsely rejected over the number of the discoveries and a correct detection proportion, which is the ratio of the number of false null hypotheses detected over the number of false null ones in the population.[11]

We assess the two most important criteria of an FDR procedure: the control of the FDR and the power. To do so, we implement the $mfFDR$ and benchmark procedures at FDR targets $\tau \in \{0.05, 0.1, \ldots, 0.95\}$ over 1000 iterations and average the false discovery proportions and the correct detection proportions to have estimates of the actual FDR and the power, i.e. the expectation of the mentioned correct detection proportion, respectively. In total, we are studying six cases corresponding to the combinations of the two forms of the function $\pi_0(u, v)$ and the three aforementioned specifications of the Beta distribution. We also demonstrate that our method retains its power when the covariates are correlated with each other and are subject to estimation errors. Finally, we illustrate the FDR control and the power of the $mfFDR$ when the $p-$values are dependent which

[10]Section IA of the Internet Appendix illustrates, via analysing a specific sample, how different the three cases are.

[11]In Section IA of the Internet Appendix, we illustrate how the use of informative covariates can enhance the power of $mfFDR$ in detecting false nulls compared to other methods. Briefly, by partitioning the sample into groups based on value of the covariates, we observe that: i) the $mfFDR$ correctly rejects more false nulls compared to others in the groups with rich false nulls; and ii) having more covariates helps.

will be the case when each hypothesis test represents a technical trading rule that stands for a specific combination of parameters. To assess the first criterion, we compare the estimated actual FDR of the $mfFDR$ to the given FDR targets, while for the second one, we compare the power of $mfFDR$ against that of the $fFDR$, $StdFDR$, and BH approaches.

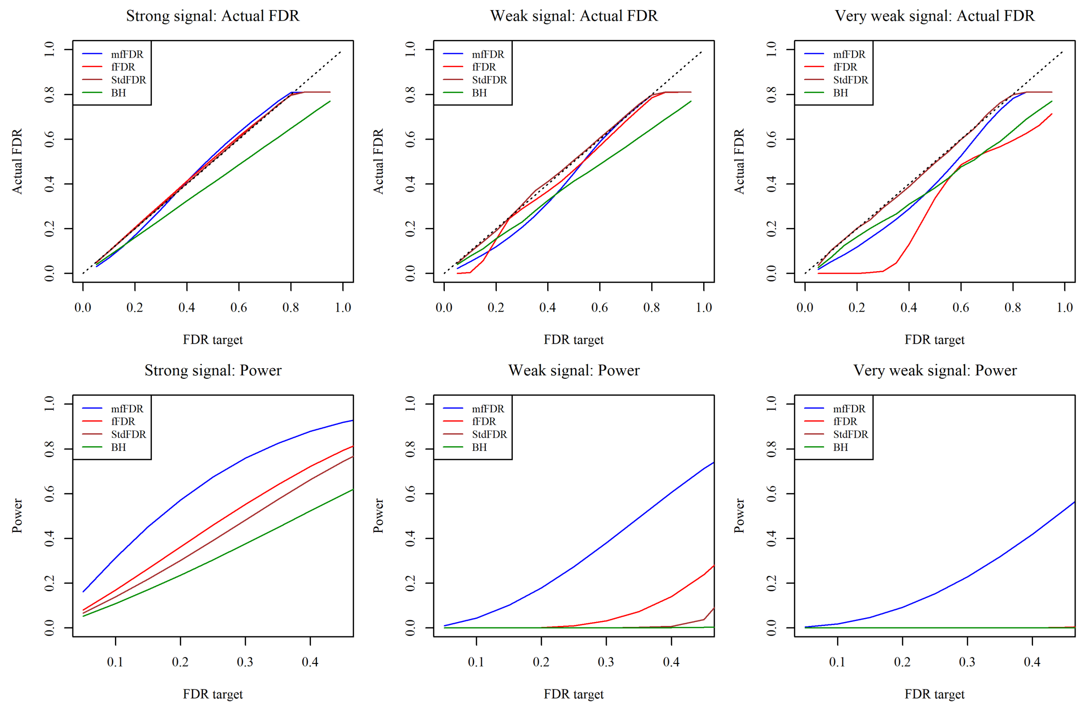
In Figure 1, we show the estimate of the actual FDR and the power of all procedures under the sine (Panel A) and monotonic (Panel B) form of $\pi_0(\boldsymbol{z})$. In each panel, the top three sub-figures exhibit the estimated actual FDR corresponding to the three cases of the signal (strong, weak, and very weak). In each of those sub-figures, each line presents the estimated actual FDR of a procedure at the given FDR targets. Ideally, a procedure perfectly (strictly) controls the FDR at a target if its estimated actual FDR at that target lies on (below) the 45° dotted line. For instance, all procedures, either perfectly or strictly, control for FDR at target 20%. As controlling FDR is a sample property, it is acceptable to observe a point positioning slightly above the dotted line since we estimate the actual FDR over only 1000 iterations. In general, from the sub-figures, we see that all procedures control well for the FDR at any given targets and in all considered cases.

In terms of power, it makes sense to focus on only the FDR targets less than 0.5 (i.e., up to 0.45 in our simulation). Hence, the three bottom sub-figures in each panel present the power of the four methods for only those targets. In all cases, the lines representing the power of the $mfFDR$ are always at the top regardless of the form of $\pi_0(u, v)$ as well as the strength of the signal. In other words, the $mfFDR$ beats its benchmarks in terms of power. Apparently, all procedures have higher power when the signal is strong. In this case, the $mfFDR$ has gaps up to about 21%, 29%, and 37% compared with the $fFDR$, $StdFDR$, and BH procedures, respectively. Those figures are 44%, 67%, and 76% (57%, 62%, and 62%) for the weak (very weak) signal case. At the FDR target of 20%, which will be used later in our main analysis, the gap of $mfFDR$ over $fFDR$ varies from 10% to 20%. Finally, the weak and the very weak signal cases highlight the benefit of using the $mfFDR$ when the data has a low signal-to-noise ratio. While the $fFDR$, $StdFDR$, and BH procedures can hardly detect a single false null hypothesis even at the FDR target of 20% (see the very weak signal case), the $mfFDR$ quickly gains a significant power of more than 10%.

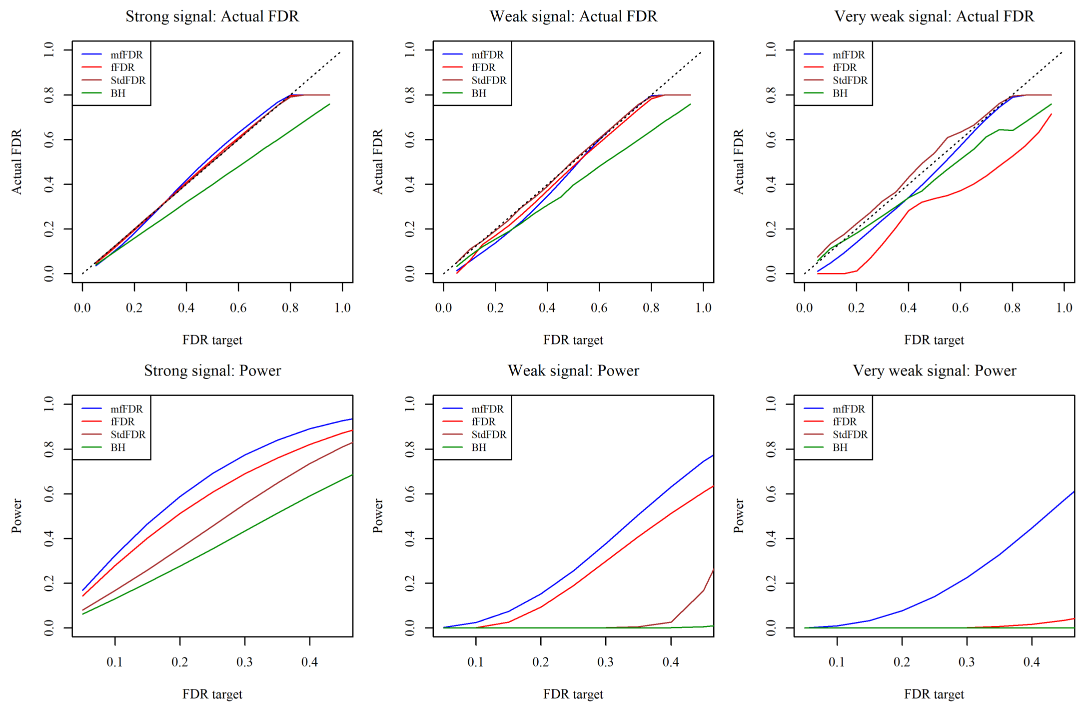
We have developed $mfFDR$, which enables multiple informative covariates to detect

Figure 1: Performance comparison of the $mfFDR$, the $fFDR$, the Standard FDR of Storey ($SdtFDR$) and the FDR procedure in BH. Panel A shows the performance when the $\pi_0(u,v)$ has a sine form, whereas Panel B is the monotonic one. In each panel, the top three sub-figures exhibit the estimated actual FDR, and the three bottom sub-figures present the power.

Panel A: $\pi_0(u,v)$ is a sine function.



Panel B: $\pi_0(u,v)$ is monotonic with respect to each covariate.



false null hypotheses and compare our method to the FDR procedures of Storey (2002) and BH, which do not use covariates. We have shown that the power of the former largely surpasses that of the latter while controlling well for the FDR at any given target. In other words, we can detect more outperforming technical rules under a testing framework that is conditional on multiple information (i.e., $mfFDR$) than the ones that are unconditional or use less information. On the other hand, we also show that when more than one informative covariate that is mutually independent exists, the $mfFDR$ gains remarkably higher power than the $fFDR$, which is the $mfFDR$ with $d = 1$, especially when the signal of the false null hypothesis is weak.

2.3. Robustness checks

Hitherto, we have studied the performance of $mfFDR$ where the covariates are independently drawn from a uniform distribution. In financial data, this is unlikely the case. To consider this issue, in Section IC of the Internet Appendix, we conduct simulations with the use of correlated covariates. We see that the $mfFDR$ controls well FDR at all levels for a pair of covariates with correlation coefficients up to 0.3 and asymptotically does for those with correlation coefficients up to 0.7. The superior power of the $mfFDR$ to other methods remains.

Alongside the concern in the correlation of covariates, the estimation errors of covariates (i.e., the noise in the estimating covariates) also potentially affect the performance of the $mfFDR$. In empirical finance, the covariates will be estimated quantities and are thus subject to estimation errors. To address this concern, we additionally conduct simulations where the covariates used as input contain noise. For the interest of space, we defer the results of this experiment to Section ID of the Internet Appendix. Generally, we find that $mfFDR$ still controls FDR well. Perhaps not surprisingly, its power is lower than the prior case with uncorrelated covariates, as presented in the previous section. Nevertheless, the power of $mfFDR$ is still remarkably higher than that of other methods. The $mfFDR$ method is robust to correlation and estimation biases.

The development of our methodology and the aforementioned simulations are conducted under the assumption that the tests are independent. However, simulations show that the method is valid under the dependent setting that is typical in trading rules. More specifically, for the category of trading rules, the rules with close parameters tend to have highly correlated returns. This leads to a dependency among the testing statistics or $p-$values of the corresponding hypotheses. Thus, we additionally present in Section IE of the Internet Appendix that our method is robust under this type of dependence.

Thus, our simulations cover most concerns about applying our $mfFDR$ method to financial data. We find that the $mfFDR$ performs well in terms of FDR control and power with correlated covariates, under typical dependent $p-$values, and with estimation errors in covariates.

3. Data and strategy universe

3.1. Data

We collect daily spot and 1-month forward exchange rates data against the U.S. dollar for 30 countries, including currencies from nine developed countries (Australian dollar, Canadian dollar, German mark/euro, Japanese yen, New Zealand dollar, Norwegian krone, Swedish krona, Swiss franc, and U.K. pound) and 21 emerging economies (Argentine peso, Brazilian real, Chilean peso, Colombian peso, Czech koruna, Hungarian forint, Indian rupee, Indonesian rupiah, Israeli shekel, Korean won, Mexican peso, Philippine peso, Polish zloty, Romanian new leu, Russian ruble, Singaporean dollar, Slovak koruna, South African rand, Taiwanese dollar, Thai baht, and Turkish lira). The data for developed economies span the period of January 4, 1971, and December 31, 2020. The sample periods for emerging economies start from various dates due to data available on exchange and short-term interest rates. Specifically, Israel starts in January 1978, South Africa in January 1981, Singapore in January 1982, and Taiwan in October 1983; all emerging economies data end in December 2020. Our series on exchange rates and short-term interest rates were kindly offered by the London branch of the asset manager BlackRock and are based on midday quotations in the London market.

Table 1 presents summary statistics of gross currency returns and short-term interest rates of developed and emerging economies. We define the gross currency return as the return from buying a foreign currency unit and holding it for one day, $r_t = ln(s_t/s_{t-1})$, where s_t represents the spot exchange rate on the day t. We define the spot exchange rate as units of U.S. dollars per foreign currency, so an increase of s is associated with an appreciation of the foreign currency. We report the mean, the standard deviation, the minimum and the maximum, and the corresponding sample period. We find that gross returns tend to be more volatile for emerging economies offering higher minimum and maximum values.

Another important aspect of our analysis is the short-term interest rate, as it affects the overall return of the trading strategies, even though technical analysis focuses more

Table 1: Summary Statistics. The table presents summary statistics of gross returns on foreign currencies and short-term interest rates. We report the mean, the volatility, the minimum and the maximum. We also report the sample period for each currency in our sample. We show results for developed and emerging economies.

Country	Gross returns on foreign currencies				Short-term interest rates				Sample Period
	mean(%)	vol	min	max	mean(%)	vol(%)	min(%)	max(%)	
Developed									
Australia	-0.0028	0.0068	-0.1925	0.1073	0.0191	0.0119	0.0001	0.1709	1/4/1971-12/31/2020
Canada	-0.0018	0.0041	-0.0434	0.0505	0.0172	0.0125	0.0005	0.0596	1/4/1971-12/31/2020
Germany/E.U.	-0.0017	0.0058	-0.0421	0.0462	0.0092	0.0078	-0.0020	0.0344	1/4/1971-12/31/2020
Japan	0.0095	0.0063	-0.0626	0.0950	0.0108	0.0124	-0.0006	0.0406	1/4/1971-12/31/2020
New Zealand	-0.0033	0.0073	-0.2050	0.0995	0.0194	0.0111	-0.0007	0.1286	1/4/1971-12/31/2020
Norway	-0.0014	0.0066	-0.0814	0.0646	0.0163	0.0106	0.0023	0.1372	1/4/1971-12/31/2020
Sweden	-0.0036	0.0065	-0.1507	0.0555	0.0171	0.0139	-0.0011	0.2090	1/4/1971-12/31/2020
Switzerland	0.0122	0.0073	-0.0892	0.1267	0.0045	0.0098	-0.0086	0.2276	1/4/1971-12/31/2020
U.K.	-0.0044	0.0058	-0.0848	0.0467	0.0160	0.0112	0.0004	0.0460	1/4/1971-12/31/2020
U.S.	-	-	-	-	0.0128	0.0097	-0.0000	0.0466	1/4/1971-12/31/2020
Emerging									
Argentina	-0.0579	0.0098	-0.3418	0.1712	0.0343	0.0368	0.0035	0.4144	4/1/1991-12/31/2020
Brazil	-0.0238	0.0097	-0.1080	0.1178	0.0428	0.0283	0.0051	0.2792	7/4/1994-12/31/2020
Chile	-0.0070	0.0064	-0.1160	0.1114	0.0014	0.0016	0.0000	0.0214	1/3/1994-12/31/2020
Colombia	-0.0224	0.0062	-0.0593	0.0562	0.0315	0.0240	0.0052	0.0908	1/6/1992-12/31/2020
Czech	0.0040	0.0073	-0.0707	0.0522	0.0104	0.0133	0.0002	0.2628	4/22/1992-12/31/2020
Hungary	-0.0177	0.0080	-0.0842	0.0520	0.0297	0.0227	0.0021	0.0834	6/4/1991-12/31/2020
India	-0.0178	0.0046	-0.1281	0.0376	0.0254	0.0149	0.0002	0.1944	1/3/1986-12/31/2020
Indonesia	-0.0276	0.0130	-0.3576	0.2361	0.0296	0.0242	0.0000	0.2054	1/1/1991-12/31/2020
Israel	-0.0683	0.0059	-0.1725	0.0645	0.0555	0.0962	0.0002	0.6309	1/4/1978-12/31/2020
South Korea	-0.0047	0.0078	-0.1809	0.2012	0.0154	0.0138	0.0012	0.0676	1/3/1992-12/31/2020
Mexico	-0.0349	0.0114	-0.2231	0.2231	0.0403	0.0416	0.0077	0.3387	1/1/1987-12/31/2020
Philippines	-0.0096	0.0044	-0.0860	0.1015	0.0198	0.0165	0.0014	0.1962	1/2/1987-12/31/2020
Poland	-0.0117	0.0079	-0.0715	0.0670	0.0212	0.0198	-0.0004	0.0792	6/4/1993-12/31/2020
Romania	-0.0368	0.0091	-0.3887	0.0953	0.0407	0.0518	0.0006	0.3667	1/2/1997-12/31/2020
Russia	-0.0515	0.0132	-0.3863	0.2779	0.0288	0.0391	0.0022	0.3208	9/1/1994-12/31/2020
South Africa	-0.0286	0.0097	-0.1030	0.1440	0.0289	0.0115	0.0000	0.0588	1/2/1981-12/31/2020
Singapore	0.0043	0.0033	-0.0276	0.0414	0.0053	0.0050	0.0000	0.0181	4/27/1993-12/31/2020
Slovakia	0.0021	0.0063	-0.1097	0.0462	0.0130	0.0164	0.0001	0.2068	1/4/1982-12/31/2020
Taiwan	0.0035	0.0029	-0.0420	0.0430	0.0082	0.0067	0.0004	0.0483	10/4/1983-12/31/2020
Thailand	-0.0022	0.0055	-0.2077	0.0635	0.0102	0.0107	0.0004	0.0664	1/2/1991-12/31/2020
Turkey	-0.1002	0.0109	-0.3348	0.1256	0.0833	0.0643	0.0130	1.0328	2/1/1990-12/31/2020

on exchange rate fluctuations. We convert our annual short-term interest rate (i^a) into daily data $i_t = ln(1+i^a)/360$. In Table 1, we report the mean, standard deviation, minimum, and maximum values of short-term interest rates for both developed and emerging countries. We find that short-term interest rates for developed countries range between 0.45 to 1.94 basis points on average. We also find that short-term interest rates are higher for most of the currencies of the emerging markets category, and they exhibit a particular variation. The highest interest rate in this group is for Turkey, which is 8.33 basis points, and the lowest is for Chile, which offers 0.14 basis points.

3.2. Trading rule universe

Based on past daily spot exchange rates, a trading rule determines positions (long, short, or neutral) that traders should take in the next day. In this study, we assess the universe of trading rules used in Hsu *et al.* (2016), which consists of the following five family trading rules widely used by traders.

The filter trading rules: The rules generate a long (short) position whenever the closing exchange spot rate has risen (fallen) by a given percentage above (below) its most recent high (low). This family rule is generally based on the momentum of the exchange rate, where traders believe the rising (falling) rate continues to rise (or fall). The threshold percentages are set so that the traders are not misled by the small fluctuations.

Moving average trading rules: These rules generate long or short positions based on comparing the closing spot exchange rate to one or three simple moving averages of given different lengths or comparing the moving averages of two different lengths. For example, the simplest moving average rule generates a long (short) position when the spot exchange rate moves up (down) at least a certain percent above (below) the moving average of a specific length.

Relative strength indicators: A relative strength indicator is a popular form of an oscillator, which aims to identify imminent market corrections after rapid exchange rate movements. Generally, the indicator of a given length has values from 0 to 100 and generates an oversold or overbought signal when it crosses a predetermined lower or upper extremity, respectively.

Support-resistance rules: These rules rely on determining a support or resistance level for which the exchange rate appears to have difficulty in penetrating in a previously given number of days and a premise that when the closing exchange rate breaches the level, it will trigger further movement in the same direction.

Channel breakouts: The rules establish time-varying support and resistance levels, forming a trading channel with upper and lower bounds. Once a bound is breached, a long or short position is initiated in a similar way as the support-resistance rules.

Given the described family rules, a number of their variants are generated by varying plausible parameters, the delay time, and the position's holding time. Ultimately, we obtain 2,835 filter rules, 12,870 moving average rules, 600 relative strength indicators, 1,890 support-resistance rules, and 3,000 channel breakout rules, making up 21,195 trading rules in total. Readers are referred to Appendix A in Hsu *et al.* (2016) for the detailed specifications of our technical rules.

4. Measures of predictive ability and profitability

In this study, we distinguish the excess return and net excess return gained by a trading strategy before and after counting for transaction cost, respectively. The excess return from buying one unit of foreign currency (against the U.S. dollar) and holding it for one day is calculated as the summation of returns due to appreciation/depreciation of the foreign currency and the return obtained from lending the money in foreign currency, minus the benchmark return which is the return would gain if the money is deposited in the U.S., that is

$$r_t = \ln(s_t/s_{t-1}) + \ln(1 + i^*_{t-1}) - \ln(1 + i_{t-1}) \tag{6}$$

where s_t and s_{t-1} are spot rates at the midday of the days t and $t-1$, respectively, and i_{t-1} and i^*_{t-1} are daily interest rates on U.S. dollar deposits and foreign currency deposits on the day $t-1$, respectively.

The daily excess return of the trading rule j, earned from day $t-1$ to day t, $R_{j,t}$, is determined as

$$R_{j,t} = S_{j,t-1}.r_t \tag{7}$$

where $S_{j,t-1}$ is a position guided by the trading rule based on past data up to day $t-1$ and taking value in the interval $[-1, 1]$. For most of the trading rules in our universe, $S_{j,t-1}$ takes value $+1$ for the long, (-1) for the short, and 0 for the neutral position on the foreign currency. For some moving average trading rules, where the signal guides for positioning one-third of funds on a long position while the rest is kept in U.S. dollar or on a short position while the rest is kept in foreign currency, $S_{j,t-1}$ takes value $+1/3$ or $-1/3$, respectively.[12]

In this study, we assess the performance of a trading rule j based on its Sharpe ratio. From day T_1 to day T_2, it is defined as

$$SR_j = \frac{\overline{R}_j}{\sigma_j}, \tag{8}$$

[12] Generally, with 1 dollar fund, the excess return for a position guided by a signal $S_{t-1} \in [-1, 1]$ is $S_{t-1}.r_t$. When $S_{t-1} > 0$, we invest S_{t-1} U.S. dollar on the foreign currency and the remaining $1 - S_{t-1}$ on U.S. bond. When $S_{t-1} < 0$, we borrow an amount equivalent to $|S_{t-1}|$ U.S. dollar in the foreign currency and convert it to U.S. dollar, then we invest the total of $|S_{t-1}| + 1$ dollar on U.S. bond.

where $\overline{R}_j$ and σ_j are the mean and standard deviation of excess returns of the trading rule over the mentioned period, that is $\overline{R}_j = \sum_{t=T_1}^{T_2} R_t/N$, $\sigma_j = \sqrt{\sum_{t=T_1}^{T_2}(R_{j,t} - \overline{R}_j)^2/(N-1)}$ where $N = T_2 - T_1 + 1$, respectively.

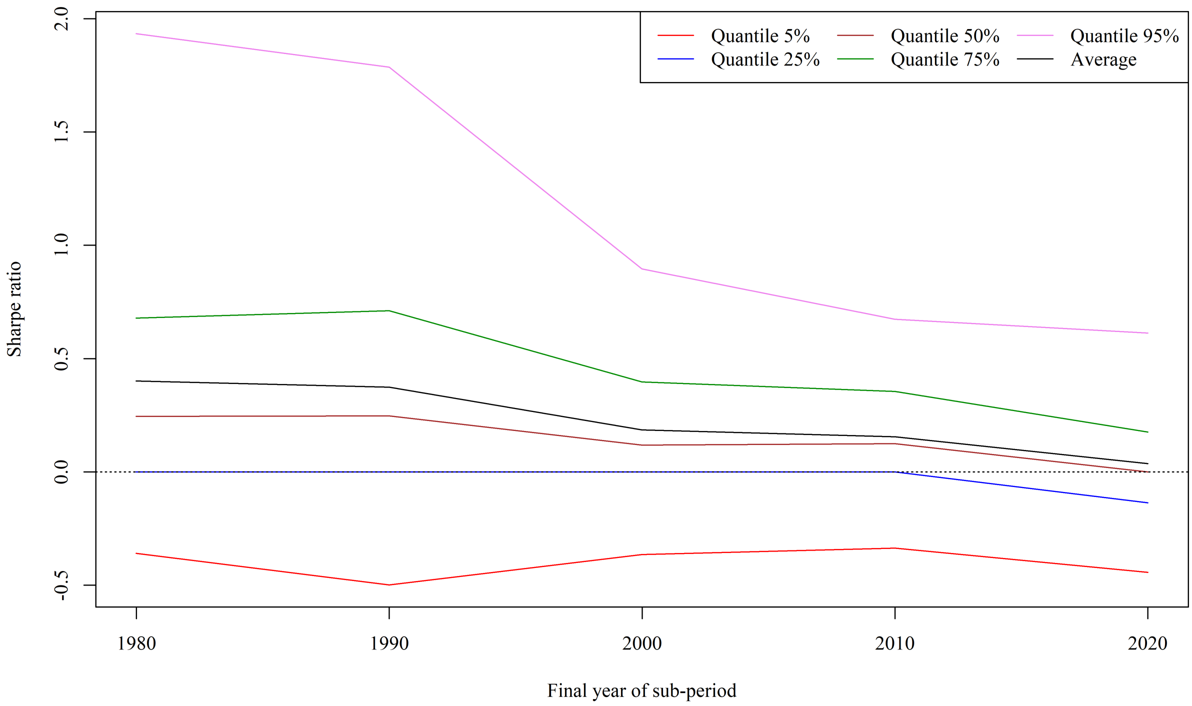
When a trading rule changes its guiding signal, a new or existing position is triggered or closed, and transaction costs occur consequently. The net excess return on a daily basis, therefore, is the daily excess return less the cost caused by the transactions, and the net excess return and Sharpe ratio after transaction costs of a trading rule are calculated accordingly. As in Neely and Weller (2013), we use one-third of the quoted one-month forward rate bid-ask spread in each currency as an estimate of the one-way transaction costs on any particular day. For periods before the forward data is available, we set fixed transaction costs for each period in the same way as in Neely and Weller (2013): for developed country currencies, we set the transaction cost at a flat 5 basis points in the 1970s, 4 basis points in the 1980s and 3 basis points in the 1990s, and for emerging market currencies we set the daily cost at one-third of the average of the first 500 bid-ask observations.

To have the first overview of the performance of the rules, we consider the Sharpe ratio distribution of all rules in each sub-period of 10 years. For each sub-period, all Sharpe ratios of 190,755 to 635,850 trading rules applied on a maximum of 30 currencies (depending on the starting point of currencies) are collected, and then the quantiles and an average of the population are calculated. We plot the average and quantiles of Sharpe ratios based on excess and net returns in panels A and B of Figure 2, respectively. We observe from both graphs that the Sharpe ratios of the 95^{th}, 75^{th}, and 50^{th} quantiles dramatically decline before 2000 and then continue the trend gradually since 2000. We note that the 95^{th} quantile becomes stable in the recent decade. Figure 3 presents similar information for annualized average returns instead of the Sharpe ratio. We observe similar patterns.

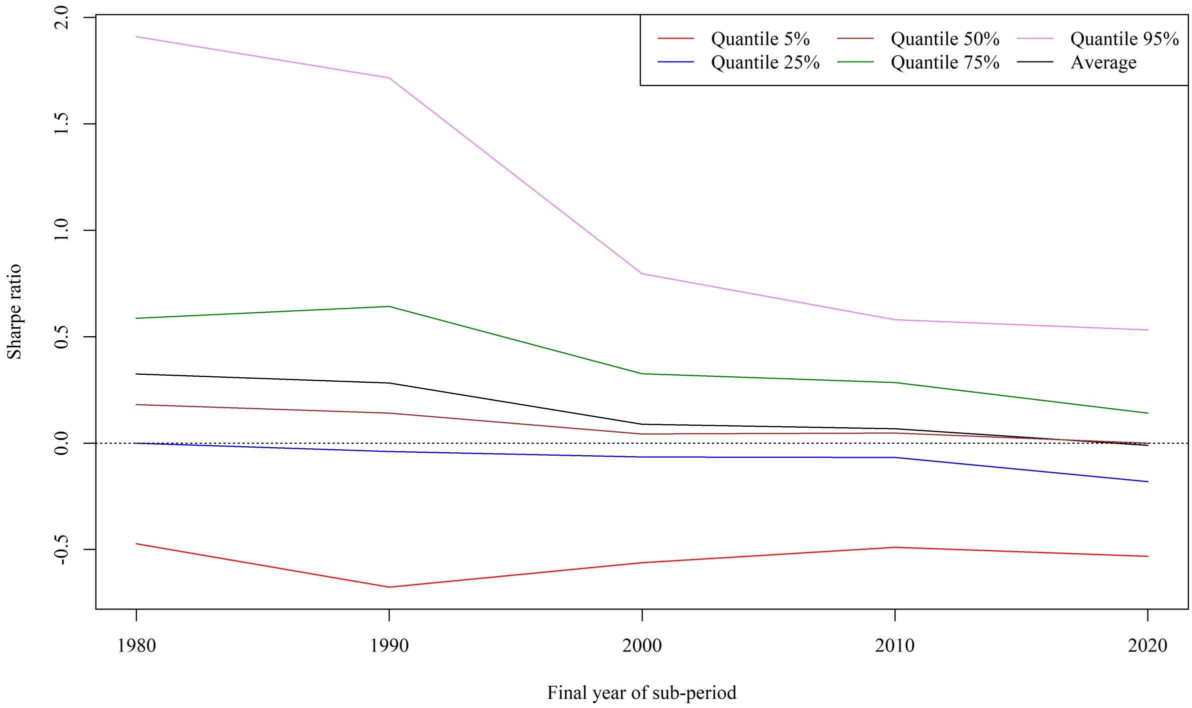
Table 2 provides some statistical metrics of significantly performing rules. It captures important features of trading rules in terms of performance corresponding to each exchange rate after transaction costs. It is noted that the number of significant rules in this particular table is merely based on a conventional $p-$value of 5%. If all rules are independent random signals, one should expect there are roughly 530 significantly out-performing

Figure 2: Evolution of Sharpe ratios (annualized) of the trading rules applying on the basket of 30 currencies in each sub-period. Panel A presents the quantiles 5%, 25%, 50%, 75%, 95%, and the average annualized Sharpe ratio before the transaction cost of strategies calculated based on sub-periods of ten years. Panel B shows those numbers after the transaction cost.

Panel A: Sharpe ratio before transaction cost.



Panel B: Sharpe ratio after transaction cost.

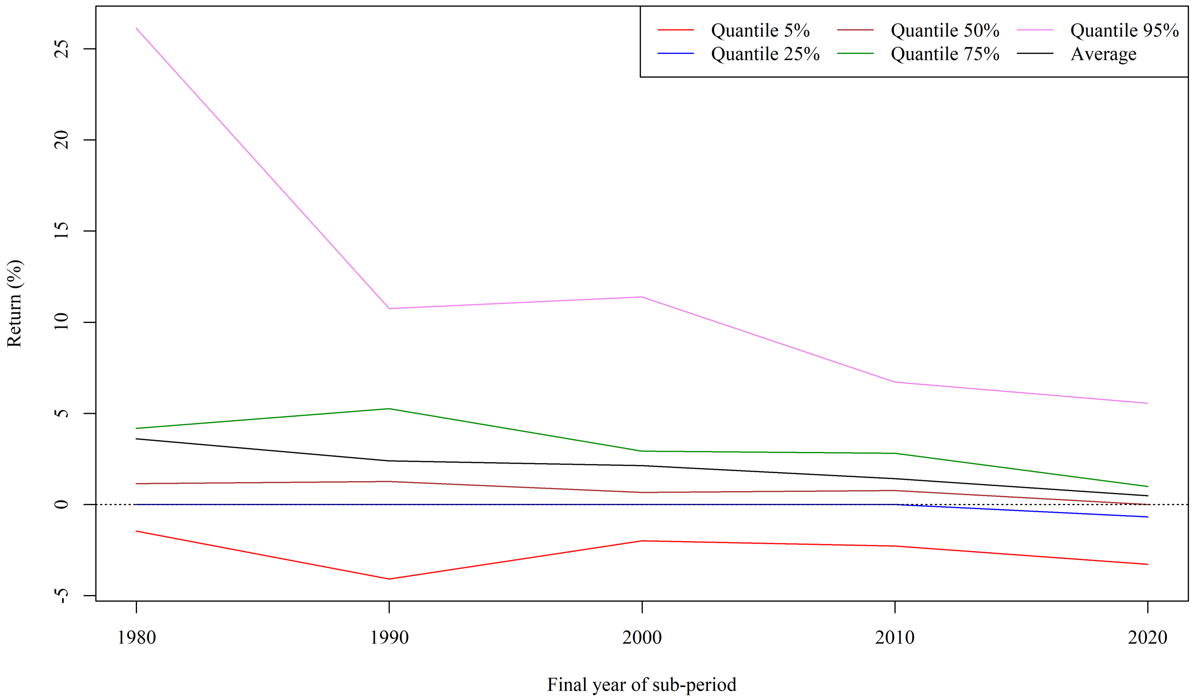


rules for each exchange rate.[13] With both performance measures under study, there are

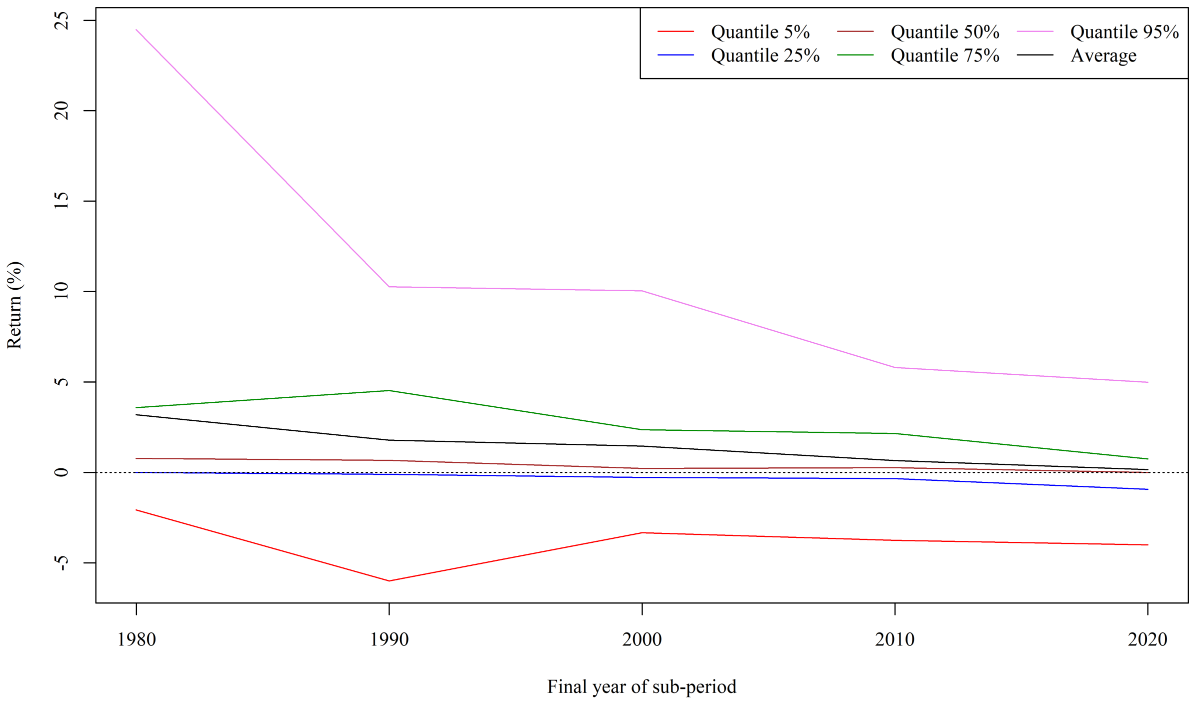
[13]That is, as the $p-$values of true null hypotheses have a uniform distribution, there are approximate

Figure 3: Evolution of returns (annualized, in %) of the trading rules applying on the basket of 30 currencies in each sub-period. Panel A presents the quantiles 5%, 25%, 50%, 75%, 95%, and the annualized average return before the transaction cost of strategies calculated based on sub-periods of ten years. Panel B shows those numbers after the transaction cost.

Panel A: Return before transaction cost.



Panel B: Return after transaction cost.



$21195 \times 0.05/2 \approx 530$ significantly out-performing rules.

Table 2: Summary performance of trading rules in the whole sample period. The table shows the number of significant rules, which are those having bootstrap p-value < 0.05 (under conventional individual tests) and positive estimated performance, the highest performing rule in each currency with its performance. The column "Best rule" presents the class with the technical rules that provide the highest performance metric among all trading rules in a currency in the sample period. See Appendix A of Hsu *et al.* (2016) for the details of the various trading rules and corresponding class codes (F3, MA4, etc). We report two performance metrics - the Sharpe ratio and the mean of net excess return (both are annualized), which is the excess return after transaction cost. "*", "**" and "***" respectively indicate statistical significance of the individual test of the rules at levels of 10%, 5% and 1%.

	Mean net return			Sharpe ratio		
Country	Number of significant rules	Highest return (%)	Best rule	Number of significant rules	Highest ratio	Best rule
Australia	900	5.73***	F3	1155	0.54***	F3
Canada	196	2.67***	MA4	217	0.41***	MA4
Germany/E.U.	8585	5.33***	F3	8536	0.66***	MA5
Japan	7862	6.13***	MA4	7954	0.61***	MA4
New Zealand	6204	6.14***	MA1	6747	0.56***	MA1
Norway	443	4.24***	SR1	449	0.41***	SR1
Sweden	6827	5.99***	MA4	7038	0.59***	MA5
Switzerland	601	5.05***	F3	599	0.48***	CB1
U.K.	2161	4.58***	MA2	2217	0.54***	CB1
Argentina	9254	11.22***	SR1	10973	0.77***	MA5
Colombia	3915	11.04***	MA1	4025	1.17***	MA5
India	1844	4.44***	F3	2287	0.67***	F3
Indonesia	1744	10.71***	MA4	2401	0.56***	MA5
Israel	15271	9.26***	MA1	15452	1.22***	MA5
Philippines	6401	4.74***	SR1	6444	0.69***	MA5
Romania	67	5.53***	F3	118	0.52***	MA5
Russia	3897	15.63**	F1	6246	0.83***	MA5
Slovak	1331	5.69***	MA4	1348	0.58***	MA4
Brazil	4938	10.97***	SR1	6623	0.77***	MA5
Chile	1545	7.16***	SR1	1431	0.85***	MA5
Czech	107	5.87***	MA4	114	0.52***	MA5
Hungary	34	5.39***	SR1	80	0.45***	RSI
Korea	200	8.71***	F3	456	0.7***	F3
Mexico	8	4.66**	SR1	190	0.38***	CB1
Poland	39	6.3***	F3	74	0.51***	F3
Singapore	739	2.45***	CB1	673	0.56***	MA5
South Africa	914	7.82***	CB1	1013	0.53***	CB1
Taiwan	9782	4.3***	MA3	9827	1***	MA5
Thailand	3850	6.85***	MA1	3842	0.87***	MA5
Turkey	13516	16.07***	F3	14568	1.1***	MA5

22 over 30 exchange rates having more than 530 significant rules. This fact suggests that a portion of technical trading rules that are truly profitable could indeed exist. All the best rules of the 30 exchange rates are significant, with five of them gaining more than 10% per annum and four with a Sharpe ratio of at least one. However, we do not know if these rules are truly profitable or just lucky before we implement formal tests to correct

data snooping biases.

5. Empirical results

For each trading rule i, we assess its performance based on a test

$$H_0 : \phi_i = 0, \qquad H_1 : \phi_i \neq 0 \tag{9}$$

where ϕ_i is the Sharpe ratio of the trading rule before transaction cost.[14]

In the follows, we first discuss our covariates in technical trading rules and the information they potentially convey. Based on the $mfFDR$ test incorporating these covariates, we select outperforming rules in each currency separately and in all currencies combined. We then construct rolling portfolios by the selected rules and show their in-sample and out-of-sample performance. If the performance of technical trading rules in the foreign currency market is predictable to some degree, investing in technical rules that are truly profitable in an in-sample period is expected to generate an out-of-sample (OOS) profit. In particular, we demonstrate the OOS performance of the $mfFDR$ in detecting profitable rules using the following procedure: we implement the $mfFDR$ on the group of *positive* estimated SR trading rules to control for FDR at the given target. We name this procedure the multivariate functional false discovery rate "plus" ($mfFDR^+$). Similarly, we consider $fFDR^+$ of Hsu *et al.* (forthcoming) as a benchmark that accommodates one covariate and focuses on positive performance. For the interest of space, we present the results with control of FDR at a target of 20% in the main text. In Section IB of the Internet Appendix, we show that our conclusions are robust for other FDR targets.

5.1. Covariates

We apply within our $mfFDR$ framework a set of informative continuous variables that are derived from technical rules' return series. These covariates reflect the performance persistence, financial risk, activeness, and out-performance compared to a passive strategy of the trading rules under study. The first covariate we consider is the autocorrelation of a trading rule's return (ρ) that reflects its performance persistence. For

[14] In Section IJ of the Internet Appendix, we show that the results are robust when ϕ_i is the strategy's mean return before transaction cost instead.

the remaining covariates, we estimate the following equation for each trading rule in a particular currency using a rolling window:

$$r_{i,t} = \alpha_{i,bh} + \beta_{i,bh} r_{bh,t} + \varepsilon_{bh,t} \tag{10}$$

where the $r_{i,t}$ is the excess return of the trading rule i, $r_{bh,t}$ is the excess return of the buy-and-hold strategy in a currency and $\epsilon_{bh,t}$ is the noise at day t. The $\alpha_{i,bh}$, $\beta_{i,bh}$ and the R–squared of the regression (10), respectively, represent the alpha, the riskiness level, and the variation of the trading rule i compared to the buy-and-hold strategy. We denote the three covariates as α_{bh}, β_{bh} and R^2_{bh} in the following context. The buy-and-hold strategy is an alternative passive strategy that traders have. It has been also used in literature to benchmark the performance of active trading strategies (e.g., see Olson, 2004). The use of the covariates is inspired by the asset pricing models. In this particular case, if a trading rule has a positive α_{bh}, it beats the buy-and-hold strategy after adjusting for risk. The trading rule is might be risky compared to the buy-and-hold one, and this is reflected via $\beta_{i,bh}$. Finally, if the R^2_{bh} is close to 1, the trading rule is just similar to the buy-and-hold strategy, which resembles a passive one. If it is close to 0, the trading rule is very active and completely opposite to the buy-and-hold strategy.

When the target is to select outperforming trading rules from the universe of all trading rules of all currencies (or else, when we examine all currencies together), which consists of $21195 \times 30 = 635850$ rules, we compare a particular trading rule to the average currency excess return factor of Lustig *et al.* (2011), denoted by RX. That is the return of a strategy that invests equally weighted on all 30 currencies in our sample (henceforth, currency market factor). In these cases, instead of using regression (10), the latter three covariates are obtained from the regression model:

$$r_{i,t} = \alpha_{i,mk} + \beta_{i,mk} RX_t + \varepsilon_{mk,t} \tag{11}$$

where the RX_t is the currency market factor on the day t. Analogously, we denote the three new covariates as α_{mk}, β_{mk} and R^2_{mk}, respectively. The explanation of these covariates is similar to the previous three covariates, where the buy-and-hold strategy is applied to the market factor, RX, instead of a particular currency.

In the following analyses, the informativeness of the covariates will be examined, and the benefit of using multiple covariates under the $mfFDR$ framework will be discussed. We first focus on constructing a portfolio on each currency and use ρ α_{bh}, β_{bh} and R^2_{bh}.

Then, we examine all currencies together and consider ρ, α_{mk}, β_{mk}, and R^2_{mk}. Finally, we investigate how profitable trading rules vary across currencies.

5.2. Individual Currencies

As a first exercise, we examine each currency individually. Therefore, the covariates of $mfFDR^+$ are $r^2_{bh}, \alpha_{bh}, \beta_{bh}$ and ρ as described in Section 5.1. We use daily data and construct portfolios in a monthly frequency as follows: for each currency at the end of each month, we utilize the daily data in the most recent 12 months data as the in-sample period to calculate the $p-$value of the test (9), which tests Sharpe ratio before transaction cost against zero, and covariates of the trading rules. We then implement the $mfFDR^+$ (as explained at the beginning of Section 5) to detect outperforming strategies in the in-sample period at the FDR target of 20%. We combine the signals of these outperforming rules to determine the position of that day by neutralizing the opposite ones. For instance, suppose there are 100 trading rules selected as profitable right before a day. Among them, 20 of them indicate a buy signal with a weight $+1$, 30 of them indicate sell signals with a weight of $-1/3$, and the remaining 50 provide neutral signals. After combining, we have $10(= 20 \times (+1) + 30 \times (-1/3))$ long signals out of 100 profitable rules. Thus, the trader takes a long $1/10(= 10/100)$ position in the foreign currency.[15] We follow the signals of these portfolios to determine the position of each trading day in the following month (i.e., the OOS period). We then compare the performance metrics of these portfolios to those based on rules selected by FDR^+ and $fFDR^+$ with one covariate. The truly profitable rules are detected based on the use of the excess return (before transaction costs) and we mainly assess the OOS performance of the portfolio's net excess return (after transaction costs).

Table 3 provides the quantiles of pairwise correlation coefficients of input covariates and those after combining all coefficients of the pairs as a pool. Those numbers capture an overview of the dependencies among the input covariates. The final row indicates that

[15]This approach is based on the idea of the $1/N$ portfolio strategy, where we invest equally funds into each of the selected rules. We choose this method for three reasons. First, DeMiguel *et al.* (2007) show that such an approach is hard to be beaten by more sophisticated approaches (e.g., approaches weight funds differently on selected rules). Second, it reflects directly the performance of its components (e.g., the selected rules). Lastly, the funds allocated to opposite signals should be neutralized to avoid transaction costs. In a different context, Burnside *et al.* (2011) follow a similar approach to construct carry trade and momentum portfolios, and Filippou *et al.* (2018) build an equally-weighted portfolio of a trading rule that is based on political risk.

a majority of 98% coefficients are in $[-0.69, 0.59]$ for which, as shown in Section IC of the Internet Appendix, the $mfFDR$ controls very well for any FDR targets. We emphasize that there are only 1% of the coefficients having an absolute value from or above 0.7.

Table 3: The pairwise correlation of covariates. The table presents the quantiles of the correlation coefficients (calculated monthly with the use of one-year in-sample data) of six covariate pairs which are combinations of the four covariates: $r^2_{bh}, \alpha_{bh}, \beta_{bh}$ and ρ. The final row shows the numbers for the set of all correlated pairs.

Covariate pairs	min	1%	5%	25%	50%	75%	95%	99%	max
r^2_{bh}, α_{bh}	-0.73	-0.44	-0.32	-0.12	0.05	0.21	0.40	0.54	0.70
r^2_{bh}, β_{bh}	-0.82	-0.75	-0.68	-0.51	-0.25	0.18	0.59	0.70	0.81
r^2_{bh}, ρ	-0.92	-0.68	-0.40	-0.09	0.02	0.12	0.28	0.47	0.91
α_{bh}, β_{bh}	-0.81	-0.54	-0.39	-0.20	-0.06	0.09	0.30	0.45	0.72
α_{bh}, ρ	-0.86	-0.45	-0.22	-0.10	-0.04	0.02	0.12	0.21	0.72
β_{bh}, ρ	-0.52	-0.32	-0.19	-0.06	0.01	0.08	0.22	0.34	0.58
All pairs	-0.92	-0.69	-0.51	-0.15	-0.02	0.10	0.36	0.59	0.91

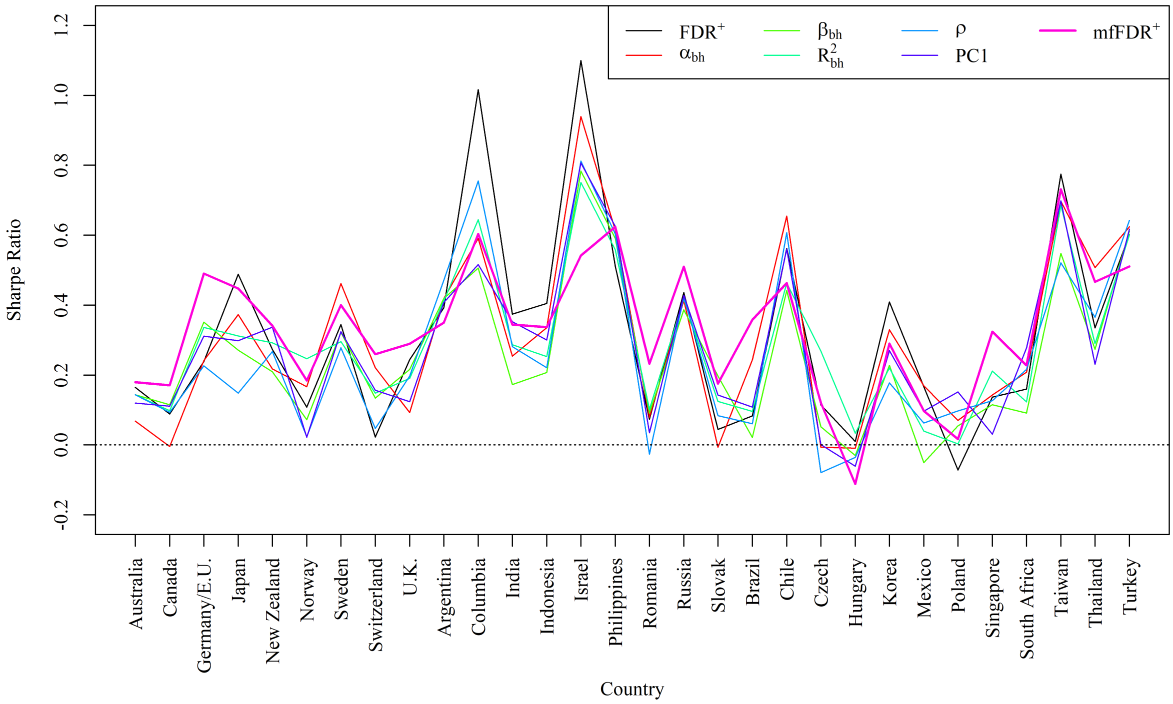
In Table 4, we present the average number of trading rules selected by FDR^+ and $mfFDR^+$, controlling for FDR at 20%, at the end of each month in different sample periods. As mentioned earlier, we measure technical trading rules' in-sample performance using the Sharpe ratio based on the portfolios' excess returns. We find that controlling at the same level of FDR, the $mfFDR^+$ detects more outperforming rules than FDR^+ does. This finding strongly holds for all currencies both in the whole period and in the sub-periods.

In Figure 4 we depict the OOS performance of the portfolios in terms of Sharpe ratios before (Panel A) and after (Panel B) transaction cost. The Panel A shows that all portfolios do well and they are indistinguishable. However, when the transaction cost is counted, the FDR^+-based portfolio becomes inferior compared to others, and the $mfFDR^+$-based portfolios almost always on the top. The details of the metric are reported in Tables 5 and 6.

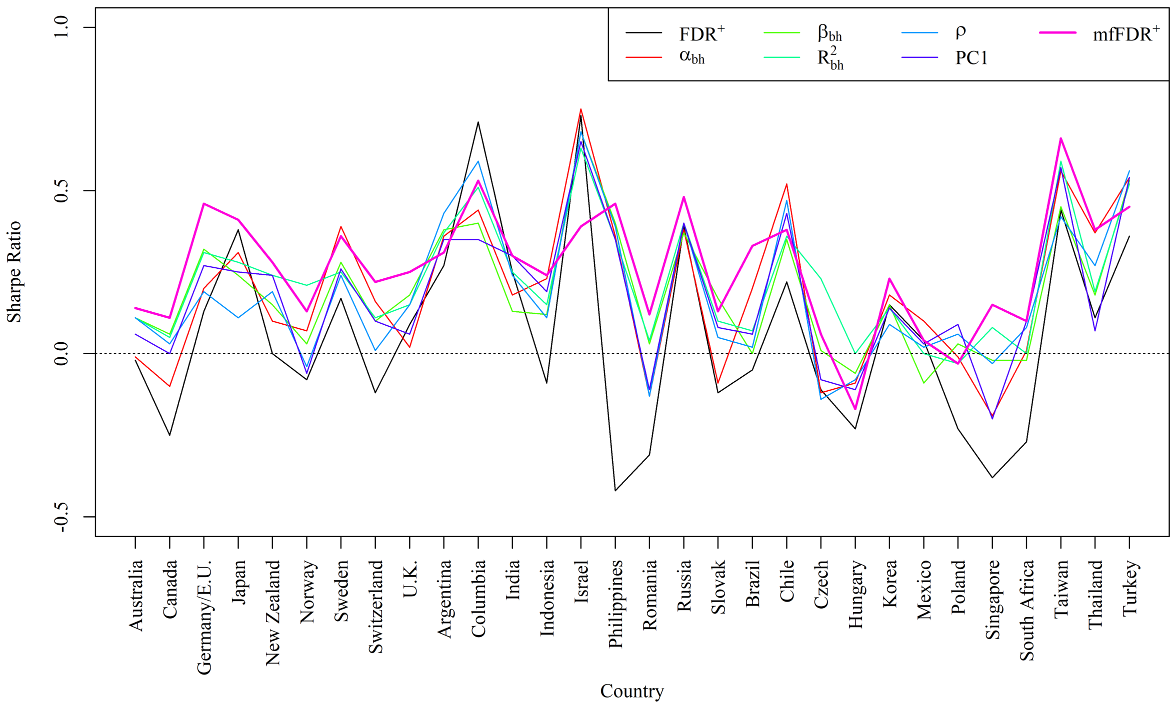
In Table 5, we present the OOS performance of the portfolios, including the FDR^+, $fFDR^+$, and $mfFDR^+$ in terms of Sharpe ratio based on the portfolios' excess return before transaction costs. We also study the performance of $fFDR^+$ having as covariate the first principal component (PC1) of the four aforementioned covariates. We find that most portfolios are able to produce positive Sharpe ratios. It is noteworthy that the $mfFDR^+$ surpasses the other two methods in terms of Sharpe ratios. On the other hand, the performance of the $fFDR^+$ portfolio with PC1 as a covariate does not sur-

Figure 4: OOS performance of portfolios. The graphs show the OOS Sharpe ratios before (Panel A) and after (Panel B) transaction cost of the portfolios FDR^+, $fFDR$ with use each of the four covariates $\alpha_{bh}, \beta_{bh}, R^2_{bh}, \rho$, the PC1 of the four, and the $mfFDR^+$ with use of all the four. The input $p-$value for the procedures is obtained by the test (9), which tests Sharpe ratio before transaction cost against zero.

Panel A: Sharpe ratio of portfolios before transaction cost



Panel B: Sharpe ratio of portfolios after transaction cost



pass that of $fFDR^{+}$-based portfolios with the original individual covariates. The fact that the $mfFDR^{+}$ portfolio based on four covariates outperforms the $fFDR^{+}$ portfolio based on PC1 suggests that the former is effective in extracting useful new information from different covariates in detecting outperforming rules. This is possible because the $mfFDR^{+}$ method enables us to test conditional null hypotheses and allows investors to evaluate technical rules' performance in a timely matter.

Table 6 presents the portfolios' excess returns after transaction costs and shows a similar pattern. On average, the Sharpe ratio of the FDR^{+} portfolio is now indistinguishable from zero. The final row of the table presents the $t-$statistic comparing the mean of Sharpe ratios of the $fFDR^{+}$ and $mfFDR^{+}$ portfolios across currencies to that of the FDR^{+}. All covariate-augmented portfolios statistically significantly outperform the FDR^{+}. The $mfFDR^{+}$ using all four covariates performs the best. This portfolio gains a positive Sharpe ratio for all currencies except the Hungarian forint and Polish zloty. 16 of these positive Sharpe ratios are statistically significant. The implications from Tables 5 and 6 are threefold. First, the covariates are informative in a way that they help us to discover more out-performing rules that are able to deliver profits after trading costs are accounted for. Second, considering more covariates indeed enhances the OOS performance of technical trading rules portfolio, which justifies the advantage of $mfFDR^{+}$ and the importance of conditional hypotheses. Third, the fact that our method outperforms the $fFDR^{+}$ suggests that the information contents of four covariates may be non-linear and cannot be summarized by principal components, which again supports the strength of $mfFDR^{+}$.

Table 4: Empirical power comparison. The table presents the average number of trading rules in portfolios based on the FDR^+ and $mfFDR^+$ at the beginning of each month, controlling for FDR at 20%. The rules are selected based on the the two procedures using as inputs the $p-$value of the test (9), which tests Sharpe ratio before transaction cost against zero, and the four covariates ($\alpha_{bh}, \beta_{bh}, \rho$ and R^2_{bh}). The first five columns report the numbers in five sub-periods, while the last column shows those numbers across the months from the first time forming portfolios till December 2020.

Country	Period											
	1973-1980		1981-1990		1991-2000		2001-2010		2011-2020		1973-2020	
	FDR^+	$mfFDR^+$	FDR^+	$mfFDR^+$	FDR^+	$mfFDR^+$	FDR^+	$mfFDR^+$	FDR^+	$mfFDR^+$	FDR^+	$mfFDR^+$
Australia	4138	8710	2451	8804	359	6615	1104	8095	114	6191	1530	7636
Canada	4603	8089	1349	6810	335	5826	370	6497	1099	5751	1433	6533
Germany/E.U.	2835	8244	3865	10832	496	7225	655	7707	864	6012	1697	7984
Japan	6121	11179	3453	9309	1173	8791	25	6211	724	6499	2145	8282
New Zealand	3227	10921	4483	10066	1141	8163	2232	8298	137	5047	2195	8388
Norway	2218	7348	3914	10194	127	6278	1243	6592	990	5615	1675	7192
Sweden	2485	9306	3050	10704	2142	8036	688	8146	1288	6079	1906	8410
Switzerland	3255	9296	1531	8960	1053	6513	259	6299	353	4684	1210	7056
U.K.	6303	10826	2889	9646	321	4828	22	6389	349	5413	1803	7280
Argentina					1674	2132	1565	3390	8213	10230	4011	5524
Colombia					3711	9102	3723	9918	1821	6589	3010	8463
India					1305	9112	3399	9262	387	5249	1716	7774
Indonesia			20	3936	988	5291	2801	9641	3084	8593	2090	7493
Israel	15551	3570	9365	5464	1882	8918	1542	7876	620	6223	3658	7021
Philippines			10529	13022	4671	9466	3644	8230	357	6962	3348	8509
Romania					11420	14408	1915	5954	896	6175	2312	6823
Russia					6787	11236	3806	9383	2597	8139	3836	9198
Slovak					345	5616	1643	8526	890	5813	1061	6818
Brazil					4185	5536	1350	10093	650	8451	1583	8580
Chile					5425	10570	3694	8572	210	6692	2635	8214
Czech					16	5919	2646	8221	1670	5788	1619	6726
Hungary					331	6843	473	6934	133	4711	310	6096
Korea					3098	9697	1899	8850	124	5346	1547	7761
Mexico			63	1388	628	5454	905	5326	147	6579	526	5518
Poland					2427	7506	1617	6950	359	5365	1298	6447
Singapore			1060	6313	2422	6740	1014	6655	1068	5400	1415	6268
South Africa			2045	9723	2817	8579	1661	8192	187	6635	1649	8199
Taiwan			8678	9708	4078	9220	2860	8490	116	6551	3278	8317
Thailand					1247	7120	2917	9792	2483	7851	2284	8330
Turkey					13393	11654	1477	8028	1935	9302	5334	9596

Table 5: Gross Sharpe ratios of portfolios. The table presents the OOS annualized Sharpe ratios before transaction costs of seven portfolios based on the FDR^+, the $fFDR^+$ and $mfFDR^+$ controlling FDR at 20%. The input $p-$value for the procedures is obtained by the test (9), which tests Sharpe ratio before transaction cost against zero. For the $fFDR^+$, we first consider four covariates: α_{bh}, β_{bh}, R^2_{bh}, ρ and the first principal component of the four mentioned covariates. For the $mfFDR^+$ we study $d = 4$ with all four covariates. The last row is the average Sharpe ratio across currencies. The numbers in parentheses are the corresponding $p-$values. "*", "**" and "***" respectively indicate statistical significance at levels of $10\%, 5\%$ and 1%.

Country	FDR^+	$fFDR^+$					$mfFDR^+$
		α_{bh}	β_{bh}	R^2_{bh}	ρ	PC1	
Australia	0.16 (0.15)	0.07 (0.60)	0.14 (0.30)	0.14 (0.30)	0.14 (0.30)	0.12 (0.34)	0.18 (0.17)
Canada	0.09 (0.53)	0.00 (0.94)	0.11 (0.44)	0.10 (0.52)	0.09 (0.59)	0.11 (0.46)	0.17 (0.23)
Germany/E.U.	0.24 (0.06)*	0.24 (0.11)	0.35 (0.01)***	0.34 (0.01)***	0.23 (0.13)	0.31 (0.03)**	0.49 (0.00)***
Japan	0.49 (0.00)***	0.37 (0.01)***	0.27 (0.07)*	0.31 (0.03)**	0.15 (0.29)	0.30 (0.03)**	0.45 (0.00)***
New Zealand	0.27 (0.00)***	0.22 (0.08)*	0.21 (0.07)*	0.29 (0.01)***	0.27 (0.02)**	0.34 (0.00)***	0.34 (0.00)***
Norway	0.11 (0.44)	0.17 (0.22)	0.07 (0.61)	0.25 (0.08)*	0.02 (0.87)	0.02 (0.90)	0.18 (0.17)
Sweden	0.34 (0.00)***	0.46 (0.00)***	0.32 (0.02)**	0.30 (0.02)**	0.28 (0.03)**	0.32 (0.02)**	0.40 (0.00)***
Switzerland	0.02 (0.89)	0.22 (0.13)	0.13 (0.35)	0.15 (0.27)	0.05 (0.70)	0.16 (0.28)	0.26 (0.06)*
U.K.	0.24 (0.09)*	0.09 (0.52)	0.22 (0.13)	0.19 (0.20)	0.20 (0.17)	0.12 (0.37)	0.29 (0.05)**
Argentina	0.39 (0.12)	0.42 (0.01)***	0.42 (0.01)***	0.41 (0.02)**	0.47 (0.01)***	0.41 (0.02)**	0.35 (0.06)*
Colombia	1.02 (0.00)***	0.59 (0.00)***	0.51 (0.02)**	0.64 (0.00)***	0.75 (0.00)***	0.52 (0.02)**	0.60 (0.00)***
India	0.37 (0.01)***	0.25 (0.15)	0.17 (0.32)	0.29 (0.08)*	0.28 (0.09)*	0.35 (0.04)**	0.34 (0.03)**
Indonesia	0.40 (0.04)**	0.34 (0.02)**	0.21 (0.17)	0.25 (0.08)*	0.22 (0.16)	0.30 (0.03)**	0.34 (0.02)**
Israel	1.10 (0.00)***	0.94 (0.00)***	0.78 (0.00)***	0.75 (0.00)***	0.81 (0.00)***	0.81 (0.00)***	0.54 (0.00)***
Philippines	0.51 (0.00)***	0.61 (0.00)***	0.59 (0.00)***	0.56 (0.00)***	0.60 (0.00)***	0.63 (0.00)***	0.62 (0.00)***
Romania	0.07 (0.72)	0.08 (0.65)	0.09 (0.62)	0.10 (0.58)	-0.03 (0.94)	0.03 (0.88)	0.23 (0.19)
Russia	0.44 (0.08)*	0.41 (0.06)*	0.39 (0.06)*	0.42 (0.05)**	0.42 (0.06)*	0.43 (0.05)**	0.51 (0.01)***
Slovak	0.04 (0.77)	-0.01 (0.98)	0.20 (0.31)	0.12 (0.51)	0.08 (0.67)	0.14 (0.46)	0.18 (0.35)
Brazil	0.08 (0.65)	0.24 (0.20)	0.02 (0.91)	0.10 (0.63)	0.06 (0.72)	0.11 (0.60)	0.36 (0.07)*
Chile	0.56 (0.01)***	0.65 (0.00)***	0.44 (0.02)**	0.46 (0.01)***	0.61 (0.01)***	0.56 (0.01)***	0.46 (0.02)**
Czech	0.11 (0.51)	-0.01 (0.98)	0.05 (0.72)	0.27 (0.11)	-0.08 (0.68)	0.00 (0.93)	0.12 (0.46)
Hungary	0.01 (0.89)	-0.01 (1.00)	-0.03 (0.89)	0.03 (0.85)	-0.04 (0.90)	-0.06 (0.76)	-0.11 (0.60)
Korea	0.41 (0.07)*	0.33 (0.09)*	0.23 (0.22)	0.22 (0.23)	0.18 (0.34)	0.27 (0.16)	0.29 (0.16)
Mexico	0.17 (0.20)	0.17 (0.25)	-0.05 (0.76)	0.04 (0.75)	0.06 (0.65)	0.10 (0.52)	0.10 (0.49)
Poland	-0.07 (0.70)	0.07 (0.67)	0.05 (0.76)	0.00 (0.98)	0.10 (0.60)	0.15 (0.45)	0.02 (0.92)
Singapore	0.14 (0.38)	0.14 (0.40)	0.12 (0.53)	0.21 (0.20)	0.13 (0.45)	0.03 (0.88)	0.32 (0.02)**
South Africa	0.16 (0.33)	0.21 (0.18)	0.09 (0.53)	0.12 (0.41)	0.22 (0.14)	0.28 (0.07)*	0.23 (0.14)
Taiwan	0.77 (0.00)***	0.70 (0.00)***	0.55 (0.00)***	0.69 (0.00)***	0.52 (0.00)***	0.70 (0.00)***	0.73 (0.00)***
Thailand	0.34 (0.07)*	0.51 (0.01)***	0.27 (0.13)	0.29 (0.12)	0.37 (0.04)**	0.23 (0.22)	0.47 (0.02)**
Turkey	0.60 (0.00)***	0.62 (0.00)***	0.61 (0.00)***	0.60 (0.00)***	0.64 (0.00)***	0.62 (0.00)***	0.51 (0.00)***
Average	0.32	0.30	0.25	0.29	0.26	0.28	0.33

Table 6: Net Sharpe ratios of portfolios. The table presents the annualized Sharpe ratio after transaction cost of seven portfolios based on the FDR^+, the $fFDR^+$ and $mfFDR^+$ controlling FDR at 20%. The input $p-$value for the procedures is obtained by the test (9), which tests Sharpe ratio before transaction cost against zero. For the $fFDR^+$, we first consider four covariates: α_{bh}, β_{bh}, R^2_{bh}, ρ and the first principal component of the four mentioned covariates. For the $mfFDR^+$ we study $d = 4$ with all four covariates. The second last row is the average of the Sharpe ratio across currencies. The last row is the $t-$statistic of the test comparing the (paired) means of the portfolios $fFDR^+/mfFDR^+$ to the portfolios FDR^+. The numbers in parentheses are the corresponding $p-$values. "*", "**" and "***" respectively indicate statistical significance at levels of $10\%, 5\%$ and 1%.

Country	FDR^+	$fFDR^+$					$mfFDR^+$
		α_{bh}	β_{bh}	R^2_{bh}	ρ	PC1	
Australia	-0.02 (0.42)	-0.01 (0.97)	0.11 (0.42)	0.11 (0.42)	0.11 (0.42)	0.06 (0.63)	0.14 (0.29)
Canada	-0.25 (0.80)	-0.10 (0.50)	0.06 (0.69)	0.05 (0.80)	0.03 (0.89)	0.00 (0.98)	0.11 (0.46)
Germany/E.U.	0.13 (0.02)**	0.20 (0.17)	0.32 (0.02)**	0.31 (0.02)**	0.19 (0.19)	0.27 (0.07)*	0.46 (0.00)***
Japan	0.38 (0.06)*	0.31 (0.02)**	0.24 (0.11)	0.28 (0.06)*	0.11 (0.43)	0.25 (0.09)*	0.41 (0.00)***
New Zealand	0.00 (0.05)**	0.10 (0.43)	0.15 (0.24)	0.24 (0.05)**	0.19 (0.11)	0.24 (0.05)**	0.28 (0.02)**
Norway	-0.08 (0.14)	0.07 (0.64)	0.03 (0.88)	0.21 (0.14)	-0.04 (0.78)	-0.06 (0.64)	0.13 (0.36)
Sweden	0.17 (0.04)**	0.39 (0.01)***	0.28 (0.05)**	0.25 (0.04)**	0.24 (0.07)*	0.26 (0.05)**	0.36 (0.01)***
Switzerland	-0.12 (0.41)	0.16 (0.26)	0.10 (0.44)	0.11 (0.41)	0.01 (0.91)	0.10 (0.47)	0.22 (0.12)
U.K.	0.09 (0.30)	0.02 (0.87)	0.18 (0.22)	0.15 (0.30)	0.15 (0.29)	0.06 (0.64)	0.25 (0.08)*
Argentina	0.27 (0.03)**	0.36 (0.04)**	0.38 (0.02)**	0.37 (0.03)**	0.43 (0.01)***	0.35 (0.04)**	0.31 (0.11)
Colombia	0.71 (0.01)***	0.44 (0.04)**	0.40 (0.07)*	0.51 (0.01)***	0.59 (0.00)***	0.35 (0.10)*	0.53 (0.01)***
India	0.25 (0.12)	0.18 (0.29)	0.13 (0.45)	0.25 (0.12)	0.24 (0.15)	0.30 (0.07)*	0.30 (0.07)*
Indonesia	-0.09 (0.29)	0.23 (0.12)	0.12 (0.40)	0.15 (0.29)	0.11 (0.47)	0.19 (0.18)	0.24 (0.09)*
Israel	0.73 (0.00)***	0.75 (0.00)***	0.68 (0.00)***	0.63 (0.00)***	0.68 (0.00)***	0.65 (0.00)***	0.39 (0.01)***
Philippines	-0.42 (0.05)**	0.36 (0.05)**	0.40 (0.03)**	0.35 (0.05)**	0.39 (0.03)**	0.35 (0.07)*	0.46 (0.01)***
Romania	-0.31 (0.81)	-0.13 (0.53)	0.03 (0.86)	0.04 (0.81)	-0.13 (0.58)	-0.11 (0.60)	0.12 (0.51)
Russia	0.39 (0.07)*	0.38 (0.07)*	0.37 (0.07)*	0.40 (0.07)*	0.40 (0.07)*	0.40 (0.05)**	0.48 (0.01)***
Slovak	-0.12 (0.62)	-0.09 (0.68)	0.17 (0.38)	0.10 (0.62)	0.05 (0.76)	0.08 (0.63)	0.13 (0.48)
Brazil	-0.05 (0.71)	0.20 (0.29)	0.00 (0.98)	0.07 (0.71)	0.02 (0.94)	0.06 (0.78)	0.33 (0.10)*
Chile	0.22 (0.07)*	0.52 (0.01)***	0.35 (0.07)*	0.36 (0.07)*	0.47 (0.03)**	0.43 (0.05)**	0.38 (0.06)*
Czech	-0.11 (0.17)	-0.12 (0.55)	0.01 (0.89)	0.23 (0.17)	-0.14 (0.42)	-0.08 (0.70)	0.06 (0.65)
Hungary	-0.23 (0.97)	-0.09 (0.67)	-0.06 (0.75)	0.00 (0.97)	-0.08 (0.67)	-0.11 (0.55)	-0.17 (0.41)
Korea	0.15 (0.41)	0.18 (0.33)	0.15 (0.39)	0.14 (0.41)	0.09 (0.62)	0.14 (0.42)	0.23 (0.25)
Mexico	0.04 (0.97)	0.10 (0.49)	-0.09 (0.58)	0.00 (0.97)	0.02 (0.86)	0.03 (0.85)	0.04 (0.73)
Poland	-0.23 (0.89)	-0.01 (0.98)	0.03 (0.87)	-0.03 (0.89)	0.06 (0.77)	0.09 (0.64)	-0.03 (0.89)
Singapore	-0.38 (0.67)	-0.19 (0.25)	-0.02 (0.93)	0.08 (0.67)	-0.03 (0.87)	-0.20 (0.21)	0.15 (0.32)
South Africa	-0.27 (0.99)	0.01 (0.93)	-0.02 (0.93)	0.00 (0.99)	0.08 (0.62)	0.10 (0.54)	0.10 (0.49)
Taiwan	0.44 (0.00)***	0.56 (0.00)***	0.45 (0.00)***	0.59 (0.00)***	0.42 (0.01)***	0.57 (0.00)***	0.66 (0.00)***
Thailand	0.11 (0.30)	0.37 (0.05)**	0.18 (0.35)	0.19 (0.30)	0.27 (0.13)	0.07 (0.75)	0.38 (0.04)**
Turkey	0.36 (0.00)***	0.54 (0.00)***	0.53 (0.00)***	0.52 (0.00)***	0.56 (0.00)***	0.54 (0.00)***	0.45 (0.00)***
Average	0.06	0.19	0.19	0.22	0.18	0.18	0.26
t-stats		3.9	3.5	4.8	3.6	3.4	5.1

5.3. Baskets of Currencies

In the prior section, we select out-performing trading rules separately in each currency. The profitability of technical trading rules is found to vary by currency and time period. It, therefore, might be beneficial for a trader to assess the performance of trading rules across currencies simultaneously. In doing so, the trader will be able to diversify and/or switch her/his funds across currencies. Thus, we assume that the trader can trade all currencies and select, with control of luck, out-performing rules in a pool of technical trading rules across currencies. Depending on the availability of the data, the pool consists of 190,755 ($= 9 \times 21,195$) to 635,850 ($= 30 \times 21,195$) trading rules. We construct a monthly rolling portfolio as in the previous section with the use of the new trading rule set. The four input covariates of the $mfFDR^+$ now are ρ, α_{mk}, β_{mk} and r^2_{mk} as described in Section 5.1. The quantiles of correlation coefficients of covariate pairs are shown in Table 7. Here, again, a majority of about 98% of the coefficients are in $[-0.73, 0.52]$.

Table 7: Summary of correlation coefficients of covariate pairs: the case of all currencies. The table presents the quantiles of the correlated coefficient of six covariate pairs, which are combinations of the four covariates: $r^2_{mk}, \alpha_{mk}, \beta_{mk}$ and ρ. The final row shows the numbers for the set of all correlated pairs.

Covariate pairs	min	1%	5%	25%	50%	75%	95%	99%	max
r^2_{mk}, α_{mk}	-0.44	-0.38	-0.27	-0.08	0.02	0.10	0.37	0.67	0.71
r^2_{mk}, β_{mk}	-0.81	-0.76	-0.70	-0.52	-0.23	0.12	0.50	0.54	0.55
r^2_{mk}, ρ	-0.26	-0.24	-0.18	-0.08	-0.02	0.07	0.25	0.30	0.36
α_{mk}, β_{mk}	-0.93	-0.90	-0.68	-0.31	-0.16	-0.03	0.31	0.53	0.63
$\alpha_{mk}\rho$	-0.20	-0.15	-0.11	-0.03	0.04	0.12	0.26	0.33	0.36
β_{mk}, ρ	-0.40	-0.36	-0.27	-0.11	-0.03	0.05	0.15	0.20	0.21
All pairs	-0.93	-0.73	-0.55	-0.16	-0.03	0.07	0.32	0.52	0.71

At the beginning of each month, the trading rules selected by the $mfFDR^+$ procedure (control for FDR at 20%) are pooled together based on currency. For instance, suppose the $mfFDR^+$ identifies a set of out-performing rules that contains only rules from two currencies, namely A and B. The numbers of these out-performing rules in the two currencies are k_A and k_B, respectively. The wealth is then split into k ($= k_A + k_B$) portions, and k_A (k_B) of them are invested on the corresponding rules in currency A (B). We then calculate the performance of this portfolio for each month.

The performance of the $mfFDR^+$-based portfolios in terms of annualized Sharpe ratios and mean returns before and after transaction costs are exhibited in Table 8. The first row of each table reveals the performances over the whole sample period from 1973 to the end of 2020. We find impressive Sharpe ratios with values of about 1.06 and 0.95 before and after transaction costs, respectively. The one-way break-even point (reported in the rightmost column), which is a fixed transaction cost at which the Sharpe ratio of the $mfFDR^+$ portfolio is set to zero, is as high as 60 basis points. The net Sharpe ratio indicates that in the whole sample period, the $mfFDR^+$ portfolio that consists of all currencies performs much better than any considered FDR^+-, $fFDR^+$-, and $mfFDR^+$-based portfolios that are traded on a single currency. The next rows of the table break down the performance of the portfolio into sub-periods of roughly ten years. While we find deterioration in Sharpe ratios over time, the $mfFDR^+$ portfolio is still fairly profitable even in the most recent decade (2011-2020). For instance, the net Sharpe ratio is 0.21 with a break-even cost of 14 basis points.

Table 8: Performance of $mfFDR^+$ portfolios. The table shows the annualized Sharpe ratios (SR) and mean returns (before and after transaction cost) of implementing the $mfFDR^+$ on all strategies in all currencies to control the FDR at 20%. The last column is the related break-even point.

Period	Excess SR	Net SR	Excess Return (%)	Net Return (%)	Break-even (bps)
Whole Period	1.06	0.95	3.80	3.40	60
1973-1980	1.45	1.35	4.47	4.18	69
1981-1990	2.08	1.97	7.30	6.93	128
1991-2000	0.92	0.81	4.15	3.64	72
2001-2010	0.69	0.53	2.37	1.82	34
2011-2020	0.29	0.21	0.88	0.63	14

To better understand the evolution of time-series variation of technical profitability, we conduct the following three analyses. The first one examines the proportion of technical rules being selected as profitable by the $mfFDR^+$ in each sub-period. The second one examines the portion of technical rules being selected as profitable by the $mfFDR^+$ in each currency in each sub-period. The third one investigates the evolution in the Sharpe ratio distribution of all rules.[16]

For the first purpose, we calculate the ratio of the numbers of trading rules selected as profitable by $mfFDR^+$ in each month divided by the overall number of the input technical rules which varies from 190,755 (= 9×21,195) at the beginning of the sample

[16] We note that, while the two first analyses reflect how prevalent the out-performing rules (in in-sample of one-year periods) are under our framework, they do not reflect how well those rules perform individually in OOS. In our study, the OOS performance is checked after the selected rules are combined so that the transaction cost will be reduced by neutralizing the opposite trading signals. The third analysis is not related to our framework, it merely shows the nature of profitability of all rules which could be detected or not by our framework.

period, when only nine currencies are considered, to 635,850 (= 30×21,195) at the end of the sample. These ratios are averaged so we have a selected ratio per month. We report the results for both the whole sample period and sub-periods in the first column of Table 9. We observe that overall, there are 27% of rules are profitable. This number is 36% in the first decade, then reduced over time to 18% in the most recent decade (2011-2020). The ratio of selected genuinely profitable rules reveals a declining pattern before 1990 and stabilizes since 1991.

Table 9: Ratios of selected technical rules in each category selected by the $mfFDR^{+}$ under controlling for FDR at 20%. The table exhibits the average of the ratios of technical rules (in %) in each category (Channel breakout (CB), Filter rule (FR), Moving average (MA), Relative strength relative (RSI), and Support-resistance (SR)) and in a whole pool of strategies (All) have been selected to invest each month over the whole period (first row) and sub-periods (remaining rows).

	Selected strategy rate					
Period	All	CB	FR	MA	RSI	SR
Whole period	27	8	21	34	13	19
1973-1980	36	11	30	45	19	26
1981-1990	36	10	27	46	15	26
1991-2000	23	7	18	29	12	16
2001-2010	23	7	19	29	11	16
2011-2020	18	6	13	23	11	13

Table 9 also provides the selected ratio in each category of trading rules. This measure shows how rich each technical rule category is in terms of containing outperforming rules and thus being useful for practitioners. We observe that, in the whole sample period, the moving average is the most profitable category, with 34% of the set to be outperforming, followed by filter rule, support-resistance, RSI, and channel breakout categories with 21%, 19%, 13%, and only 8%, respectively. More useful facts are revealed when breaking down the rates into sub-periods. Although the portions vary across periods, the cross-category distribution of the ratios appears consistent over time. Again, the ratio of selected genuinely profitable rules was declining before 1990 but has stabilized since 1991.

Tables 8 and 9 collectively offer two important insights: first, a substantial part of technical trading rules are still able to predict FX rates in recent years. Second, we still find substantial profits from technical rules in the most recent decade (2011-2020).

In Section IG of the Internet Appendix, we further construct portfolios conditional on trading rule category. We find that, first, the portfolio conditional on filter rules is the most profitable OOS before transaction cost and remains so alongside the moving average after transaction cost. Second, the filter and support-resistance categories are trading in higher frequencies and thus generate more trading costs than channel breakout and moving average. Thus, in times when transaction costs are high, traders should avoid using or containing the filter and support-resistance rules in the portfolio. And last but not least, the RSI category performs worst with negative profit for all sub-samples and thus, if traders construct a $mfFDR$ portfolio based on all rules, they should exclude the RSI rules from the pool.

5.4. Evolution of Profitability and Fund Redistribution of the $mfFDR$

It is well-known that technical profitability could decline over time with the improvement of efficiency in foreign exchange markets (especially in developed currencies). As documented in prior studies, the profitability of technical trading rules seems to have vanished in developed currencies in the recent decades (Qi and Wu, 2006; Neely *et al.*, 2009); nevertheless, it still exists in several emerging currencies (Hsu *et al.*, 2016). It is, therefore, important for us to examine the evolution of the profitability of the technical rules in the FX market. As such, we conduct two experiments. First, we track the OOS profitability of the selected rules in each decade. Second, we investigate the ratio of selected profitable trading rules across currencies over time. We argue that the profitability of the technical rules varies across the currencies and across sub-periods but the $mfFDR$ can detect and allocate funds efficiently and, therefore, generates profit over time.

Thereby, for each considering sub-period, we first partition the rules selected by the $mfFDR$ into currencies and track their OOS performance (after combining at the currency level). The gross and net (of transaction cost) Sharpe ratios of the combined rules are presented in Table 10. Second, we report in Table 11 the share (in percentage) of profitable rules selected by the $mfFDR^{+}$ among currencies in the corresponding sub-periods. Except for the first sub-period, in the analyses below, we *do not* consider a currency during its first available sub-period performance record. This is due to the fact that the OOS return of rules selected in this currency does not cover the whole sub-period but for a shorter period compared to other currencies.

In the period from 1973 to 1980, only developed currencies (the Israeli shekel covers only a short period) are considered. We see that all selected rules generate positive OOS Sharpe ratios both before and after transaction cost. The selected rules in the Japanese Yen and British Pound are the most profitable ones with Sharpe ratios of more than one both before and after transaction cost (see Table 10). These two currencies are also allocated with the highest portions of funds (14.8% and 12.9% per month, respectively, see Table 11).

From 1981 to 1990, only the developed currencies and the Israeli shekel are considered. We observe that the rules selected in the Israeli shekel and Euro perform best in OOS

Table 10: OOS performance of trading rules selected by the $mfFDR^{+}$ partitioned into currencies (control for FDR at 20%) in sub-periods. The table displays the OOS Sharpe ratios (before and after transaction cost) of trading rules selected by implementing the $mfFDR^{+}$ on the basket of currencies then partitioned into currencies in five sub-periods: 1973-1980 (P1), 1981-1990 (P2), 1991-2000 (P3), 2001-2010 (P4) and 2011-2020 (P5). For instance, at the end of each month in the sub-period 2011-2020, the outperforming rules selected by the $mfFDR$ portfolio, which applies to the whole 635,850 rules, are partitioned into currencies. As such, there are a number of rules that come from those rules applied to the Australian dollar. Those rules are combined, and a return series of signals generated in the following month is recorded. The Sharpe ratio of the return series before and after transaction cost are presented in columns P4 under the excess Sharpe ratio and net Sharpe ratio groups, respectively.

Country	Excess Sharpe Ratio					Net Sharpe Ratio				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
Australia	0.17	0.44	0.07	0.30	-0.26	0.12	0.38	-0.01	0.28	-0.29
Canada	0.83	0.49	0.12	0.10	-0.11	0.62	0.36	0.01	0.07	-0.15
Germany/E.U.	0.51	1.14	0.16	0.41	-0.17	0.43	1.12	0.13	0.38	-0.19
Japan	1.28	0.87	0.27	-0.13	0.16	1.23	0.81	0.23	-0.15	0.14
New Zealand	0.65	0.65	0.54	0.39	-0.65	0.61	0.54	0.44	0.36	-0.69
Norway	0.36	0.86	-0.14	0.24	-0.27	0.27	0.79	-0.22	0.21	-0.30
Sweden	0.52	1.05	0.33	0.31	-0.28	0.45	1.00	0.25	0.29	-0.31
Switzerland	0.67	0.74	0.15	-0.00	-0.33	0.63	0.70	0.11	-0.03	-0.38
U.K.	1.09	0.88	-0.50	0.00	-0.26	1.02	0.84	-0.57	-0.02	-0.28
Argentina			1.05	0.04	1.09			0.50	-0.01	1.05
Colombia			0.81	0.74	0.30			0.76	0.66	0.20
India			0.47	0.75	-0.19			0.44	0.68	-0.23
Indonesia		-0.01	0.38	0.54	0.67		-0.04	0.30	0.13	0.56
Israel	0.84	2.03	0.61	0.30	-0.10	0.83	1.92	0.27	0.10	-0.15
Philippines		0.78	0.80	0.59	0.09		0.72	0.54	0.41	-0.01
Romania			0.54	0.15	-0.17			0.52	-0.02	-0.24
Russia			0.37	0.81	0.44			0.37	0.76	0.39
Slovak			-0.23	0.61	-0.23			-0.31	0.56	-0.25
Brazil			-0.49	0.63	0.19			-0.52	0.60	0.17
Chile			1.07	0.66	0.00			0.90	0.57	-0.07
Czech			-0.06	0.35	-0.28			-0.11	0.30	-0.37
Hungary			-0.05	0.20	-0.41			-0.09	0.14	-0.47
Korea			0.45	0.33	-0.32			0.45	0.23	-0.52
Mexico		-0.07	0.26	0.04	-0.03		-0.17	0.17	0.01	-0.04
Poland			0.18	0.28	-0.53			0.14	0.22	-0.57
Singapore		0.32	0.42	0.38	-0.04		0.20	0.11	0.19	-0.09
South Africa		0.40	0.36	0.35	-0.13		0.29	0.19	0.17	-0.18
Taiwan		1.91	0.86	0.69	0.34		1.88	0.79	0.52	0.29
Thailand			0.43	0.75	0.17			0.33	0.65	0.08
Turkey			1.01	0.40	0.36			0.99	0.27	0.33

Table 11: Distribution by currency of trading rules selected by the $mfFDR^{+}$ (control for FDR at 20%) in sub-periods. The table displays the average proportion (%) of trading rules selected by implementing the $mfFDR^{+}$ on the basket of currencies in five sub-periods: 1973-1980 (P1), 1981-1990 (P2), 1991-2000 (P3), 2001-2010 (P4) and 2011-2020 (P5). For instance, the average number of outperforming rules selected at the end of each month in the last sub-period, 2011-2020, by the $mfFDR^{+}$ is 112,647 rules (the bottom number of the final column) and, on average, 3.1% of those rules are applied on Australian dollar (the top number of the final column). The bold numbers are the highest ones in a sub-period.

	Period				
Country	P1	P2	P3	P4	P5
Australia	9.6	6.7	3.0	3.5	3.1
Canada	9.3	4.2	2.5	2.4	2.3
Germany/E.U.	9.0	9.3	3.9	3.3	3.0
Japan	**14.8**	7.9	4.8	2.2	3.2
New Zealand	12.1	8.4	3.7	3.9	2.3
Norway	7.7	8.1	3.5	2.9	2.8
Sweden	9.8	9.1	4.4	3.5	3.1
Switzerland	11.4	7.9	4.0	2.6	2.3
U.K.	12.9	7.7	2.7	2.4	2.5
Argentina			1.1	2.8	**10.1**
Colombia			2.8	4.8	3.2
India			3.7	3.4	2.6
Indonesia		1.7	2.8	4.1	3.7
Israel	3.2	**10.9**	4.4	2.7	2.3
Philippines		2.9	6.1	3.2	3.1
Romania			1.5	3.4	3.2
Russia			2.9	4.0	4.5
Slovak			1.1	4.0	2.9
Brazil			1.2	**4.9**	5.0
Chile			2.2	3.5	3.0
Czech			1.7	3.9	3.2
Hungary			2.5	3.4	2.6
Korea			3.1	3.7	2.2
Mexico		0.2	3.0	1.9	3.3
Poland			1.6	3.2	2.9
Singapore		2.9	3.2	2.1	2.2
South Africa		6.2	4.3	4.1	4.1
Taiwan		5.9	4.7	2.6	2.5
Thailand			1.9	3.8	3.2
Turkey			**11.6**	3.8	5.4
Average number of selected rules (per month)	69,028	93,317	122,476	146,532	113,160

with a net Sharpe ratio of 1.92 and 1.12, respectively. Those three currencies are also allocated with the highest proportions of funds, with 10.9% and 9.3% of total funds per month, respectively.

From 1991 to 2000, Indonesia, Philippines, Mexico, Singapore, South Africa, Taiwan, and somewhat Turkey (whose OOS return series starts from January 1992) are added to the considering currencies. The top 20% OOS performance of the 18 considering currencies includes Turkey, Taiwan, and the Philippines, with net Sharpe ratios of 0.99, 0.79, and 0.54, respectively. The three currencies also occupy the highest, fourth, and second highest proportion of invested funds with 11.6%, 4.7%, and 6.1%.

From 2001 to 2010, the OOS return series of all 30 currencies is considered. The top 20% of the most profitable OOS currencies are Russia, India, Colombia, Thailand, Brazil, and Chile, with Sharpe ratios varying from 0.76 to 0.57. An amount of 24.4% of funds are invested in these 20% currencies. To compare, the worst OOS performance rules are those applied in Japan Yen, Switzerland, British Pound, Romania, Argentine Peso, and Mexico Peso with net Sharpe ratios indistinguishable from zero and varying from -0.15 to 0.01. The $mfFDR^+$ allocates only 15% of funds on those currencies.

In the final sub-period, the top 20% of the most profitable OOS countries are Argentina, Indonesia, Russia, Turkey, Taiwan, and Colombia, with net Sharpe ratios varying from 1.05 to 0.2. The $mfFDR^+$ invests 29.4% of the funds on those currencies. In contrast, it invests only 15.5% of funds on the 20% currencies, which turn out to generate the worst OOS net Sharpe ratios.

When we consider all currencies having their OOS performance records, across the five sub-periods, there are 100%, 88%, 77%, 83%, and 30% of currencies generating positive OOS net Sharpe ratio, and the $mfFDR^+$ distributes 100%, 98%, 85%, 87% and 41% of total funds on those currencies, respectively. The fact that the percentages of funds allocated to the positive outperforming currencies are always greater than the proportion of those currencies in our sample suggests that the $mfFDR^+$ correctly puts more weight on those currencies where the technical rules are still profitable. Those analyses suggest that the $mfFDR^+$ is able to detect out-performing rules and allocate funds efficiently across the currencies. Thus, the profitability of the $mfFDR$ portfolio consists of two components: the profitability of technical rules and the ability of $mfFDR$ to distribute funds correctly. This explains why the portfolio has still been profitable over most recent decades, though the profitability of technical rules is challenged by market efficiency in most of the currency markets.

6. Economic Fundamentals for Profitability

As highlighted previously, a large strand of the literature offers different explanations of the profitability of technical trading rules (e.g., Hsu *et al.*, 2016; Menkhoff and Taylor, 2007). For example, the reduction of their profitability can be attributed to the increase in market efficiency over time. However, their pervasiveness and profitability, even accounting for these factors, remain a puzzle. In this section, we provide alternative explanations of the performance of technical trading rules. For this purpose, we use the return of the $mfFDR$ portfolio partitioned into each currency as mentioned in Table 10. As such, each currency has a portfolio return, which represents the OOS performance of the technical rules. Thus, the analyses below explain the persistence in the profitability of technical rules, which is the main interest of practitioners.

6.1. The profitability of trading rules and computational power

Throughout our study, we implement technical trading rules with the use of daily data, and we trade at a daily frequency. Using the advantage in the development of computational power, however, traders have been using high-frequency trading (HFT) and machine learning in trading for decades. It is reasonable to account for the reduction of the profitability of the technical rules evidenced in this study to those developments as competitors.

The use of computational power is difficult to directly measure. In this study, we instrument it by using the exposure of closely related innovations. Particularly, we look at publications in the related area as potential candidates. For this purpose, we collect data about publications on related topics on the Scopus website. More specifically, we retrieve data at the country level on the number of publications with their title, abstract, and keywords containing at least one of the phrases: “algorithmic trading”, “artificial intelligence in trading”, “machine learning in trading,”, ”neural networks in trading”, and “high-frequency trading”.[17] This number is adjusted by the country’s population to have the number of publications per million. We label this new variable as “$Publication$”. Table 12 presents the correlation coefficient between the $Publication$ and the returns of the $mfFDR$-based portfolios. We find that 26 out of 29 countries with available data

[17]The data are exported from https://scopus.com/ in January 2024.

exhibit a negative correlation. Thus, in line with our conjecture, we find a strong impact of the *Publication* on the profitability of the technical rules.

Table 12: Correlation coefficients of *Publication* and $mfFDR$ portfolio's returns. The table presents the correlation coefficients of *Publication*, the number of publications per million people related to computational power in trading, and the return of the $mfFDR$-based portfolio before (Excess return column) and after (Net return column) transaction cost for each country. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20%), then partitioning the selected rules into currencies. We also provide the number of observations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1% of Pearson test for linear correlation.

Country	Net return	Excess return	Number of observations
Australia	-0.12	-0.1	48
Canada	-0.12	-0.1	48
Germany/E.U.	-0.25*	-0.25*	48
Japan	-0.28*	-0.27*	48
New Zealand	-0.26*	-0.24*	48
Norway	0.07	0.08	48
Sweden	-0.4***	-0.39***	48
Switzerland	-0.5***	-0.5***	48
U.K.	-0.21	-0.2	48
Argentina	0.65***	0.64***	28
Colombia	-0.15	-0.17	27
India	-0.42**	-0.42**	28
Indonesia	0.03	0.06	33
Israel	-0.22	-0.2	41
Philippines	-0.1	-0.08	32
Romania	-0.07	-0.02	22
Russia	-0.04	-0.05	25
Slovak	0.01	0.03	26
Brazil	0.04	0.05	25
Chile	-0.32	-0.31	25
Czech	0	-0.01	27
Hungary	-0.24	-0.23	28
Korea	-0.23	-0.24	27
Mexico	-0.09	-0.07	32
Poland	-0.4**	-0.39*	26
Singapore	-0.04	0.01	37
South Africa	-0.15	-0.12	38
Taiwan			0
Thailand	-0.17	-0.13	28
Turkey	-0.16	-0.15	29

To strengthen the evidence, we further collect macro factors at the country level (on a yearly basis) and conduct country-fixed effect regressions of the portfolio returns on the *Publication* in combination with the other country's macro factors. Those factors include GDP growth, unemployment rate (in percent), stock market development, which is the ratio of the country's stock market capitalization over its GDP (in percent), the international trade ratio, which is proxied by the ratio of summation of export and import over the GDP (in percent), and population (in logarithm)[18]. As the data of these macro factors are available yearly, we compute the cumulative daily currency return to obtain a yearly return of $mfFDR$ portfolios and regress it on the mentioned variables, that is:

$$R_{i,t} = a_i + \beta Publication_{i,t} + \sum_{k=1}^{4} Z_{k,i,t} + \epsilon_{i,t} \tag{12}$$

where $Publication_{i,t}$ is the number of publications per million people while $Z_{k,i,t}$s are the macro variables of country i in year t. The results are reported in Table 13.

We find that the models are significant for the use of returns both before (left panel) and after (right panel) transaction costs. The *Publication* has a significantly negative impact on the returns in both the single regression and the regression with controls. This suggests the important role of the developments in computational power in the profitability of the technical rules.

This adds a new explanation to the profitability and the persistent performance of technical rules in FX. Alongside the adaptive market hypothesis, it provides a comprehensive picture of the use of technical trading rules in the FX market. First, it is still widely used since, in some way, and traders can still make a profit from it. The $mfFDR$-based portfolio is a clear example of this. Second, the profitability of the rules, even after advanced filters (like using new FDR control tools), is declining over time due to the development of new technologies that facilitate faster, more complex and low-cost trading.

6.2. The profitability of trading rules and financial development

In this section, we consider the role of capital account openness. To this end, we employ the Chinn and Ito (2006) index. As the index is offered on an annual basis, we use the yearly return of $mfFDR$ portfolios as in Section 6.1. Table 14 reports the correlation coefficients of the Chin-Ito index and the portfolio's net and excess returns, alongside the number of observations (i.e., number of years). We find that the majority of coefficients

[18]The data are collected from https://data.worldbank.org/ in January 2024.

Table 13: Regressions of the *mfFDR*-based portfolio returns on the *Publication* and macro factors. The table reports the regressions with country-fixed effects of the $mfFDR$-based portfolio's return before and after transaction cost on the ratio of the number of publications per million people (Publication), GDP growth rate (GDP growth), the ratio of stock market capitalization to GDP (Stock market development), total export and import per GDP (International trade) and unemployment rate. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20%), then partitioning the selected rules into currencies. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Dependent variable			
	Excess Return		Net return	
Publication	−0.0101***	−0.0072**	−0.0094***	−0.0066**
	(0.0028)	(0.0031)	(0.0028)	(0.0031)
GDP growth		−0.0018**		−0.0016**
		(0.0008)		(0.0008)
Stock market development		−0.0005***		−0.0004***
		(0.0001)		(0.0001)
International trade		0.00003		−0.00001
		(0.0002)		(0.0002)
Unemployment rate		−0.0018		−0.0022**
		(0.0011)		(0.0011)
Observations	1,016	790	1,016	790
R^2	0.0129	0.0689	0.0109	0.0621
Adjusted R^2	−0.0162	0.0283	−0.0182	0.0211
F Statistic	12.8406***	11.1951***	10.9047***	10.0053***

are negative, which indicates that countries with a high degree of capital account openness tend to have lower returns. In our sample, Taiwan does not have the Chinn-Ito index, while the index is constant for Canada, Germany, Switzerland, and India. This results in 25 correlation coefficients. We find that 21 (20) correlation coefficients are negative, and four (five) are positive before (after) accounting for transaction costs. Thus, we find that a country's degree of capital openness could partly provide an explanation of the profitability of technical trading rules.

We conduct regression with country-fixed effects of the portfolio's returns on the Openness index and the macro factors mentioned in Section 6.1:

$$R_{i,t} = a_i + \beta Openness_{i,t} + \sum_{k=1}^{4} Z_{k,i,t} + \epsilon_{i,t} \tag{13}$$

Table 14: Correlation coefficients of country's openness degree and $mfFDR$ portfolio's returns. The table presents the correlation coefficients of the openness index and the $mfDR$-based portfolio's return before (Excess return column) and after (Net return column) transaction cost for each country. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20%), then partitioning the selected rules into currencies. We also provide the number of observations. Of the 30 countries, the openness index is not available for Taiwan, but it is a fixed number for Canada, Germany/EU, and India. In the table, the correlation coefficients for those countries are left blank. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1% of Pearson test for linear correlation.

Country	Net return	Excess return	Number of observations
Australia	-0.15	-0.15	48
Canada			48
Germany/E.U.			48
Japan	-0.34**	-0.33**	48
New Zealand	-0.28*	-0.29**	48
Norway	-0.23	-0.22	48
Sweden	-0.35**	-0.34**	48
Switzerland			25
U.K.	-0.34**	-0.33**	48
Argentina	-0.16	-0.15	28
Colombia	-0.02	-0.05	27
India			28
Indonesia	-0.06	-0.05	33
Israel	-0.35**	-0.33**	41
Philippines	-0.06	-0.1	32
Romania	-0.6***	-0.57***	22
Russia	-0.13	-0.15	25
Slovak	-0.07	-0.04	25
Brazil	0.23	0.22	25
Chile	-0.12	-0.1	25
Czech	0.08	0.07	25
Hungary	-0.03	-0.04	28
Korea	-0.44**	-0.45**	27
Mexico	-0.02	-0.01	32
Poland	-0.23	-0.22	26
Singapore	-0.38**	-0.27	37
South Africa	-0.01	0.01	38
Taiwan			0
Thailand	-0.02	-0.05	28
Turkey	-0.17	-0.14	29

where $Openness_{i,t}$ is the openness index while $Z_{k,i,t}$s are the macro variables of country i in year t. The results for both returns before and after transaction cost are reported in Table 15.

We find that the model is significant for the use of returns both before (left panel) and

Table 15: Regression of the return on the openness index and macro factors. The table reports the regressions with country-fixed effects of the $mfFDR$-based portfolio's return before and after transaction cost on the openness index (Openness), GDP growth rate (GDP growth), the ratio of stock market capitalization to GDP (Stock market development), total export and import per GDP (International trade) and unemployment rate. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20%), then partitioning the selected rules into currencies. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Dependent variable			
	Excess Return		Net return	
Openness	-0.0148^{***}	-0.0071^{**}	-0.0143^{***}	-0.0068^{**}
	(0.0026)	(0.0033)	(0.0026)	(0.0033)
GDP growth		-0.0017^{**}		-0.0015^{**}
		(0.0008)		(0.0008)
Stock market development		-0.0005^{***}		-0.0005^{***}
		(0.0001)		(0.0001)
International trade		0.00005		0.00001
		(0.0002)		(0.0002)
Unemployment rate		-0.0013		-0.0018^{*}
		(0.0011)		(0.0011)
Observations	990	772	990	772
R^2	0.0326	0.0682	0.0303	0.0617
Adjusted R^2	0.0033	0.0266	0.0010	0.0197
F Statistic	32.3200^{***}	10.8090^{***}	29.9747^{***}	9.7038^{***}

after (right panel) transaction cost, and the beta coefficient is significantly negative in throughout cases. The results support the hypothesis that the persistence in profitability of technical rules is weaker in a country or at a time with a higher degree of financial openness.

6.3. The profitability of technical trading rules and macro factors

In this section, we investigate the relationship between the portfolio return and common macro variables used in FX literature. Following Filippou *et al.* (2023), we include local and global variables that are theoretically motivated. Specifically, we include the 3-month T-Bill (BILL), 5-year note (NOTE), 10-year bond (BOND), inflation (INF), unemployment gap (UN), the dividend price ratio (DP), the price to earnings ratio (PE), the stock market capitalization (MKTCAP). All variables are calculated against the U.S., so they are measured as the difference between the local and the U.S. variables. We also include idiosyncratic volatility (IV), volatility (VOL), and illiquidity (ILL).[19] As the data of these macro variables are available monthly, we sum up the daily returns to have monthly returns of the portfolios. We conduct a panel data regression where we use all N mentioned macro variables as explanatory, i.e.:

$$R_{i,t} = a_i + \sum_{k=1}^{11} \beta_k X_{k,i,t} + \epsilon_{i,t} \tag{14}$$

where $R_{i,t}$ is the return of the $mfFDR$ portfolio partitioned onto currency i at time t.[20] We consider both gross and net returns. The estimations are reported in Table 16.

We find that the model is highly significant regardless of the use of returns before (Panel A) or after (Panel B) transaction cost. Of the considered variables, the stock market capitalization, illiquidity, and volatility are highly significant at the 1% level; the price-to-earnings ratio is significant at the level of 5% while the idiosyncratic volatility is so with the use of returns after transaction cost. The fact that coefficients of the market capitalization and price-to-earnings ratio are negative supports a hypothesis that profitability of technical rules is more popular in a country with a lower level of stock market development. The positive coefficients of the illiquidity and volatility variables support the findings that the technical rules are usually helpful in illiquid markets or in highly volatile market conditions.[21] Our analyses thus help uncover the puzzle of profitability of technical trading rules in FX markets.

In Section IJ we repeat all exercises with use of mean return (before transaction cost) as performance metric (ϕ) of the trading rules. We see that our conclusions are unchanged.

[19]Following Filippou, Gozluklu and Taylor (2018), we define volatility (VOL) and illiquidity (ILLIQ) as: $VOL_{i,t} = (1/T_t)\sum_{d\in T_t}(|\Delta s_{i,d}|)$ and $ILLIQ_{i,t} = (1/T_t)\sum_{d\in T_t}(BAS_{i,d})$, respectively, where $|\Delta s_{i,d}|$ represents the absolute change in the log spot exchange rate of currency i on day d. Similarly, $BAS_{i,d}$ is the bid-ask spread in percentage points of currency i on day d. T_t is the total number of days in month t. Thus, an increase in this measure is associated with higher levels of illiquidity.

[20]The data for the mentioned macro variables are not available for five out of 30 countries under study, including Argentina, Colombia, Romania, Chile, and Turkey.

[21]In Section IH of the Internet Appendix, we repeat this exercise with predictive regressions. We see that the INF, MKTCAP, IV, ILL, and VOL are strong predictors of the selected out-performing rules, but they explain a small part of the variation of the returns.

Table 16: Regression of $mfFDR^{+}$-based portfolio returns on macro variables. The table reports the results of the regression with country-fixed effects of the $mfFDR^{+}$-based portfolio returns before and after transaction cost on macro variables. The numbers in parentheses are the corresponding standard deviations. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^{+}$ on all rules traded on the basket of currencies (control for FDR at 20%), then partitioning the selected rules into currencies. "*", "**" and "***" respectively indicate statistical significance at levels of $10\%, 5\%$ and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Excess return	Net return
BOND	0.024^{*}	0.020
	(0.013)	(0.013)
INF	0.056	0.057
	(0.039)	(0.040)
DP	-0.030	-0.028
	(0.020)	(0.021)
PE	-0.00002^{**}	-0.00002^{**}
	(0.00001)	(0.00001)
MKTCAP	-0.020^{***}	-0.020^{***}
	(0.003)	(0.003)
UN	0.006	0.006
	(0.010)	(0.011)
IV	-0.079^{*}	-0.093^{**}
	(0.043)	(0.043)
ILL	0.026^{***}	0.024^{***}
	(0.004)	(0.004)
VOL	0.009^{***}	0.009^{***}
	(0.001)	(0.001)
Observations	7,230	7,230
R^2	0.030	0.028
Adjusted R^2	0.025	0.023
F Statistic (df = 9; 7196)	24.464^{***}	22.897^{***}

7. Conclusion

We introduce the $mfFDR$ testing method, which estimates FDR as a function of multiple covariates for detecting the false null hypotheses with control of FDR for large-scale multiple testing problems. We show that the method works well in controlling FDR and gains a considerably higher power than existing methods in detecting false null hypotheses. Our use of multiple informative covariates helps us examine predictors' conditional performance by incorporating a comprehensive set of information and is applicable to important finance research questions.

Empirically, we apply the $mfFDR$ method to a large universe of technical trading rules to detect truly profitable ones with control of data snooping biases. We first study the trading rules of each currency individually. With the use of multiple informative covariates, the results show that the $mfFDR$-based portfolio is much more powerful than the existing methods that do not use covariates. More importantly, this method quantifies the conditional performance of technical rules, which is more realistic because currency traders and portfolio managers review and select trading strategies based not only on a single performance metric but also on a set of comprehensively updated information.

We then find that the $mfFDR$-based portfolio of selected rules generates positive profits higher than those based on prior data snooping control methods with and without using a sole covariate. We then study 30 currencies together, where more than 600 thousand trading rules are generated. We implement the $mfFDR$ method on this set of rules to construct a portfolio that generates a Sharpe ratio of roughly one for roughly 50 years. Moreover, this portfolio has generated out-of-sample profits even over the recent decades.

Further analyses of the returns of the $mfFDR$-based portfolio reveal the source of its profitability. Thereby, it is the combination of two components: the existence of currencies where the technical rules are still profitable and the ability of the $mfFDR$ to correctly allocate funds to those profitable currencies. By separating the second component out of the analysis, we unfold the factors that significantly explain the profitability of technical rules in the FX market, which is a puzzle in the literature.

We also analyze the economic fundamentals behind the profitability of technical rules in FX markets. We find that such profitability decreases with computational power, capital account openness, equity market development and FX illiquidity. These results collectively attribute the profitability to the barriers to arbitragers who speed up price adjustments and reduce technical analysts' exploitation of the trends in equilibrium FX rates.

The development of the $mfFDR$ framework will contribute to complex problems in Finance, Economics, and other fields of Social Sciences that are plagued by multiple competing models and hypotheses. It has higher power, and is easy to implement and robust to noise and correlation. More importantly, it enables researchers to examine the conditional performance that is more realistic and applicable.

References

Allen, H. and Taylor, M. P. (1990) Charts, noise and fundamentals in the london foreign exchange market. *The Economic Journal*, **100**, 49–59.

Bajgrowicz, P. and Scaillet, O. (2012) Technical trading revisited: False discoveries, persistence tests, and transaction costs. *Journal of Financial Economics*, **106**, 473–491.

Barras, L., Scaillet, O. and Wermers, R. (2010) False discoveries in mutual fund performance: Measuring luck in estimated alphas. *The Journal of Finance*, **65**, 179–216.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**, 289–300.

Burnside, C., Eichenbaum, M. S. and Rebelo, S. (2011) Carry trade and momentum in currency markets. Tech. rep., National Bureau of Economic Research.

Chen, X., Robinson, D. G. and Storey, J. D. (2021) The functional false discovery rate with applications to genomics. *Biostatistics*, **22**, 68–81.

Chinn, M. D. and Ito, H. (2006) What Matters for Financial Development? Capital Controls, Institutions, and Interactions. *Journal of Development Economics*, **81**, 163–192.

Chinn, M. D. and Meese, R. A. (1995) Banking on currency forecasts: How predictable is change in money? *Journal of International Economics*, **38**, 161–178.

DeMiguel, V., Garlappi, L. and Uppal, R. (2007) Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, **22**, 1915–1953.

Filippou, I., Gozluklu, A. E. and Taylor, M. P. (2018) Global political risk and currency momentum. *Journal of Financial and Quantitative Analysis*, **53**, 2227–2259.

Filippou, I., Rapach, D., Taylor, M. P. and Zhou, G. (2023) Out-of-sample exchange rate prediction: A machine learning perspective. *Available at SSRN 3455713.*

Geenens, G. (2014) Probit transformation for nonparametric kernel estimation on the unit interval. *Journal of the American Statistical Association*, **109**, 346–358.

Hansen, P. R. (2005) A test for superior predictive ability. *Journal of Business & Economic Statistics*, **23**, 365–380.

Hsu, P.-H., Hsu, Y.-C. and Kuan, C.-M. (2010) Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias. *Journal of Empirical Finance*, **17**, 471–484.

Hsu, P.-H., Kyriakou, I., Ma, T. and Sermpinis, G. (forthcoming) Mutual funds' conditional performance free of data snooping bias.

Hsu, P.-H., Taylor, M. P. and Wang, Z. (2016) Technical trading: Is it still beating the foreign exchange market? *Journal of International Economics*, **102**, 188–208.

Ignatiadis, N. and Huber, W. (2021) Covariate powered cross-weighted multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **83**, 720–751.

Ignatiadis, N., Klaus, B., Zaugg, J. B. and Huber, W. (2016) Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nature Methods*, **13**, 577–580.

Levich, R. M. and Thomas, L. R. (1993) The significance of technical trading-rule profits in the foreign exchange market: a bootstrap approach. *Journal of International Money and Finance*, **12**, 451–474.

Loader, C. (1999) *Local Regression and Likelihood.* Springer.

Lustig, H., Roussanov, N. and Verdelhan, A. (2011) Common Risk Factors in Currency Markets. *The Review of Financial Studies*, **24**, 3731–3777.

Meese, R. A. and Rogoff, K. (1983) Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, **14**, 3–24.

Menkhoff, L. and Taylor, M. P. (2007) The obstinate passion of foreign exchange professionals: Technical analysis. *Journal of Economic Literature*, **45**, 936–972.

Neely, C., Weller, P. and Dittmar, R. (1997) Is technical analysis in the foreign exchange market profitable? a genetic programming approach. *The Journal of Financial and Quantitative Analysis*, **32**, 405–426.

Neely, C. J. (2002) The temporal pattern of trading rule returns and exchange rate intervention: intervention does not generate technical trading profits. *Journal of International Economics*, **58**, 211–232.

Neely, C. J. and Weller, P. A. (2013) Lessons from the evolution of foreign exchange trading strategies. *Journal of Banking and Finance*, **37**, 3783–3798.

Neely, C. J., Weller, P. A. and Ulrich, J. M. (2009) The adaptive markets hypothesis: Evidence from the foreign exchange market. *The Journal of Financial and Quantitative Analysis*, **44**, 467–488.

Newton, M. A., Noueiry, A., Sarkar, D. and Ahlquist, P. (2004) Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics*, **5**, 155–176.

Olson, D. (2004) Have trading rule profits in the currency markets declined over time? *Journal of Banking & Finance*, **28**, 85–105.

Qi, M. and Wu, Y. (2006) Technical trading-rule profitability, data snooping, and reality check: Evidence from the foreign exchange market. *Journal of Money, Credit and Banking*, **38**, 2135–2158.

Romano, J. P. and Wolf, M. (2005) Stepwise multiple testing as formalized data snooping. *Econometrica*, **73**, 1237–1282.

Storey, J. D. (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **64**, 479–498.

Storey, J. D. (2003) The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics*, **31**, 2013–2035.

Storey, J. D., Akey, J. M. and Kruglyak, L. (2005) Multiple locus linkage analysis of genomewide expression in yeast. *PLOS Biology*, **3**, 1380–1390.

Sullivan, R., Timmermann, A. and White, H. (1999) Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance*, **54**, 1647–1691.

Sullivan, R., Timmermann, A. and White, H. (2001) Dangers of data mining: The case of calendar effects in stock returns. *Journal of Econometrics*, **105**, 249–286.

Taylor, M. P. and Allen, H. (1992) The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance*, **11**, 304–314.

White, H. (2000) A reality check for data snooping. *Econometrica*, **68**, 1097–1126.

Appendix A. The multivariate functional false discovery rate

To complement Section 2.1 of the manuscript, we present the procedures for estimating the $\pi_0(\boldsymbol{z})$, $f(p,\boldsymbol{z})$ and $q-$value.

To estimate $\pi_0(\boldsymbol{z})$ one can extend the "bin" approach presented in Hsu *et al.* (forthcoming), where we partition m tests into some groups based on covariates and estimate $\pi_0(\boldsymbol{z})$ as a constant in each group. In this study, we additionally utilize another approach which is based on density estimation as follows. Firstly, for some $\lambda \in [0,1)$ we define:

$$\pi_0(\boldsymbol{z},\lambda) = \frac{\mathbb{P}(P>\lambda|\boldsymbol{Z}=\boldsymbol{z})}{1-\lambda}. \tag{A.1}$$

Conditional on $\boldsymbol{Z}=\boldsymbol{z}$, $\mathbb{P}(P>\lambda) \geq \mathbb{P}(P>\lambda|h=0).\mathbb{P}(h=0)$. Thus, we have

$$\begin{aligned} \mathbb{P}(P>\lambda|\boldsymbol{Z}=\boldsymbol{z}) &\geq \mathbb{P}(P>\lambda|\boldsymbol{Z}=\boldsymbol{z},h=0).\mathbb{P}(h=0|\boldsymbol{Z}=\boldsymbol{z}) \\ &= (1-\lambda)\mathbb{P}(h=0|\boldsymbol{Z}=\boldsymbol{z}) \\ &= (1-\lambda)\pi_0(\boldsymbol{z}) \end{aligned} \tag{A.2}$$

where the second step comes from the uniform distribution of $P|(\boldsymbol{Z}=\boldsymbol{z},h=0)$. This turns out $\pi_0(\boldsymbol{z},\lambda)\geq\pi_0(\boldsymbol{z})$, i.e. $\pi_0(\boldsymbol{z},\lambda)$ is an conservative estimate of $\pi_0(\boldsymbol{z})$. If we express $\pi_0(\boldsymbol{z},\lambda)$ as

$$\pi_0(\boldsymbol{z},\lambda) = \mathbb{P}(\boldsymbol{Z}=\boldsymbol{z}|P>\lambda).\frac{\mathbb{P}(P>\lambda)}{1-\lambda} \tag{A.3}$$

then, the first term $\frac{\mathbb{P}(P>\lambda)}{1-\lambda}$ is conservatively estimated by $\hat{\pi}_0(\lambda)=\frac{\#\{p_i>\lambda\}}{n(1-\lambda)}$, here $\#$ returns the number of elements in the set, as in Storey (2002). The remain $\mathbb{P}(\boldsymbol{Z}=\boldsymbol{z}|P>\lambda)$, which is the density function of $\boldsymbol{Z}$ conditional on $p-$value $>\lambda$, will be estimated as a function $\hat{h}_\lambda(\boldsymbol{z})$ such that

$$\hat{\pi}_0(\boldsymbol{z},\lambda) = \hat{h}_\lambda(\boldsymbol{z}).\hat{\pi}_0(\lambda) \tag{A.4}$$

is a conservative estimate of $\pi_0(\boldsymbol{z})$.

Next, to estimate the density functions $\hat{h}_\lambda(\boldsymbol{z})$ and $f(p,\boldsymbol{z})$, we use a local likelihood kernel density estimation (KDE) approach in which a probit transformation in Geenens (2014) is adopted. More specifically, let $\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}e^{-x^2/2}dx$ and Φ^{-1} its inverse. We transform the variables $(p,\boldsymbol{z})$ to $(\tilde{p},\tilde{\boldsymbol{z}})$ by using $\tilde{p}_i=\Phi^{-1}(p_i)$ and $\tilde{z}_i^k=\Phi^{-1}(z_i^k)$, $k=1,\dots,d$. We denote by $\tilde{f}(\tilde{p},\tilde{\boldsymbol{z}})$ and $\tilde{h}_\lambda(\tilde{\boldsymbol{z}})$, respectively, the joint density function of $(\tilde{p},\tilde{\boldsymbol{z}})$ and the conditional density function of $\tilde{z}$ on the group of null hypotheses having

$p-$value $< \lambda$, $\{i|P_i > \lambda\}$. We estimate them by using the local likelihood KDE method in Loader (1999). When the number of variables in the estimating function is greater than two (i.e., $d \geq 3$ for $\hat{h}_\lambda(\tilde{\boldsymbol{z}})$ and $d \geq 2$ for $f(p, \tilde{\boldsymbol{z}})$), we drop the cross-product terms from the local model. This allows us to overcome the curse of dimensionality as we are working with one-dimensional integrals instead of a multivariate ones. The estimation can be implemented easily via the freely available R package `locfit`. When the dimension is less than three, the bandwidth of the KDE is chosen locally via a k-Nearest-Neighbor approach using generalized cross-validation similar to CRS.

The desired density functions $\hat{h}_\lambda(\boldsymbol{z})$ and $f(p, \boldsymbol{z})$, respectively, are then estimated, as $\hat{h}_\lambda(\boldsymbol{z}) = \frac{\tilde{h}_\lambda(\tilde{\boldsymbol{z}})}{\Pi_{k=1}^d \phi(\tilde{z}^k)}$ and $\hat{f}(p, \boldsymbol{z}) = \frac{\tilde{f}(\tilde{p}, \tilde{\boldsymbol{z}})}{\phi(\tilde{p})\Pi_{k=1}^d \phi(\tilde{z}^k)}$ where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

For each λ, the $\hat{h}_\lambda(\boldsymbol{z})$ is plugged to (A.4) to acquire the estimate $\hat{\pi}_0(\boldsymbol{z}, \lambda)$. The λ is chosen in the set $\{0.4, \ldots, 0.8\}$ such that the mean integrated squared error defined as in CRS is minimal.

As argued mentioned in footnote 7, if the $f(p, z)$ is non-increasing with respect to p for each fixed $\boldsymbol{z}$, then we should adjust the $\hat{f}(p, z)$ so that it carries this property. In practice, this can be acquired by resetting the value of $\hat{f}(p_i, z_i)$

$$\hat{f}(p_i, z_i) := \min\left\{\hat{f}(p_i, z_i), \min\left\{\hat{f}(p_l, z_l)|p_l < p_i, \|z_i - z_l\| \leq \epsilon\right\}\right\}$$

where $\|.\|$ is the Euclid distance and ϵ is a small positive number which is set at 0.01 in this study.

Finally, we calculate $\hat{r}(p, \boldsymbol{z}) = \hat{\pi}_0(\boldsymbol{z})/\hat{f}(p, \boldsymbol{z})$ and estimate the so-called $q-$value function (see CRS) as

$$\hat{q}(p_i, \boldsymbol{z}_i) = \frac{1}{\#S_i} \sum_{k \in S_i} \hat{r}(p_k, \boldsymbol{z}_k), \tag{A.5}$$

where $S_i = \{k|\hat{r}(p_k, \boldsymbol{z}_k) \leq \hat{r}(p_i, \boldsymbol{z}_i)\}$, p_i is the $p-$value of test i and $\boldsymbol{z}_i = (z_i^1, \ldots, z_i^d)$ is the covariate bundle of strategy i.

We recall here the "positive false discovery rate", a type I error introduced in Storey (2003), $pFDR = \mathbb{E}\left(\frac{V}{R}\middle| R > 0\right)$ where R is the number of rejected null hypotheses in n tests and V the wrongly rejected ones. For a given target $\tau \in [0, 1]$ of $pFDR$, a null hypothesis $H_{0,i}$ is rejected if and only if $\hat{q}(p_i, \boldsymbol{z}_i) \leq \tau$. One can replicate the arguments in CRS to show that the $mfFDR$ procedure controls at the target τ of $pFDR$. We emphasize that control for $pFDR$ is equivalent to control for the FDR at the same level when n is large, which is the case in this study.

Internet Appendix for

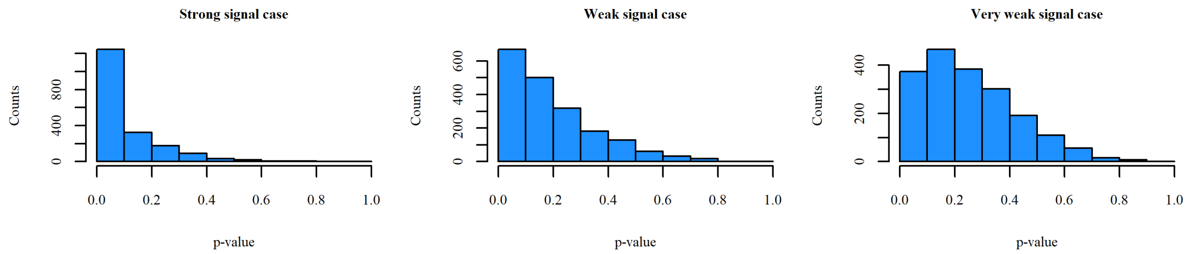
"Conditional Tests for the Profitability of Technical Analysis in Currency Trading and Its Economic Fundamentals"

IA. The role of informative covariates in detecting false nulls

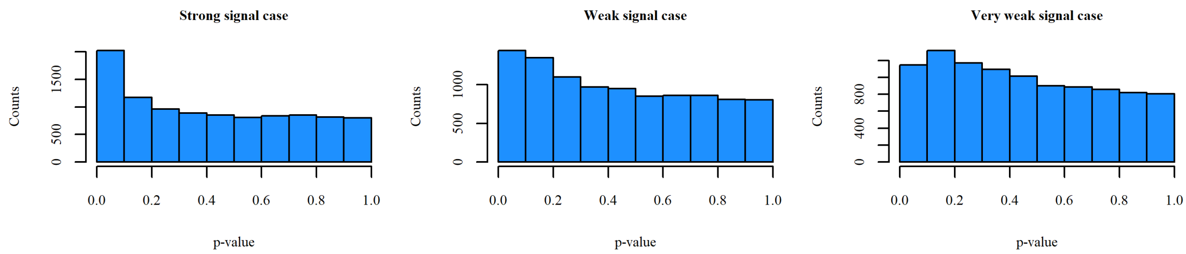
First, we illustrate the differences among the three levels of signals. Figure I depicts from left to right in Panel A (Panel B) the distribution of $p-$values of the false null hypotheses (the whole population) drawn from the strong, weak, and very weak signal cases. In these panels, the probability of being null each hypothesis is generated from the same sine form of $\pi_0(u, v)$, and the $p-$values of the true null hypotheses are drawn from the uniform distribution $[0, 1]$. In the strong signal case ($\alpha = 0.5$), the $p-$values of the false null ones are mostly concentrated near the zero point. In contrast, those $p-$values under the weak signal setting are less condensed at the zero point and dispersed remarkably up to 0.6. In the very weak case, the peak departs from the zero point. The false null hypotheses in the weak and the very weak cases are more difficult to be detected. For example, if we reject a null hypothesis whenever its $p-$value is less than 0.05, then we detect half of the false null hypotheses in the strong signal case while in the very weak signal case, the detected portion is much smaller.[22] In the context of the technical trading rule, a weak signal means that the truly out/under-performing rules have small absolute

Figure I: Distribution of $p-$values of false null hypotheses component (Panel A) and of the whole population (Panel B) in three scenarios: the $p-$values of the false null hypotheses are drawn from a strong signal case, a weak signal case, and a very weak signal case.

Panel A: Distribution of $p-$values of the false nulls



Panel B: Distribution of $p-$values of the population (i.e., the pool of false and true nulls)



[22] We also note that the same number of true null hypotheses will be wrongly rejected at this threshold for all cases. Consequently, the false discovery proportion is much higher in the very weak signal case.

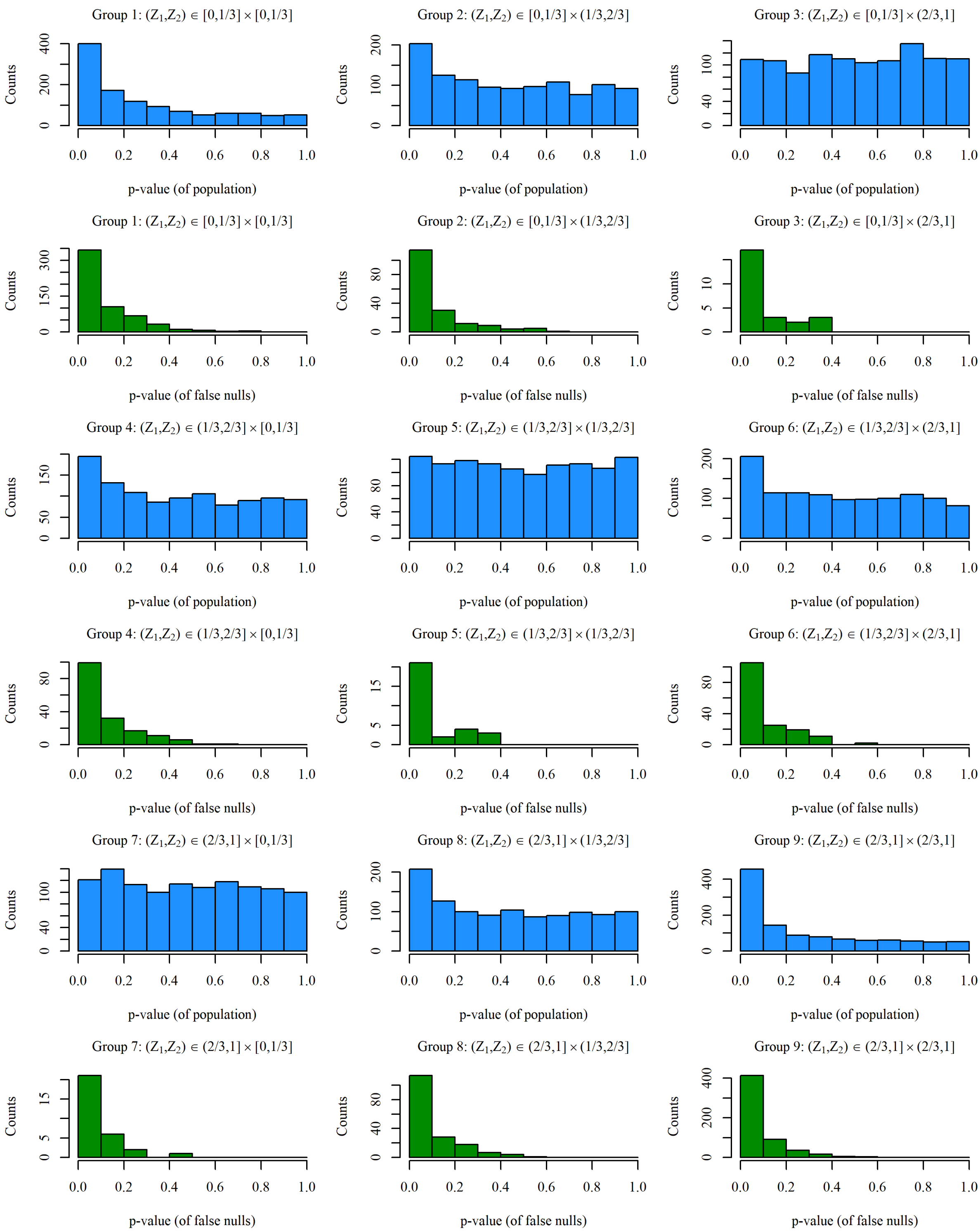
Sharpe ratios, and thus, they have large $p-$values, which make them more difficult to be detected from a random walk.

We analyze a simulated sample to show the usefulness of informative covariates and how they work in the FDR framework. Particularly, we select the one depicted in Panel A of Figure I. This sample is generated under the strong signal setting with the sine form of the $\pi_0(u, v)$. To see how the covariates convey the information of the hypotheses, we partition the sample into nine groups based on dividing each of the u and v into three equal segments. For convenience, the groups are named from 1 to 9, corresponding to the nine areas of (u, v), i.e., each group 1, 2, ... and 9 includes hypotheses having (u, v) in $[0, 1/3] \times [0, 1/3]$, $[0, 1/3] \times (1/3, 2/3]$, ... and $(2/3, 1] \times (2/3, 1]$, respectively. In doing so, each group contains roughly the same number of hypotheses. The $p-$value distributions of the groups are depicted in Figure II.

To control the FDR at a target τ, which is $\tau = 0.1$ in this particular example, the BH and $StdFDR$ reject null hypotheses based on the $p-$values by seeking a threshold at which all null hypotheses with smaller $p-$value are rejected. Thus, this threshold is fixed for all groups. Different from these methods, the blue (green) sub-figures in Figure II present frequencies of the $p-$values of all hypotheses (false null hypotheses). We find more skewed patterns in $p-$values in groups 1, 2, 4, 6, 8, and 9 (especially in green sub-figures for false nulls), which suggests that there are many more false null hypotheses in these groups. On the other hand, the uniform histograms, such as the blue ones of groups 3 and 7, indicate that most hypotheses in these groups are true nulls. This is consistent with the fact that we observe only a handful of false nulls (among more than 1,000 hypotheses) in groups 3 and 5, as shown in green histograms. In other words, the true null proportion in groups 3 and 5 is very high. Presenting the different patterns of $p-$value distributions in different groups confirms that the $p-$values to reject null hypotheses are indeed related to informative covariates.

Hence, when the purpose is to maximize the number of discoveries (while controlling for the FDR at the given target), instead of rejecting all null hypotheses up to a single threshold as in a traditional approach, say 0.05, we could use different thresholds for different groups and informative covariates. For instance, we can use a threshold of 0.2 for hypothesis testing in group 1 and 0 in group 3. With the use of informative covariates in an even finer way under the $mfFDR$, which will be explained below, the number of

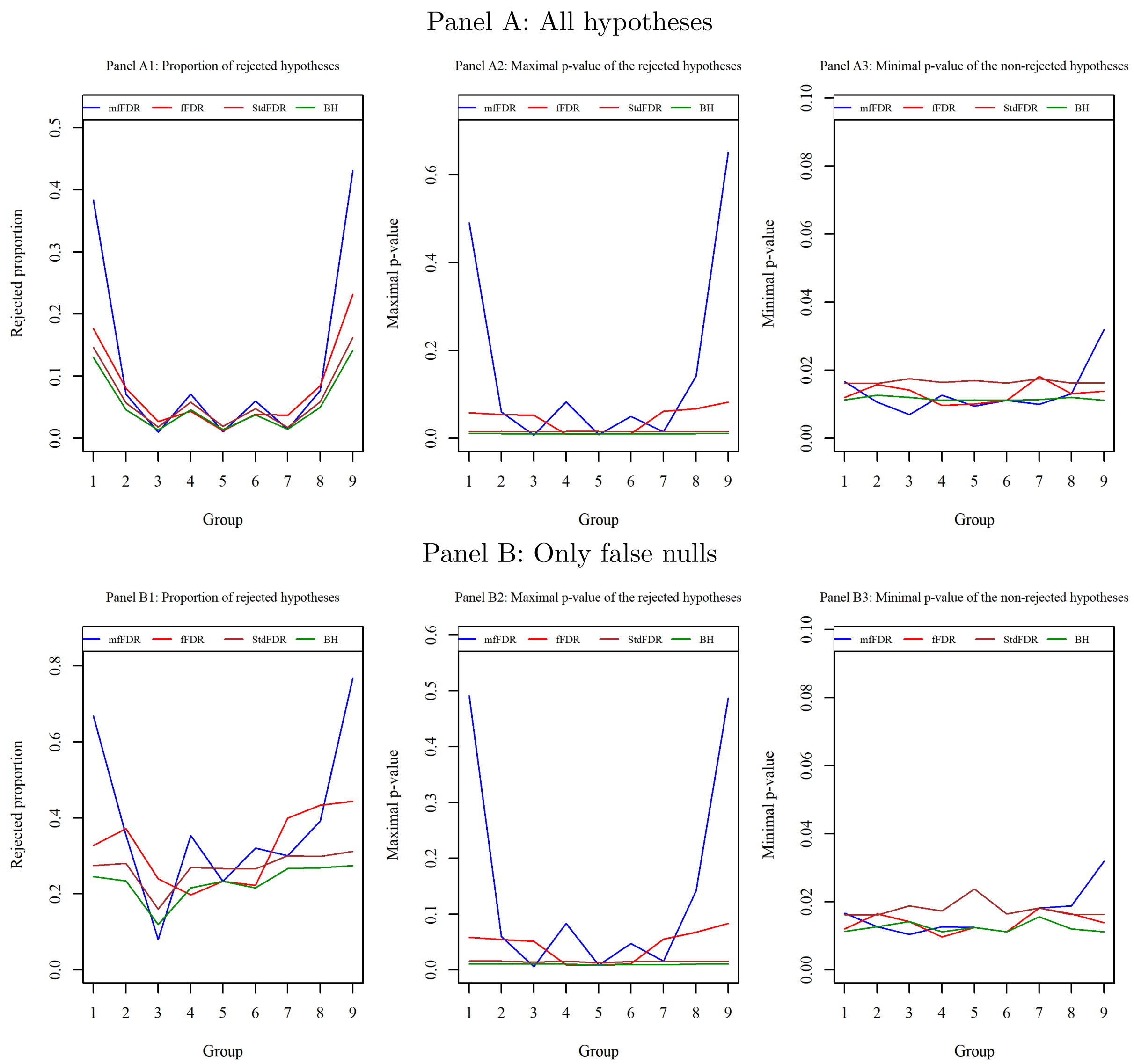
Figure II: Distribution of $p-$values of nine groups which are partitioned from the sample in Panel A of Figure I based on the value of the covariates u and v. Each blue (green) sub-figure represents the frequency distribution of the $p-$values of a group containing the mixture of true and false null hypotheses (false nulls) corresponding to the values of (u, v) shown in its title.



true alternatives detected is much higher and demonstrated clearer in Figure III.

In Panel A1 (B1) of Figure III, each line represents the proportion of null hypotheses (false null hypotheses) rejected by the four procedures across groups 1 to 9. Here, the rejected proportion is the ratio of the number of null hypotheses rejected in a group over the number of null hypotheses (false null hypotheses) in the group. It is clear that all procedures reject more null hypotheses in groups 1 and 9 than they do in groups 3, 5, and 7. Especially, we witness that the rejected proportions of the $mfFDR$ and $fFDR$

Figure III: Comparison of the procedures across the 9 groups described in Figure II. Panel A1 presents the proportion of rejections of each procedure, whereas Panel A2 the corresponding maximal $p-$value of those null hypotheses rejected in each group. Panel A3, in contrast, shows the minimal $p-$value of those hypotheses whose nulls are not rejected. Those metrics calculated in a subgroup of false nulls are presented in corresponding Panels B1, B2, and B3 of Panel B.



are much higher than those of the BH and $StdFDR$ in groups 1 and 9, while the figures of the former are slightly less than that of the $StdFDR$ in group 7. In Panel A2 (B2), we show the maximal $p-$value of rejected null hypotheses (false null hypotheses) of each procedure. As discussed, a null hypothesis is rejected by the BH and $StdFDR$ if its $p-$value is less than some threshold (these thresholds are 0.011 for the BH and 0.016 for the $StdFDR$, which virtually coincide as shown in the green and brown lines in Panel A2, respectively). Hence, the maximal $p-$values of the null hypotheses rejected in the nine groups for the BH and $StdFDR$ are roughly the same. In contrast, the $mfFDR$ rejects some null hypotheses having $p-$values up to 0.491 in group 1, while in group 3, it does not reject any null hypotheses having $p-$values more than 0.01, which is less than the significant threshold of the $StdFDR$. Thus, a few null hypotheses are rejected by the $StdFDR$ but not by the $mfFDR$. Also, it is not uncommon in a group that a null hypothesis with a high $p-$value is rejected while another with a lower $p-$value is not. For instance, in group 1, there is a null hypothesis with a $p-$value of 0.017 that is not rejected by the $mfFDR$; while in the same group, there are null hypotheses with $p-$values up to 0.491 that are rejected.

In summary, our experiments indicate that the $mfFDR$ rejects more false null hypotheses and more null hypotheses in groups where the false null ones are rich. It is worth mentioning that, while partitioning the hypotheses into several groups as presented above illustrates the role of informative covariates in the $mfFDR$ framework, the $mfFDR$ method does not rely on grouping hypotheses into just a few groups.[23] Instead, the $mfFDR$ method treats each hypothesis as a group and establishes a particular rejection threshold for each hypothesis as mentioned in Section 2.1. This is implementable using the $q-$value concept presented in Appendix A.

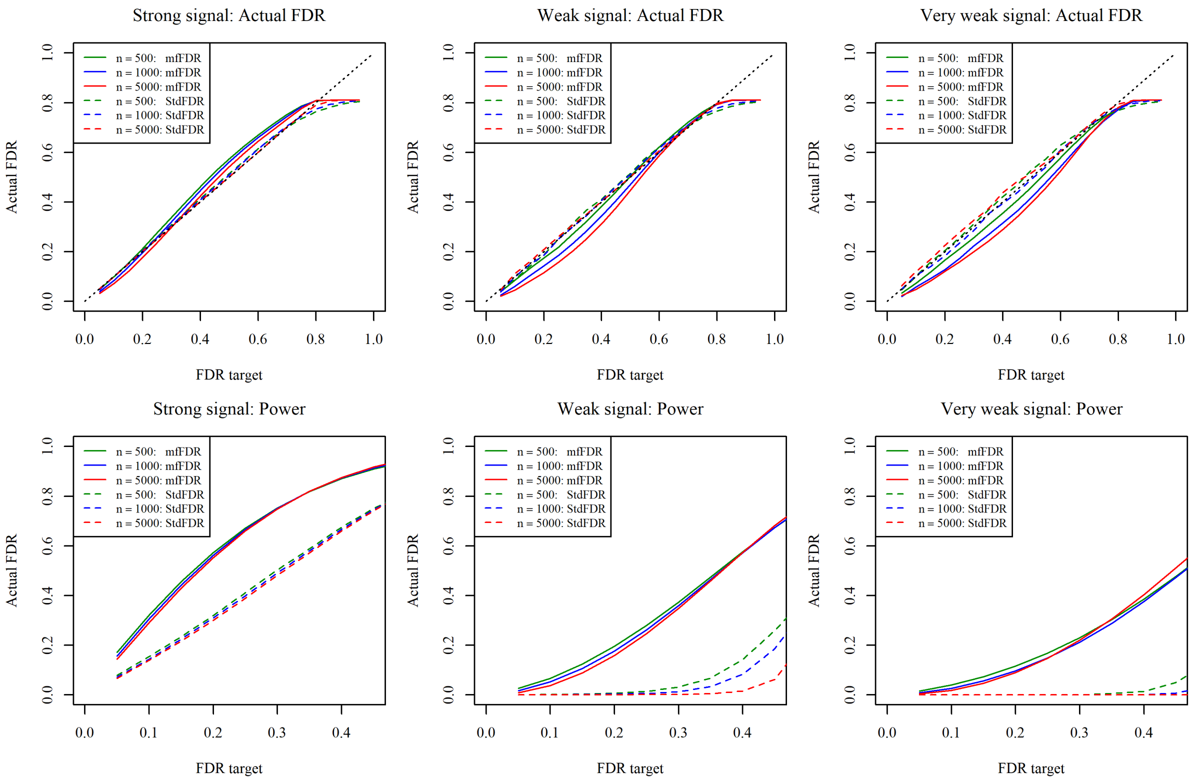
IB. Performance of $mfFDR$ under varying number of tests

In the main manuscript, we conduct simulation studies with a population of n =10,000 tests. This number is chosen since it is close to that of the actual input of $mfFDR$ in the empirical experiments. Aiming for a wider range in the application of the $mfFDR$

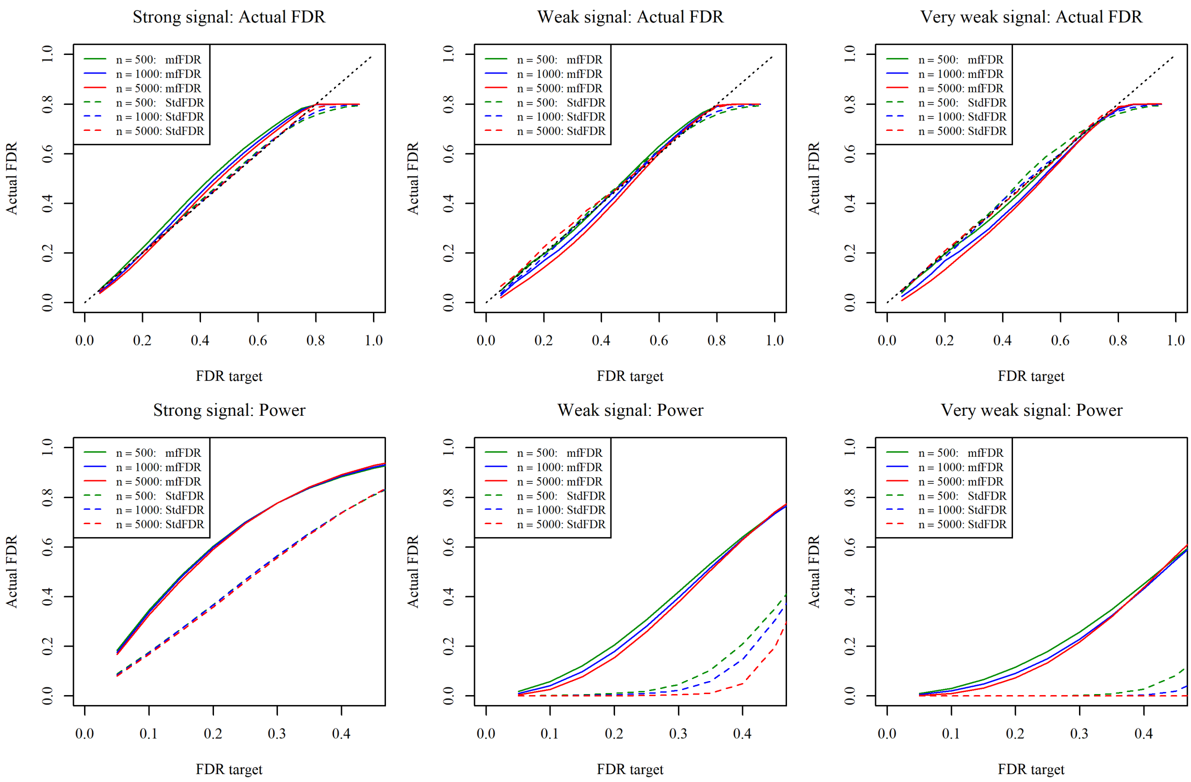
[23] Ignatiadis *et al.* (2016) and Ignatiadis and Huber (2021) recently introduce a group weighting approach where they partition hypotheses to groups based on a single covariate and determine the rejection threshold (of $p-$value) in each group. This approach, however, is less powerful than the $fFDR$ of CRS (see Section 1 of the Supplementary Materials of CRS).

Figure IV: Performance comparison of the $mfFDR$ against the Standard FDR of Storey ($SdtFDR$) with varying population size (n). Panel A shows the performance when the $\pi_0(u,v)$ has a sine form whereas Panel B is the monotonic one.

Panel A: $\pi_0(u,v)$ is a sine function.



Panel B: $\pi_0(u,v)$ is monotonic with respect to each covariate.



framework, we additionally conduct a robustness check with use of smaller numbers of tests. Particularly, we repeat the simulation with: $n = 500, 1000$ and 5000 tests.

In Figure IV, we present the performance of the $mfFDR$ against its benchmark, the $StdFDR$, in terms of FDR control and power. First, under the weak and very weak signal cases, the $mfFDR$ controls well for any given FDR targets. Second, for strong signal data, the $mfFDR$ slightly violates the FDR control at high targets, especially when the number of tests is small. More specifically, when $n = 500$, the $mfFDR$ strictly controls well for FDR at targets up to 0.2. When $n = 3000$ and $n = 5000$, these numbers are 0.3 and 0.4. Thus, if the number of considering tests is smaller than 500 and the aim is controlling for an FDR target higher than 0.2, the method should be used on weak or very weak signal rather than strong signal data. Although, it should be noted that controlling FDR at a high target with a small number of tests, has no practical value in finance, economics and most fields of science. Last, in terms of power, the $mfFDR$ is superior to that of the $StdFDR$ in all cases regardless of the number of tests.

IC. Correlation and estimation errors of covariates

In this section, we design a simple model where the covariates are positively correlated.[24] Given a correlation coefficient of r, the two covariates studied in the previous section are transformed (with the use of Cholesky factorization) into two new covariates having a correlation coefficient of approximately r. The simulated data are then generated similarly to the previous section. For the interest of space, we present the results for only the sine form of the null proportion function.[25] Aiming to study the impact of the correlation in covariates on the performance of the $mfFDR$, we consider varying values of r from 0.1 to 0.8.

Figure V depicts the performance of the $mfFDR$ in terms of FDR control (Panel A) and its power compared to others (Panels B and C). For the former aspect, the FDR is well controlled when $r < 0.7$ regardless of signal level. This coefficient range covers most cases in our real data in which, as shown in Section 5 of the main manuscript, more than 95% of the empirical coefficients have an absolute value less than 0.7. When $r \geq 0.7$,

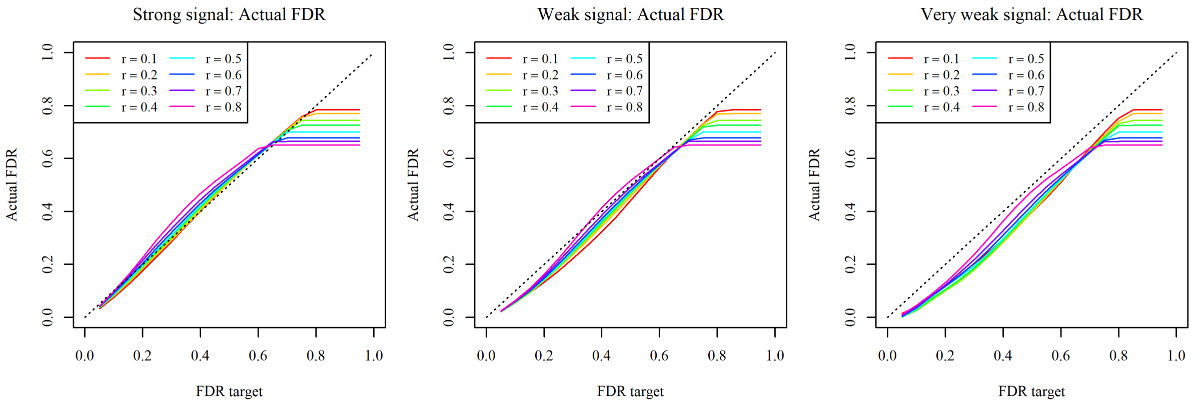
[24]The performances of the procedures under negatively correlated covariates are similar.

[25]These two new covariates are not uniformly distributed on $[0, 1]$, thus, the null proportion function is slightly modified, $\pi_0(u, v) = \min\{1, \max\{0, \sin(\pi(u + v)/2)\}\}$, so that its values are in $[0, 1]$.

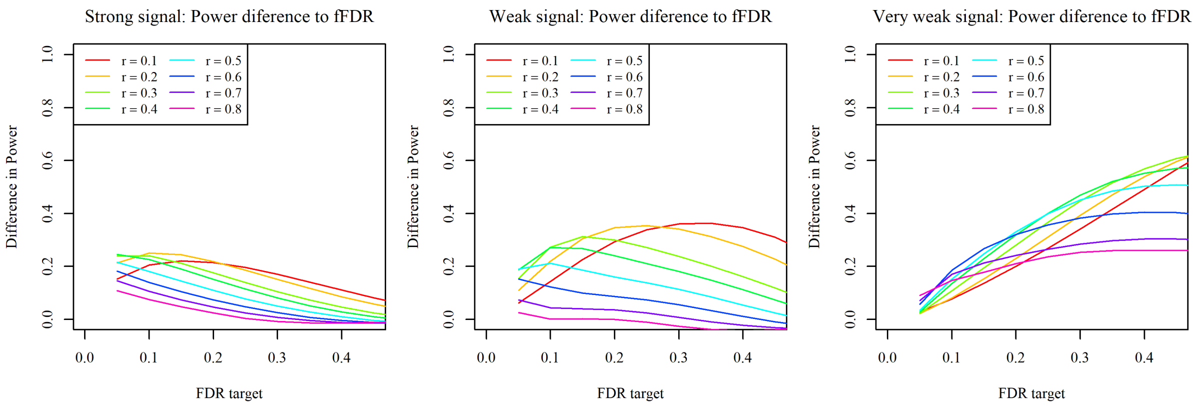
the FDR is controlled well under the weak and very weak signals and asymptotically controlled in the strong case.

Figure V: Performance of the $\boldsymbol{mfFDR}$ under correlated covariates. Panel A exhibits the performance of the procedures in terms of FDR control whereas Panels B and C present the power differences of the $mfFDR$ over the $fFDR$ and the FDR procedure of Storey (2002) ($StdFDR$), respectively.

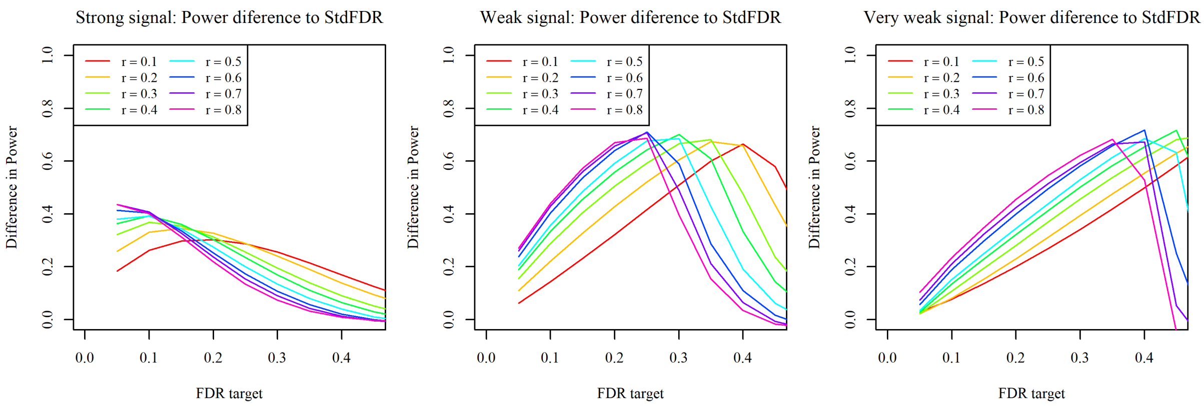
Panel A: FDR control of $mfFDR$ under correlated covariates



Panel B: Gap in power of $mfFDR$ over $fFDR$ under correlated covariates



Panel C: Gap in power of $mfFDR$ over $StdFDR$ under correlated covariates



It is noted that the powers of the $mfFDR$ under different correlation coefficients are incomparable due to the differences in the level of signals. That is, the signal in higher correlation settings might be stronger due to the transformed covariates as the $p-$values of tests are generated based on them. Consequently, to assess the performance in terms of power, we calculate the gaps in the power of the $mfFDR$ over the $fFDR$ and $StdFDR$ under each case of r and depict them in Panels B and C, respectively. In the sub-figures of these panels, a line above zero indicates a higher power of the $mfFDR$ compared to its benchmarks. This is the case for all of the considered settings, with the peak varying across cases and could reach about 70% or 30%, under the FDR target of 20%, when the benchmark is the $StdFDR$ or $fFDR$, respectively.

ID. Performance of $mfFDR$ under noisy covariates

As mentioned in Section C, the covariates are estimated quantities and, thus, have inherited some noise that might affect the power of our method. In this section, we address this concern by considering a simple setting where the input covariates contain noise. More specifically, we use the covariates $u = (u_1, \dots, u_n), v = (v_1, \dots, v_n)$ as in our previous simulations in Section 2.2 to generate the simulated data, but now we use inputs of $mfFDR$ the observed $u' = (u'_1, \dots, u'_n)$ and $v' = (v'_1, \dots, v'_n)$ defined as

$$u'_i = u_i + \eta_i \tag{D.1}$$

and

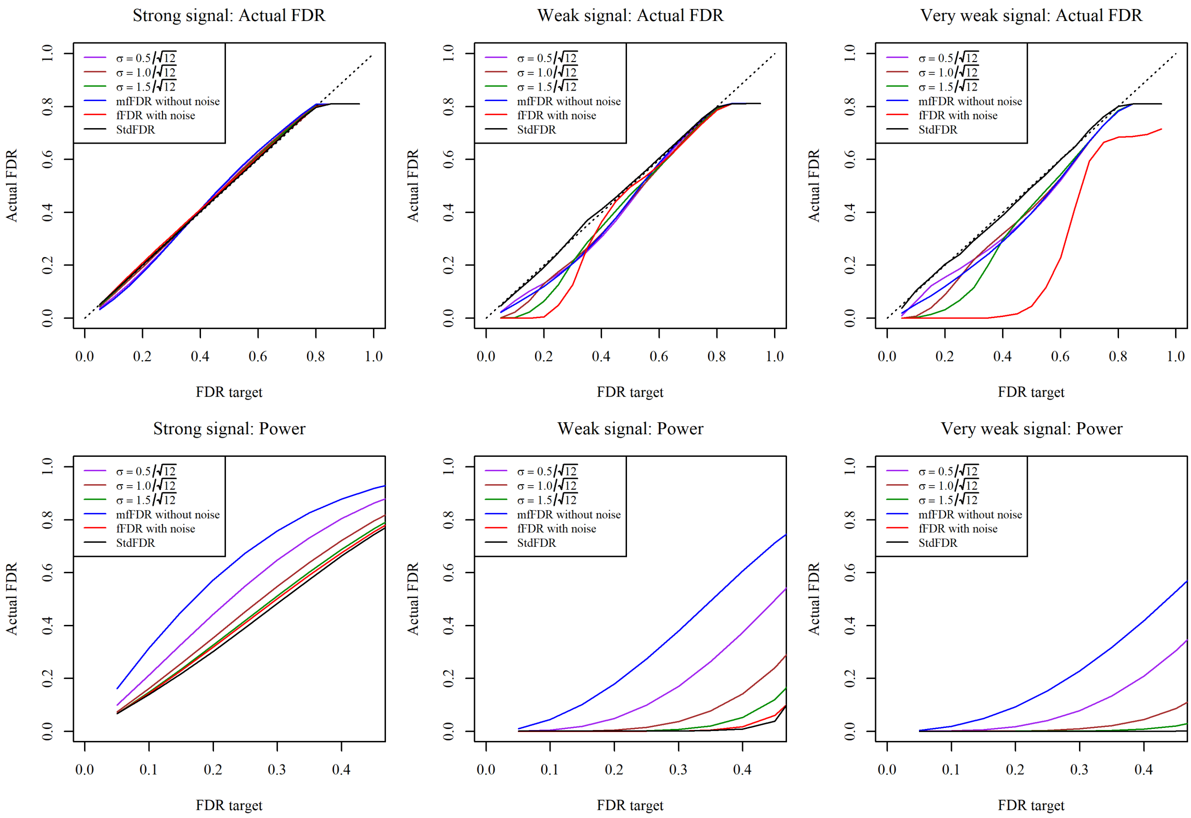
$$v'_i = v_i + \epsilon_i \tag{D.2}$$

where η_i and ϵ_i are noise generated independently from a normal distribution $N(0, \sigma^2)$, $i = 1, \dots, n$. To study the performance of the $mfFDR$, we consider three different values of σ including $\sigma_1 = 0.5/\sqrt{12}$, $\sigma_2 = 1.0/\sqrt{12}$ and $\sigma_3 = 1.5/\sqrt{12}$. The determination of those values is based on the fact that the covariates $u, v \sim U[0, 1]$, which has a standard deviation of $1/\sqrt{12}$.

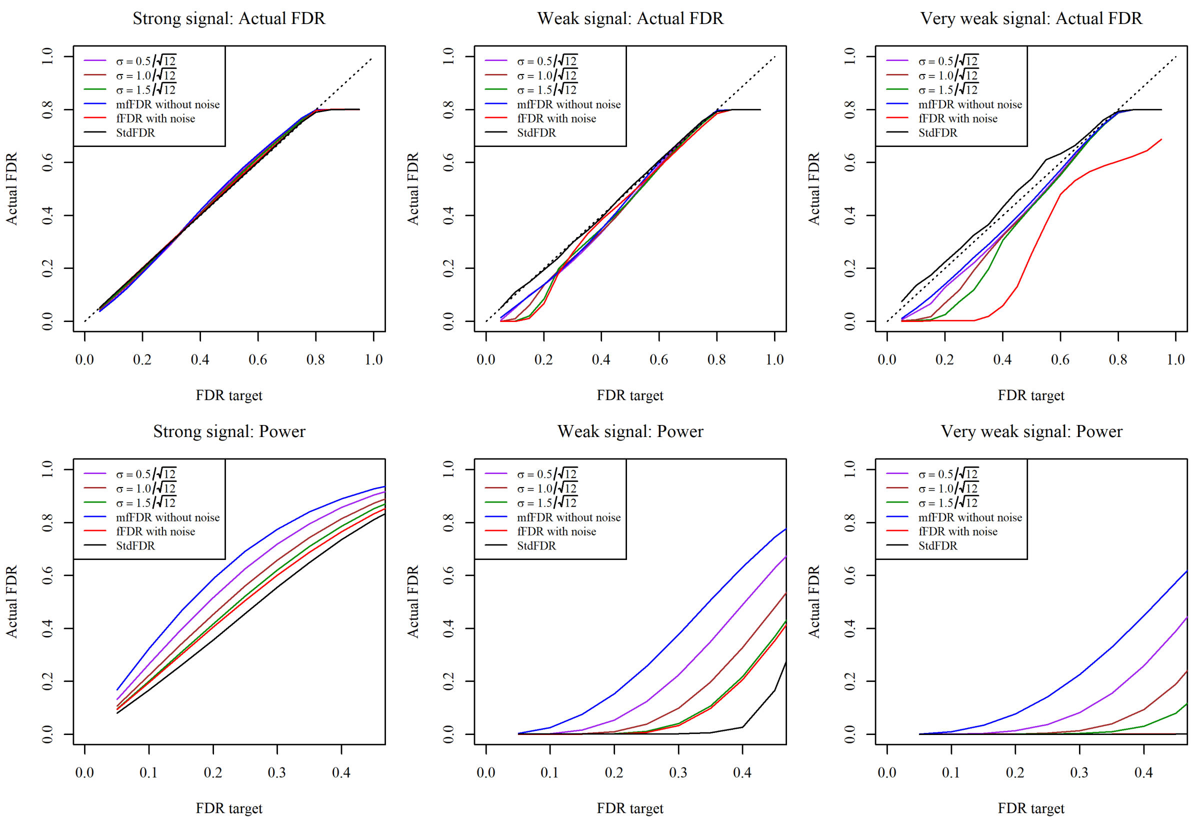
The magnitude of the noise in this setting reflects the level of informativeness of the covariates (the higher variation of the noise is, the less informative the covariates are) or the strength of the relationship between the observed covariates and the true status of hypotheses (the higher variation of the noise is, the weaker the relationship is). In Figure VI, we present the performance of $mfFDR$ in terms of FDR control and its power (at FDR target up to 0.45).

Figure VI: Performance comparison of the $mfFDR$ and $fFDR$ with noisy covariates and the Standard FDR of Storey ($SdtFDR$) procedures. Here the the input covariates for the $mfFDR$ are $u' = u+\varepsilon, v' = v+\eta$, where $\varepsilon, \eta \sim N(0, \sigma^2)$ and $\sigma \in \{0.5/\sqrt{12}, 1.0/\sqrt{12}, 1.5/\sqrt{12}\}$, whereas the $fFDR$ the u'. Panel A shows the performance when the $\pi_0(u, v)$ has a sine form, whereas Panel B is the monotonic one.

Panel A: $\pi_0(u, v)$ is a sine function.



Panel B: $\pi_0(u, v)$ is monotonic with respect to each covariate.



We see that the FDR is controlled well at any given target from 0.05 to 0.95. To have a complete picture of the impact of the noise on the power of the covariate augmented methods, we add the performance of the $fFDR$ when its covariate is u' and compare it to the $mfFDR$ using original (u, v) and the $StdFDR$. Evidently, the power of the $mfFDR$ with noise in covariates is lower than the case with original ones but still remarkably higher than that of the $fFDR$ (with noised covariate u') and $StdFDR$.

IE. Dependence in p-values

In a particular category of trading rules, the rules with close parameters tend to have highly correlated returns. This leads to a dependency among the testing statistics or $p-$values of the corresponding hypotheses. This section shows that our method is robust under this type of dependence. Specifically, we are generating data such that the hypotheses are partitioned into groups with the same size k. In each group, the $p-$values of testing hypotheses are mutually dependent at the same level, which is characterized by a covariance matrix $\mathbf{\Sigma}$. The $p-$values corresponding to the hypotheses from different groups are independent.

The data-generating process is as follows. The simulated covariates, $\pi_0(u, v)$, and the true status of null hypotheses H are as described in Section 2.2. For the sake of space, we present only the results for the sine form of $\pi_0(u, v)$. To account for the dependence in hypotheses, we first partition the true null hypotheses into groups of size k. Secondly, we generate the $z-$scores for the null hypotheses of each group from the multivariate normal distribution $\mathcal{N}(0, \mathbf{\Sigma})$. We process similarly for the false null ones, but the $z-$scores of each group are drawn from $\mathcal{N}(2, \mathbf{\Sigma})$. To simplify, the matrix $\mathbf{\Sigma} = (\Sigma_{ij})_{k\times k}$ is set as $\Sigma_{ii} = 1$ and $\Sigma_{ij} = c$ for $i \neq j = 1, \ldots, k$ for some c. By considering various values of the parameters k and c, we reveal the impact of the dependence at different levels on the performance of the $mfFDR$. Here, we choose $c \in \{0, 0.25, 0.5, 0.75\}$ where the case $c = 0$ indicates the absence of the dependence among $p-$values and will be used as a benchmark for a comparison purpose, and $k \in \{10, 100, 500\}$. The parameter k represents the dependence scale. A larger k indicates the presence of more hypotheses that are mutually dependent.[26] Finally, the $p-$value of each (two-sided) test is calculated

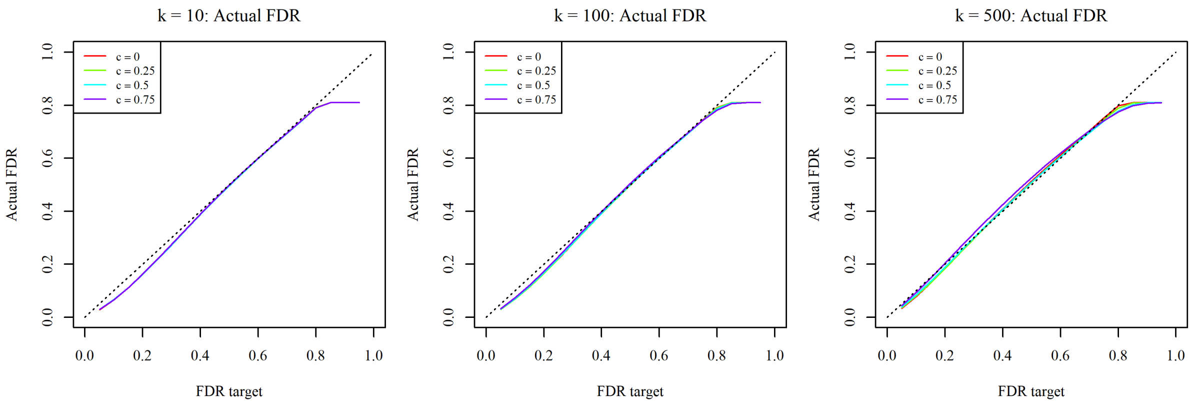
[26]This type of dependent setting is also studied in Storey (2003) with $k = 10$. Here, we extend the cases of k to study its impact on the method's performance.

from its $z-$score by using the cumulative distribution function of the standard normal distribution. Note that, in this setting, the covariates convey only the information on the probability of being true null of the hypotheses (i.e. via $\pi_0(u, v)$) and not on the $p-$values (i.e. we do not generate $p-$values via $f_{alt}(p|(u, v))$, but via the dependent $z-$scores, which does not depend on the covariates) as in the simulations conducted in the previous sections.

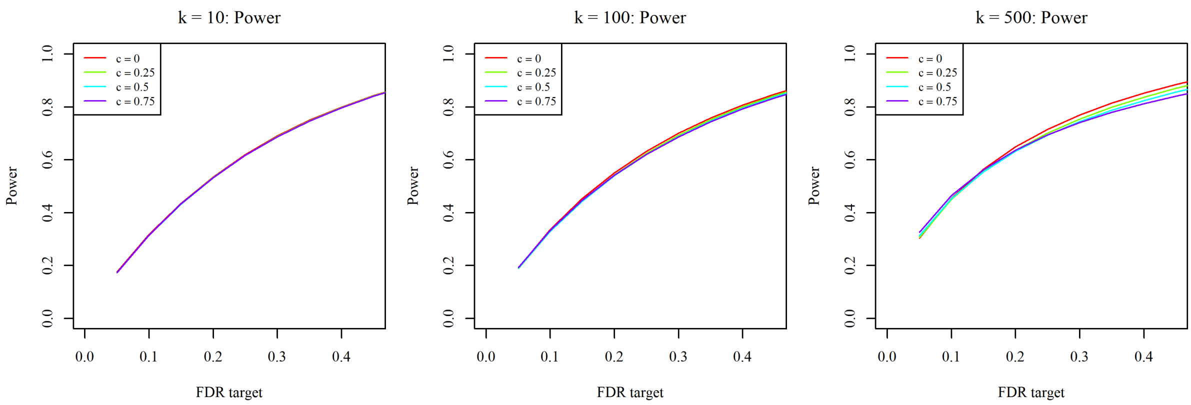
Figure VII presents the performance of $mfFDR$ against $fFDR$ and $StdFDR$. Panel A of the figure shows that the considering dependence type does not affect the FDR control of all methods. In terms of power, the $mfFDR$ outperforms its benchmarks. It is also clear that $fFDR$ is better than $StdFDR$. When k is small, the first two sub-figures of Panel B show an insignificant difference between the power of $mfFDR$ with different levels of dependence among $z-$scores in each group. Therefore, the lines corresponding to different values of c are virtually identical and covered by a blue line. Their powers are only distinguishable in the third sub-figure, where k is large. In this case, the power

Figure VII: Performance of the $mfFDR$ under correlated $p-$values. Panel A (B) presents the FDR control (power) of the $mfFDR$ under different levels of dependencies among $p-$value of the tests.

Panel A: FDR control of the $mfFDR$ under dependent $p-$values



Panel B: Power of the $mfFDR$, $fFDR$ and $StdFDR$ under dependent $p-$values



is decreasing concerning c, i.e. the level of dependence level among the $z-$scores.

IF. Performance of $\boldsymbol{mfFDR^+}$ based portfolios with various FDR targets

In this section, we present the performance of the $mfFDR^+$ based portfolios both when it is applied to individual currency as well as on all currencies on various FDR targets. More specifically, we repeat the experiment in the main manuscript with FDR target τ varying from 10% to 40%. The results for the former portfolios are presented in Table I while those of all currencies together are exhibited in Table II. The results corresponding to $\tau = 0.2$ are represented for convenience in comparison. Overall, the performance of the portfolios is stable across the considered targets. When all currencies are examined together, we observe a higher Sharpe ratio and smaller annual return for higher targets. This observation implicitly indicates that the volatility of the portfolio's daily return is higher when we control FDR at a small target.

Table I: Performance of $mfFDR$ based portfolios on individual currency with varying FDR target. The table shows annualized Sharpe ratios of the $mfFDR$ based portfolio with FDR target $\tau = \{0.1, 0.2, 0.3, 0.4\}$ based on portfolios' returns before (left side) and after transaction cost (right side). The final row shows the average Sharpe ratio across 30 portfolios corresponding to the 30 currencies. The numbers in parentheses are the corresponding $p-$values. "*", "**" and "***" respectively indicate statistical significance at levels of $10\%, 5\%$ and 1%.

Countries	Before transaction cost				After transaction cost			
	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$
Australia	0.17 (0.21)	0.18 (0.17)	0.18 (0.17)	0.18 (0.17)	0.13 (0.34)	0.14 (0.29)	0.14 (0.30)	0.14 (0.31)
Canada	0.19 (0.17)	0.17 (0.23)	0.17 (0.25)	0.17 (0.24)	0.13 (0.37)	0.11 (0.46)	0.11 (0.47)	0.11 (0.46)
Germany/E.U.	0.48 (0.00)***	0.49 (0.00)***	0.48 (0.00)***	0.48 (0.00)***	0.45 (0.00)***	0.46 (0.00)***	0.45 (0.00)***	0.45 (0.00)***
Japan	0.46 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.44 (0.00)***	0.43 (0.00)***	0.41 (0.00)***	0.41 (0.00)***	0.41 (0.00)***
New Zealand	0.32 (0.01)***	0.34 (0.00)***	0.35 (0.00)***	0.35 (0.00)***	0.27 (0.02)**	0.28 (0.02)**	0.29 (0.01)***	0.29 (0.01)***
Norway	0.20 (0.15)	0.18 (0.17)	0.18 (0.19)	0.18 (0.18)	0.15 (0.29)	0.13 (0.36)	0.13 (0.38)	0.13 (0.37)
Sweden	0.37 (0.00)***	0.40 (0.00)***	0.40 (0.00)***	0.40 (0.00)***	0.33 (0.02)**	0.36 (0.01)***	0.36 (0.00)***	0.36 (0.00)***
Switzerland	0.26 (0.06)*	0.26 (0.06)*	0.25 (0.07)*	0.25 (0.07)*	0.22 (0.11)	0.22 (0.12)	0.22 (0.12)	0.21 (0.13)
U.K.	0.29 (0.05)**	0.29 (0.05)**	0.28 (0.05)**	0.28 (0.06)*	0.26 (0.07)*	0.25 (0.08)*	0.24 (0.10)*	0.24 (0.10)*
Argentina	0.35 (0.06)*	0.35 (0.06)*	0.34 (0.06)*	0.35 (0.06)*	0.32 (0.10)*	0.31 (0.11)	0.29 (0.11)	0.30 (0.12)
Columbia	0.56 (0.00)***	0.60 (0.00)***	0.62 (0.00)***	0.61 (0.00)***	0.48 (0.01)***	0.53 (0.01)***	0.55 (0.00)***	0.54 (0.01)***
India	0.35 (0.03)**	0.34 (0.03)**	0.34 (0.03)**	0.34 (0.04)**	0.31 (0.06)*	0.30 (0.07)*	0.30 (0.07)*	0.30 (0.07)*
Indonesia	0.33 (0.02)**	0.34 (0.02)**	0.33 (0.02)**	0.33 (0.02)**	0.23 (0.10)*	0.24 (0.09)*	0.24 (0.09)*	0.24 (0.10)*
Israel	0.52 (0.00)***	0.54 (0.00)***	0.54 (0.00)***	0.54 (0.00)***	0.37 (0.02)**	0.39 (0.01)***	0.39 (0.01)***	0.39 (0.01)***
Philippines	0.64 (0.00)***	0.62 (0.00)***	0.62 (0.00)***	0.62 (0.00)***	0.48 (0.01)***	0.46 (0.01)***	0.46 (0.01)***	0.46 (0.01)***
Romania	0.19 (0.31)	0.23 (0.19)	0.24 (0.16)	0.25 (0.15)	0.09 (0.65)	0.12 (0.51)	0.13 (0.47)	0.13 (0.46)
Russia	0.47 (0.01)***	0.51 (0.01)***	0.52 (0.01)***	0.52 (0.01)***	0.44 (0.02)**	0.48 (0.01)***	0.49 (0.01)***	0.49 (0.01)***
Slovak	0.16 (0.40)	0.18 (0.35)	0.19 (0.33)	0.19 (0.31)	0.11 (0.51)	0.13 (0.48)	0.13 (0.46)	0.14 (0.43)
Brazil	0.36 (0.07)*	0.36 (0.07)*	0.35 (0.08)*	0.35 (0.08)*	0.33 (0.10)*	0.33 (0.10)*	0.32 (0.10)*	0.32 (0.10)*
Chile	0.49 (0.01)***	0.46 (0.02)**	0.45 (0.02)**	0.45 (0.02)**	0.41 (0.04)**	0.38 (0.06)*	0.37 (0.06)*	0.37 (0.06)*
Czech	0.07 (0.64)	0.12 (0.46)	0.13 (0.43)	0.14 (0.41)	0.01 (0.85)	0.06 (0.65)	0.07 (0.61)	0.08 (0.60)
Hungary	-0.12 (0.56)	-0.11 (0.60)	-0.11 (0.60)	-0.11 (0.62)	-0.17 (0.39)	-0.17 (0.41)	-0.16 (0.41)	-0.16 (0.43)
Korea	0.30 (0.15)	0.29 (0.16)	0.29 (0.16)	0.29 (0.16)	0.24 (0.23)	0.23 (0.25)	0.23 (0.26)	0.23 (0.26)
Mexico	0.10 (0.48)	0.10 (0.49)	0.09 (0.50)	0.10 (0.49)	0.05 (0.71)	0.04 (0.73)	0.04 (0.74)	0.04 (0.73)
Poland	-0.02 (0.90)	0.02 (0.92)	0.02 (0.88)	0.03 (0.85)	-0.07 (0.73)	-0.03 (0.89)	-0.03 (0.91)	-0.02 (0.93)
Singapore	0.28 (0.05)**	0.32 (0.02)**	0.32 (0.02)**	0.32 (0.02)**	0.11 (0.52)	0.15 (0.32)	0.15 (0.32)	0.15 (0.33)
South Africa	0.19 (0.19)	0.23 (0.14)	0.25 (0.10)*	0.25 (0.09)*	0.07 (0.61)	0.10 (0.49)	0.12 (0.43)	0.12 (0.43)
Taiwan	0.73 (0.00)***	0.73 (0.00)***	0.73 (0.00)***	0.73 (0.00)***	0.66 (0.00)***	0.66 (0.00)***	0.65 (0.00)***	0.65 (0.00)***
Thailand	0.46 (0.02)**	0.47 (0.02)**	0.46 (0.02)**	0.46 (0.02)**	0.37 (0.05)**	0.38 (0.04)**	0.38 (0.04)**	0.38 (0.04)**
Turkey	0.51 (0.00)***	0.51 (0.00)***	0.51 (0.00)***	0.50 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.44 (0.00)***
Average	0.32	0.33	0.33	0.33	0.25	0.26	0.26	0.26

Table II: Performance of $mfFDR^+$ based portfolio with various FDR targets. The table shows the annualized Sharpe ratios and mean returns (before and after transaction cost) and break-even point (bps) by implementing the $mfFDR^+$ on all strategies in all currencies to control the FDR at 10% (Panel A), 20% (Panel B), 30% (Panel C) and 40% (Panel D). The selected out-performing strategies are then combined by currency. The fund is allocated to trade on each of the currencies having out-performing strategies weighted by the size of the selected out-performing trading rules. The first row presents the numbers of the whole sample period while the rest are those of sub-periods.

Period	Excess Sharpe Ratio	Net Sharpe Ratio	Excess Return	Net Return	Break-even
	Panel A: FDR target of 10%				
Whole Period	0.98	0.88	4.36	3.91	63
1973-1980	1.40	1.31	4.89	4.58	76
1981-1990	2.11	2.01	8.28	7.89	141
1991-2000	0.81	0.72	4.93	4.38	84
2001-2010	0.62	0.47	2.53	1.89	34
2011-2020	0.33	0.25	1.28	0.99	10
	Panel B: FDR target of 20%				
Whole Period	1.06	0.95	3.80	3.40	60
1973-1980	1.45	1.35	4.47	4.18	69
1981-1990	2.08	1.97	7.30	6.93	128
1991-2000	0.92	0.81	4.15	3.64	72
2001-2010	0.69	0.53	2.37	1.82	34
2011-2020	0.29	0.21	0.88	0.63	14
	Panel C: FDR target of 30%				
Whole Period	1.11	0.99	3.54	3.15	57
1973-1980	1.49	1.39	4.29	4.01	76
1981-1990	2.02	1.91	6.69	6.33	118
1991-2000	1.00	0.87	3.77	3.28	66
2001-2010	0.72	0.56	2.28	1.75	34
2011-2020	0.32	0.23	0.83	0.60	8
	Panel D: FDR target of 40%				
Whole Period	1.12	1.00	3.34	2.97	54
1973-1980	1.47	1.37	4.11	3.84	75
1981-1990	1.98	1.87	6.30	5.94	112
1991-2000	1.06	0.92	3.57	3.10	63
2001-2010	0.73	0.57	2.20	1.70	34
2011-2020	0.29	0.20	0.72	0.50	7

IG. Portfolios conditional on categories of trading rules

Table 9 of our main manuscript clearly shows that the performance of rules is different across categories. In other words, the category is itself an informative covariate. It is, therefore, interesting to see how $mfFDR$ performs conditional to the category. However, this covariate is not continuous and thus cannot be an input of the $mfFDR$. To study its informativeness, we therefore repeat the experiment with a basket of currencies but on each of the categories. In doing so, we are able to answer two questions: first, which category is the most profitable OOS, and second, which categories traders should consider or avoid when constructing a portfolio. More specifically, for each category of trading rules, we pool the rules across currencies belonging to the category to form a new set of rules. Thus, we have five new sets: the largest set is that of the Moving Average category with a number of rules varying from 9×2,870 =25,830 to 30×2,870 = 86,100 rules, while the smallest set is that of RSI category containing from $9 \times 600 =$ 54,000 to $30 \times 600 =$18,000 rules. We construct five $mfFDR$-based portfolios, which control FDR at 20%, each for one of the new sets.

In Table III, we present the OOS performance of those selected rules. We learn several facts from the results. First, generally, the portfolio conditional on filter rules is the most profitable OOS before transaction cost and remains so alongside the moving average after transaction costs. Second, the fact that the break-even points corresponding with filter and support-resistance categories are lower than those of channel breakout and moving average ones implies that the former ones are trading in higher frequencies and thus generate more trading costs. The implication is that, in times when transaction costs are high, traders should avoid using or containing the filter and support-resistance rules in the portfolio. Third, the RSI category performs worst with negative profit for all sub-samples and thus for the whole sample period. This implies that the performance of the selected out-performing RSI rules is not persistent. Thus, if traders construct a $mfFDR$ portfolio based on all rules, they should exclude the RSI rules from the pool.

Table III: Performance of the $mfFDR^+$ portfolios implemented on each category. The table shows the annualized Sharpe ratios (SR) and mean returns (before and after transaction cost) of portfolios generated by implementing the $mfFDR^+$ on each category of trading rule (on all currencies) to control the FDR at 20%. The last column is the related break-even point. In each panel, we present the mentioned metrics for the portfolio implemented on a category in the whole sample period (first row) and sub-samples (next five rows).

Period	Excess SR	Net SR	Excess Return (%)	Net Return (%)	Break-even (bps)
Panel A: Channel Breakout Rule					
Whole Period	0.88	0.80	2.64	2.40	72
1973-1980	0.82	0.76	2.05	1.88	62
1981-1990	2.03	1.95	5.81	5.57	184
1991-2000	0.98	0.89	3.45	3.14	99
2001-2010	0.47	0.36	1.43	1.11	37
2011-2020	0.12	0.06	0.33	0.18	7
Panel B: Filter Rule					
Whole Period	1.26	0.91	4.41	3.19	23
1973-1980	1.77	1.53	5.62	4.86	36
1981-1990	2.17	1.81	7.14	5.97	41
1991-2000	0.98	0.66	4.90	3.27	27
2001-2010	1.02	0.44	2.95	1.28	14
2011-2020	0.67	0.37	1.71	0.95	7
Panel C: Moving Average					
Whole Period	1.01	0.92	3.80	3.47	71
1973-1980	1.39	1.32	4.49	4.24	89
1981-1990	1.99	1.91	7.42	7.12	156
1991-2000	0.92	0.83	4.23	3.82	89
2001-2010	0.65	0.53	2.43	1.97	41
2011-2020	0.18	0.11	0.57	0.37	9
Panel D: RSI					
Whole Period	-0.46	-0.96	-0.50	-1.04	-
1973-1980	-0.73	-1.16	-0.63	-0.99	-
1981-1990	-0.20	-0.70	-0.25	-0.84	-
1991-2000	-1.04	-1.67	-1.05	-1.70	-
2001-2010	-0.15	-0.73	-0.19	-0.90	-
2011-2020	-0.40	-0.74	-0.40	-0.74	-
Panel E: Support-resistance					
Whole Period	1.00	0.80	3.13	2.48	32
1973-1980	1.37	1.20	3.64	3.18	41
1981-1990	2.06	1.83	6.02	5.34	64
1991-2000	0.87	0.68	3.68	2.87	41
2001-2010	0.55	0.24	1.54	0.67	15
2011-2020	0.34	0.19	0.87	0.49	8

IH. Macro variables as predictors

This section, we repeat regressions presented in Section 6.3 in the main manuscript with use of the lag of the macro variables. The results are presented in Table IV. We see that INF, MKTCAP, IV, ILL and VOL are significantly predicting the profitability of the selected out-performing rules but the goodness-of-fit are low.

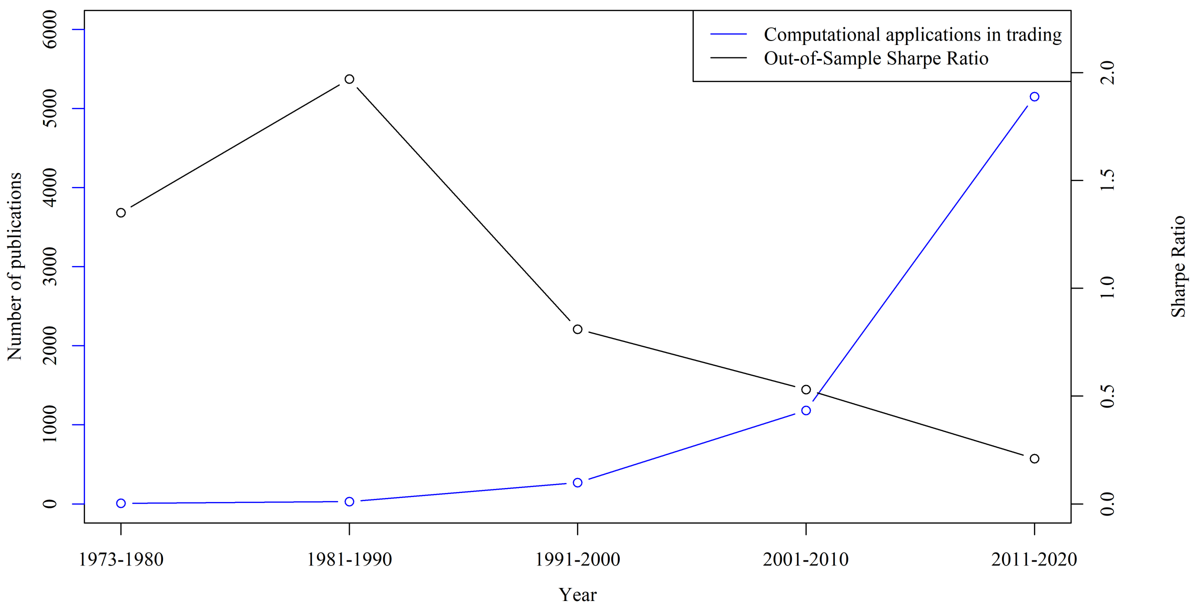
Table IV: Regression of $mfFDR^{+}$-based portfolio returns on the lag of the macro variables. The table reports the results of the regression with country-fixed effects of the $mfFDR^{+}$-based portfolio returns before and after transaction cost on the lag of the macro variables. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%.

	Dependent variable	
	Excess return	Net return
BOND	0.022^{*}	0.017
	(0.013)	(0.013)
INF	0.123^{***}	0.124^{***}
	(0.040)	(0.040)
DP	−0.009	−0.007
	(0.021)	(0.021)
PE	−0.00000	−0.00000
	(0.00001)	(0.00001)
MKTCAP	-0.008^{***}	-0.008^{***}
	(0.003)	(0.003)
UN	−0.003	−0.004
	(0.011)	(0.011)
IV	-0.127^{***}	-0.139^{***}
	(0.043)	(0.043)
ILL	-0.019^{***}	-0.021^{***}
	(0.004)	(0.004)
VOL	0.003^{***}	0.003^{***}
	(0.001)	(0.001)
Observations	7,207	7,207
R^2	0.008	0.009
Adjusted R^2	0.004	0.004
F Statistic (df = 9; 7173)	6.538^{***}	7.147^{***}

II. Number publications and $mfFDR$-based portfolio's return

Figure VIII depicts the number of publications contain in their title, abstract, and keywords containing at least one of the phrases: "algorithmic trading", "artificial intelligence in trading," "machine learning in trading,", "neural networks in trading," and "high-frequency trading" over five decades. We also add the OOS Sharpe ratio of $mfFDR$-based portfolio presented in Table 8.[27] The figure clearly shows the relationship between the Sharpe Ratio and the other two metrics. Prior to the 1990s, there were few publications in both two categories. The number of publications started to be significant in the 1990s. Coinciding with this, the OOS performance of the selected out-performing rules decline dramatically at the time. The decreasing trend continues in the next two decades, and we observe an increase in the publications in both two mentioned categories. Note that the presence of HFT and machine learning in financial trading obviously improves market liquidity. Thus, the observations are consistent with the finding on the impact of illiquidity in the previous section.

Figure VIII: Profitability of technical rules and computational power. The figure depicts the total number of publications (blue line with numbers presented on the left y-axis) on Scopus with title, abstract, and keywords containing words related to computational power in trading and the OOS Sharpe ratio of $mfFDR$-based portfolio (black line with numbers presented on the right y-axis).



IJ. Performance of $mfFDR$-based portfolios with use of mean excess return as the testing performance metric

As a robustness check, we repeat all experiments presented in the main manuscript with the use of mean excess return as a performance metric in hypothesis testing (ϕ).

[27]The data are exported from scopus.com in January 2024.

Table V presents the OOS performance of the $mfFDR$-based portfolios, both before and after transaction cost, when implementing the method on individual currencies to control FDR at targets of 10%, 20%, 30% and 40%. Table VI exhibits the results when implementing the method on all currencies together. Tables VII, VIII and IX present the results for regressions of the $mfFDR$-based portfolio's return on: i) number of publications related to computational power per million people and common macro factors; ii) openness index and common macro factors; and iii) common macro variables in FX literature, respectively.

Table V: Performance of $mfFDR$ based portfolios on individual currency with varying FDR target when mean return is used as the testing performance metric (ϕ). The table shows annualized Sharpe ratios of the $mfFDR$ based portfolio with FDR target $\tau = \{0.1, 0.2, 0.3, 0.4\}$ based on portfolios' returns before (left side) and after transaction cost (right side). The final row shows the average Sharpe ratio across 30 portfolios corresponding to the 30 currencies. The numbers in parentheses are the corresponding $p-$values. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%.

Countries	Before transaction cost				After transaction cost			
	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$
Australia	0.20 (0.11)	0.20 (0.10)*	0.20 (0.11)	0.19 (0.11)	0.16 (0.19)	0.16 (0.20)	0.15 (0.21)	0.15 (0.22)
Canada	0.19 (0.15)	0.18 (0.17)	0.18 (0.17)	0.18 (0.17)	0.13 (0.31)	0.12 (0.39)	0.12 (0.41)	0.12 (0.38)
Germany/E.U.	0.45 (0.00)***	0.46 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.42 (0.00)***	0.42 (0.00)***	0.42 (0.00)***	0.41 (0.01)***
Japan	0.46 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.45 (0.00)***	0.43 (0.00)***	0.41 (0.01)***	0.42 (0.01)***	0.41 (0.01)***
New Zealand	0.33 (0.01)***	0.35 (0.00)***	0.35 (0.00)***	0.36 (0.00)***	0.27 (0.03)**	0.29 (0.02)**	0.30 (0.02)**	0.30 (0.02)**
Norway	0.21 (0.09)*	0.19 (0.14)	0.18 (0.14)	0.18 (0.14)	0.16 (0.20)	0.14 (0.27)	0.13 (0.29)	0.13 (0.29)
Sweden	0.37 (0.01)***	0.40 (0.01)***	0.40 (0.01)***	0.40 (0.01)***	0.32 (0.01)***	0.35 (0.01)***	0.35 (0.01)***	0.35 (0.01)***
Switzerland	0.30 (0.04)**	0.30 (0.02)**	0.30 (0.03)**	0.29 (0.03)**	0.26 (0.06)*	0.26 (0.06)*	0.26 (0.07)*	0.26 (0.07)*
U.K.	0.32 (0.02)**	0.30 (0.04)**	0.30 (0.05)**	0.30 (0.05)**	0.28 (0.05)**	0.27 (0.06)*	0.26 (0.07)*	0.26 (0.07)*
Argentina	0.22 (0.29)	0.24 (0.25)	0.24 (0.23)	0.24 (0.23)	0.18 (0.37)	0.19 (0.33)	0.20 (0.31)	0.20 (0.31)
Columbia	0.52 (0.01)***	0.57 (0.00)***	0.57 (0.00)***	0.57 (0.00)***	0.45 (0.03)**	0.50 (0.01)***	0.51 (0.01)***	0.50 (0.01)***
India	0.37 (0.02)**	0.35 (0.03)**	0.35 (0.03)**	0.35 (0.03)**	0.33 (0.05)**	0.31 (0.06)*	0.31 (0.06)*	0.30 (0.06)*
Indonesia	0.33 (0.02)**	0.34 (0.02)**	0.33 (0.02)**	0.33 (0.02)**	0.24 (0.10)*	0.24 (0.10)*	0.24 (0.10)*	0.24 (0.10)*
Israel	0.30 (0.04)**	0.31 (0.03)**	0.31 (0.03)**	0.31 (0.03)**	0.12 (0.38)	0.13 (0.32)	0.13 (0.33)	0.13 (0.33)
Philippines	0.66 (0.00)***	0.64 (0.00)***	0.63 (0.00)***	0.63 (0.00)***	0.49 (0.01)***	0.47 (0.01)***	0.46 (0.02)**	0.47 (0.02)**
Romania	0.16 (0.39)	0.20 (0.28)	0.20 (0.25)	0.20 (0.26)	0.05 (0.78)	0.08 (0.65)	0.09 (0.62)	0.09 (0.62)
Russia	0.40 (0.09)*	0.41 (0.08)*	0.42 (0.08)*	0.42 (0.08)*	0.38 (0.10)*	0.39 (0.08)*	0.40 (0.08)*	0.40 (0.08)*
Slovak	0.14 (0.42)	0.15 (0.40)	0.16 (0.37)	0.17 (0.36)	0.10 (0.58)	0.10 (0.56)	0.11 (0.54)	0.12 (0.50)
Brazil	0.35 (0.08)*	0.35 (0.09)*	0.34 (0.09)*	0.34 (0.09)*	0.32 (0.11)	0.31 (0.11)	0.31 (0.11)	0.31 (0.11)
Chile	0.44 (0.03)**	0.40 (0.04)**	0.40 (0.05)**	0.40 (0.05)**	0.36 (0.08)*	0.32 (0.11)	0.31 (0.12)	0.31 (0.12)
Czech	0.07 (0.64)	0.11 (0.51)	0.12 (0.47)	0.12 (0.46)	0.01 (0.88)	0.05 (0.71)	0.06 (0.67)	0.06 (0.64)
Hungary	-0.13 (0.49)	-0.12 (0.54)	-0.12 (0.55)	-0.12 (0.57)	-0.18 (0.35)	-0.18 (0.36)	-0.17 (0.37)	-0.17 (0.38)
Korea	0.32 (0.12)	0.30 (0.14)	0.29 (0.15)	0.29 (0.15)	0.26 (0.19)	0.24 (0.21)	0.24 (0.21)	0.24 (0.21)
Mexico	0.08 (0.55)	0.08 (0.55)	0.08 (0.57)	0.08 (0.56)	0.03 (0.78)	0.03 (0.80)	0.03 (0.81)	0.03 (0.80)
Poland	-0.05 (0.78)	-0.01 (0.98)	-0.01 (0.98)	0.00 (0.98)	-0.10 (0.59)	-0.05 (0.78)	-0.06 (0.78)	-0.05 (0.81)
Singapore	0.27 (0.05)**	0.32 (0.03)**	0.32 (0.03)**	0.32 (0.03)**	0.11 (0.54)	0.15 (0.35)	0.14 (0.37)	0.14 (0.37)
South Africa	0.17 (0.24)	0.21 (0.16)	0.23 (0.12)	0.24 (0.11)	0.05 (0.71)	0.09 (0.53)	0.11 (0.46)	0.11 (0.45)
Taiwan	0.78 (0.00)***	0.77 (0.00)***	0.77 (0.00)***	0.77 (0.00)***	0.70 (0.00)***	0.70 (0.00)***	0.69 (0.00)***	0.69 (0.00)***
Thailand	0.46 (0.02)**	0.47 (0.01)***	0.47 (0.01)***	0.47 (0.01)***	0.37 (0.05)**	0.39 (0.04)**	0.39 (0.04)**	0.38 (0.04)**
Turkey	0.48 (0.00)***	0.48 (0.00)***	0.48 (0.00)***	0.48 (0.00)***	0.42 (0.00)***	0.42 (0.00)***	0.42 (0.00)***	0.42 (0.00)***
Average	0.31	0.31	0.31	0.31	0.24	0.24	0.24	0.24

Table VI: Performance of $mfFDR^{+}$ based portfolio with various FDR targets when mean return is used as the testing performance metric (ϕ). The table shows the annualized Sharpe ratios and mean returns (before and after transaction cost) and break-even point (bps) by implementing the $mfFDR^{+}$ on all strategies in all currencies to control the FDR at 10% (Panel A), 20% (Panel B), 30% (Panel C) and 40% (Panel D). The selected out-performing strategies are then combined by currency. The fund is allocated to trade on each of the currencies having out-performing strategies weighted by the size of the selected out-performing trading rules. The first row presents the numbers of the whole sample period, while the rest are those of sub-periods.

Period	Excess Sharpe Ratio	Net Sharpe Ratio	Excess Return	Net Return	Break-even
Panel A: FDR target of 10%					
Whole Period	1.00	0.90	4.48	4.04	65
1973-1980	1.43	1.34	5.01	4.69	78
1981-1990	2.09	1.99	8.36	7.96	141
1991-2000	0.85	0.76	5.27	4.71	89
2001-2010	0.63	0.47	2.51	1.89	34
2011-2020	0.35	0.28	1.37	1.08	11
Panel B: FDR target of 20%					
Whole Period	1.08	0.97	3.87	3.47	60
1973-1980	1.45	1.36	4.47	4.18	75
1981-1990	2.07	1.97	7.33	6.95	128
1991-2000	0.95	0.84	4.31	3.80	75
2001-2010	0.69	0.53	2.35	1.80	34
2011-2020	0.34	0.26	1.03	0.78	10
Panel C: FDR target of 30%					
Whole Period	1.12	1	3.55	3.17	57
1973-1980	1.47	1.38	4.24	3.97	75
1981-1990	2.02	1.91	6.71	6.34	119
1991-2000	1.02	0.89	3.85	3.37	68
2001-2010	0.72	0.55	2.24	1.73	34
2011-2020	0.34	0.25	0.88	0.65	9
Panel D: FDR target of 40%					
Whole Period	1.13	1	3.36	2.99	55
1973-1980	1.46	1.36	4.06	3.80	75
1981-1990	1.99	1.87	6.33	5.97	113
1991-2000	1.08	0.94	3.62	3.16	65
2001-2010	0.73	0.56	2.18	1.68	33
2011-2020	0.31	0.22	0.76	0.54	8

Table VII: Regressions of the $mfFDR$-based portfolio returns the with use of mean return as performance metric in testing (ϕ) on the *Publication* and macro factors. The table reports the regressions with country-fixed effects of the $mfFDR$-based portfolio's return before and after transaction cost on the ratio of the number of publications per million people (Publication), GDP growth rate (GDP growth), ratio of stock market capitalization to GDP (Stock market development), total export and import per GDP (International trade) and unemployment rate. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20% with use of mean return as performance metric (ϕ) in testing), then partitioning the selected rules into currencies. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Dependent variable			
	Excess Return		Net return	
Publication	-0.0098^{***}	-0.0072^{**}	-0.0091^{***}	-0.0067^{**}
	(0.0029)	(0.0032)	(0.0029)	(0.0032)
GDP growth		-0.0025^{***}		-0.0023^{***}
		(0.0008)		(0.0008)
Stock market development		-0.0005^{***}		-0.0004^{***}
		(0.0001)		(0.0001)
International trade		0.0001		0.0001
		(0.0002)		(0.0002)
Unemployment rate		-0.0019^{*}		-0.0024^{**}
		(0.0011)		(0.0011)
Observations	1,016	790	1,016	790
R^2	0.0115	0.0728	0.0098	0.0663
Adjusted R^2	-0.0176	0.0323	-0.0193	0.0255
F Statistic	11.4858^{***}	11.8722^{***}	9.7683^{***}	10.7306^{***}

Table VIII: Regression of the $mfFDR$-based portfolio returns with use of mean return as performance metric in testing (ϕ) on the openness index and macro factors The table reports the regressions with country-fixed effects of the $mfFDR$-based portfolio's return before and after transaction cost on the openness index (Openness), GDP growth rate (GDP growth), the ratio of stock market capitalization to GDP (Stock market development), total export and import per GDP (International trade) and unemployment rate. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20% with use of mean return as performance metric (ϕ) in testing), then partitioning the selected rules into currencies. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of $10\%, 5\%$ and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Dependent variable			
	Excess Return		Net return	
Openness	-0.0146^{***}	-0.0074^{**}	-0.0141^{***}	-0.0070^{**}
	(0.0027)	(0.0034)	(0.0027)	(0.0034)
GDP growth		-0.0024^{***}		-0.0023^{***}
		(0.0008)		(0.0008)
Stock market development		-0.0005^{***}		-0.0005^{***}
		(0.0001)		(0.0001)
International trade		0.0002		0.0001
		(0.0002)		(0.0002)
Unemployment rate		-0.0014		-0.0019^{*}
		(0.0011)		(0.0011)
Observations	990	772	990	772
R^2	0.0301	0.0717	0.0282	0.0655
Adjusted R^2	0.0008	0.0302	-0.0012	0.0237
F Statistic	29.8222^{***}	11.3973^{***}	27.8083^{***}	10.3414^{***}

Table IX: Regression of the $mfFDR$-based portfolio returns with use of mean return as performance metric in testing (ϕ) on macro variables. The table reports the results of the regression with country-fixed effects of the $mfFDR^+$-based portfolio returns before and after transaction cost on macro variables. The $mfFDR$-based portfolio for each currency is formed by first implementing the $mfFDR^+$ on all rules traded on the basket of currencies (control for FDR at 20% with use of mean return as performance metric (ϕ) in testing), then partitioning the selected rules into currencies. The numbers in parentheses are the corresponding standard deviations. "*", "**" and "***" respectively indicate statistical significance at levels of 10%, 5% and 1%. The numbers in parentheses are coefficients' estimated standard deviation.

	Excess return	Net return
BOND	0.023^{*}	0.018
	(0.013)	(0.013)
INF	0.051	0.052
	(0.040)	(0.040)
DP	−0.031	−0.028
	(0.020)	(0.021)
PE	-0.00002^{**}	-0.00002^{**}
	(0.00001)	(0.00001)
MKTCAP	-0.021^{***}	-0.021^{***}
	(0.003)	(0.003)
UN	0.004	0.004
	(0.011)	(0.011)
IV	-0.081^{*}	-0.095^{**}
	(0.043)	(0.043)
ILL	0.026^{***}	0.024^{***}
	(0.004)	(0.004)
VOL	0.009^{***}	0.009^{***}
	(0.001)	(0.001)
Observations	7,230	7,230
R^2	0.031	0.029
Adjusted R^2	0.026	0.024
F Statistic (df = 9; 7196)	25.324^{***}	23.693^{***}