

Optimal transaction filters under transitory trading opportunities: Theory and empirical illustration

Ronald Balvers^{a,c,*}, Yangru Wu^{b,c}

^a*Division of Economics and Finance, College of Business and Economics, West Virginia University, Morgantown, WV 26506, USA*

^b*Rutgers Business School-Newark and New Brunswick, Rutgers University, Newark, NJ 07102, USA*

^c*Chinese Academy of Finance and Development, Central University of Finance and Economics, China*

Available online 5 August 2009

Abstract

If transitory profitable trading opportunities exist, transaction filters mitigate trading costs. We use a dynamic programming framework to design an optimal filter that maximizes after-cost expected returns. The filter size depends crucially on the degree of persistence of trading opportunities, transaction cost, and standard deviation of shocks. For daily dollar–yen exchange trading, the optimal filter can be economically significantly different from a naïve filter equal to the transaction cost. The candidate trading strategies generate positive returns that disappear after transaction costs. However, when the optimal filter is used, returns after costs remain positive and higher than for naïve filters.

© 2009 Elsevier B.V. All rights reserved.

JEL classification: G10; G15; G11

Keywords: Transaction costs; Transaction filters; Trading strategies; Foreign exchange

1. Introduction

It is inarguable that opportunities for above-normal returns are available to market participants at some level. These opportunities may be exploitable for instance at an intra-daily frequency as a reward for information acquisition when markets are efficient, or at a

*Corresponding author at: Division of Economics and Finance, College of Business and Economics, West Virginia University, Morgantown, WV 26506, USA. Tel.: +1 304 293 7880; fax: +1 304 293 2233.

E-mail addresses: rbalvers@wvu.edu (R. Balvers), yangruwu@andromeda.rutgers.edu (Y. Wu).

lower frequency to market timers when markets are inefficient. By nature these profit opportunities are predictable but transitory, and transaction costs may be a major impediment in exploiting them.¹ This paper explores the optimal trading strategy when transitory opportunities exist and transactions are costly.

The model we present is applicable to the arbitrage of microstructure inefficiencies that require frequent and timely transactions, which may be largely riskless. An example is uncovered interest speculation in currency markets where a trader takes either one side of the market or the reverse. Alternatively, a trader arbitrages differences between an asset's return and that of one of its derivatives: going long on the arbitrage position or reversing the position and going short, as is the case in covered interest arbitrage.

Alexander (1961) and Fama and Blume (1966) introduced “filter rules” according to which traders buy (sell) an asset only if its current price exceeds (is below) the previous local minimum (maximum) by more than X (more than Y) percent, where X and Y are parameters of the rule, commonly set equal and chosen in the range of 0.5–5% (e.g., Sweeney, 1986). The parameters X (and Y if different) determine a “band of inactivity” that prompts one to trade once a realization exceeds the local minimum or is below the local maximum by a certain percentage. A larger band of inactivity (larger X) filters out more trades, thus reducing transactions costs.² The general idea of filters, in filter rules, as well as other trading rules, is that if the trade indicator is weak the expected return from the transaction may not compensate for the transaction cost. Lehmann (1990) provides an interesting alternative filter by varying portfolio weights according to the strength of the return indicators—in trading smaller quantities of the assets with the weaker trade indicators, transaction costs are automatically reduced relative to the payoff.

Knez and Ready (1996) and Cooper (1999) explore different filters and find that the after-transaction-cost returns indeed improve compared to trading strategies with zero filter. The problem with the filter approach is that there is no way of knowing a priori which filter band is reasonable because the buy/sell signal and the transaction cost are not in the same units—the filter is the percentage by which the effective signal exceeds the signal at which a change in position first appears profitable before transaction costs, but this percentage bears no relation to the percentage return expected. This also implies that there is no discipline against data mining for researchers: many filters with different bands can be tried to fabricate positive net strategy returns. While Lehmann's (1990) approach provides more discipline as it specifies a unique strategy, the filter it implies is not generally optimal.

The purpose of this paper is to design an optimal filter that a priori maximizes the expected return net of transaction cost. To accomplish this we employ a “parametric” approach (e.g., Balvers et al., 2000) that allows the trading signal and the transaction cost

¹For instance, Grundy and Martin (2001) express doubt that the anomalous momentum profits survive transaction costs, and Hanna and Ready (2005) find that the momentum profits are substantially reduced when transaction costs are accounted for. Lesmond et al. (2004) conclude more strongly that momentum profits with transaction costs are illusory. Neely and Weller (2003) reach a similar conclusion for trading profits in foreign exchange markets after transaction costs.

²Note the two usages of the term filter. We distinguish “filter rules”—the specific class of technical trading rules defined above—from a more generic use of the term “filter”—a criterion for selecting a set of trades to exclude. The latter refers typically to a “band of inactivity”. Filter rules use different size filters but filters can be applied to a much broader class of trading rules that are not filter rules. In the following we examine different size filters for ARMA and moving average trading rules, but not for filter rules.

to be in the same units. In effect we convert a filter into returns space and then are able to derive the filter's optimal band. The optimal filter depends on the exact balance between maintaining the most profitable transactions and minimizing the transactions costs.

The optimal filter (band) can be no larger than the transaction cost (plus interest). This is clear because there is no reason to exclude trades that have an immediate expected return larger than the transaction cost. In general, the optimal filter is significantly smaller than the transaction cost. This occurs when the expected return is persistent: even if the immediate return from switching is less than the transaction cost, the persistence of the expected return makes it likely that an additional return is foregone in future periods by not switching. Roughly, the filter must depend on the transaction cost, as well as a factor related to the probability that a switch occurs. Our model characterizes the determinants of the filter in general and provides an exact solution for the filter when zero-investment returns have a uniform distribution.

In exploring the effect of transaction costs when returns are predictable, this paper has the same objective as [Balduzzi and Lynch \(1999\)](#), [Lynch and Balduzzi \(2000\)](#), and [Lynch and Tan \(2009\)](#).³ The focus of these authors, however, differs significantly from ours in that they consider the utility effects and portfolio rebalancing decisions, respectively, in a life cycle portfolio choice framework. They simulate the welfare cost and portfolio rebalancing decisions given a trader's constant relative risk aversion utility function, but they do not provide analytical solutions and it is difficult to use their approach to quantify the optimal trading strategies for particular applications. Our approach, in contrast, provides specific theoretical results yielding insights into the factors affecting optimal trading strategies. Moreover, our results can be applied based on observable market characteristics that do not depend on subjective utility function specifications.

In contrast to [Balduzzi and Lynch \(1999\)](#), [Lynch and Balduzzi \(2000\)](#), and [Lynch and Tan \(2009\)](#), we sidestep the controversial issue of risk in the theory. This simplifies our analysis considerably and is reasonable in a variety of circumstances. First, we can think of the raw returns as systematic-risk-adjusted returns, with whichever risk model is considered appropriate. The systematic risk adjustment is sufficient to account for all risk as long as trading occurs at the margins of an otherwise well-diversified portfolio. Second, in particular at intra-daily frequencies, traders may create arbitrage positions so that risk is irrelevant. Third, in many applications risk considerations are perceived as secondary compared to the gains in expected return; if risk adjustments are relatively small so that the optimal trading rules are approximately correct then risk corrections can be safely applied to ex post returns.

Our optimal filter will produce higher expected returns than the naïve strategies of either using no filter or using a transaction-cost-sized filter. We apply the optimal filter to a natural case for our model: daily foreign exchange trading in the yen/dollar market. As is well-known (e.g., [Cornell and Dietrich, 1978](#); [Sweeney, 1986](#); [LeBaron, 1998](#); [Gencay, 1999](#); [Qi and Wu, 2006](#)), simple moving-average trading rules improve forecasts of exchange rates and generate positive expected returns (with or without risk adjustment) in

³There is a far more extensive literature considering investment choices under transaction costs when returns are *not* predictable. See [Liu and Loewenstein \(2002\)](#) and [Liu \(2004\)](#) for recent examples. [Marquering and Verbeek \(2004\)](#) assume predictability and adjust for transaction costs. Their approach, however, is a complement to ours in that they integrate risk into the optimal switching choice while accounting for transaction costs after the fact, whereas our approach integrates transaction costs into the optimal switching choice while accounting for risk after the fact.

the foreign exchange market. However, for daily trading, returns net of transaction costs are negative or insignificant if no filter is applied (Neely and Weller, 2003).⁴

We find that for the optimal filter, the net returns are still significantly positive and higher than those when the filter is set equal to the transaction cost. Further, the optimal filter derived from the theory given a uniform distribution and two optimal filters derived numerically under normality and bootstrapping assumptions all generate similar results that are relatively close to the ex post maximizing filter for actual data. These results are important as they suggest an approach for employing trading strategies with filters to deal with transaction costs, without inviting data mining. The results also hint that in some cases the conclusion that abnormal profits disappear after accounting for transaction costs may be worth revisiting.

Section 2 develops the theoretical model and provides a general characterization of the optimal filter for an ARMA(1,1) returns process with general shocks, as well as a specific formula for the case when the shocks follow the uniform distribution. In Section 3, we apply the model to uncovered currency speculation. We show first that the moving-average strategy popular in currency trading can be related to our ARMA(1,1) specification. We then use the first one-third of our sample to develop estimates of the returns process, which we employ to calculate the optimal filter for an ARMA(1,1), an AR(1), and two representative MA returns processes. The optimal filter is obtained from the theoretical model for the uniform distribution but also numerically for the normal distribution and the bootstrapping distribution. In Section 4, we conduct the out-of-sample test with the final two-thirds of the sample to compare mean returns from a switching strategy before and after transaction costs. The switching strategies are conducted under a variety of filters, including the optimal ones, for each of the ARMA(1,1), AR(1), and MA returns cases. Section 5 concludes the paper.

2. The theoretical model

2.1. Autoregressive conditional returns and two risky assets

An individual investor maximizes the discounted expected value of an investment over the infinite horizon. There is a proportional transaction cost and the investor chooses in each period between two assets that have autocorrelated mean returns. Each period, the investor is assumed to take a zero-cost investment position: a notional \$1 long position in one asset and a notional \$1 short position in the other. This implies that any profits or losses do not affect future investment positions. The return on the zero-cost investment position x_t (a \$1 long position in asset 1 and a \$1 short position in asset 2) is assumed to follow an ARMA(1,1) process as a parsimonious parameterization of mild return predictability. Given that either the return on the investment position has no systematic risk or the investor is risk neutral, the decision problem is

$$V_1(x_{t-1}, \varepsilon_{t-1}) = E_{t-1} \left[x_t + \text{Max} \left(\frac{V_1(x_t, \varepsilon_t)}{1+r}, \frac{V_2(x_t, \varepsilon_t)}{1+r} - 2c \right) \right] \quad (1)$$

⁴Gencay et al. (2002) show that real time trading models that employ more sophisticated techniques than the simple moving average rules can generate positive cost-adjusted returns with intra-day data.

subject to

$$x_t = \rho x_{t-1} - \delta \varepsilon_{t-1} + \varepsilon_t, \quad E_{t-1} \varepsilon_t = 0, \quad \rho > \text{Max}(0, \delta). \tag{2}$$

Each period, the investor chooses whether to hold a zero-investment position long in asset 1 and short in asset 2, or the reverse. A proportional transaction cost c is incurred whenever a position is closed out, implying a cost of $2c$ when a position is reversed.⁵ In Eq. (1), the expected net present value (NPV) of the investment strategy denoted by the value function $V_i(\cdot)$ at time $t-1$ depends on the state as given by the existing long position in asset i and short position in the other asset, and the variables describing the distribution of x_t , namely x_{t-1} and ε_{t-1} . This value equals the current expected return, given the existing asset position (which we call 1 without loss of generality), plus the expected value in the next period discounted once at rate r , which depends on the updated return state variables x_t and ε_t , as well as the new zero-investment asset position ($i = 1$ or 2 , whichever maximizes the expected NPV), and minus the up-front adjustment cost incurred if the zero-investment position is switched from long in asset 1 to long in asset 2. Wealth is not a state variable, because of risk neutrality and the assumption that the long position in each period is always \$1. Assuming that the scale of investment is low enough that wealth/margin does not become an issue preserves the relative simplicity of the decision problem.

Eq. (2) describes the return process x_t for the zero-investment position long in asset 1 and short in asset 2. The ARMA(1,1) process for x_t is assumed to be *persistent* in the sense that x_{t-1} positively affects x_t ($\rho > 0$) and that ε_{t-1} positively affects x_t ($\rho > \delta$). The unconditional mean of x_t is zero and reversing the zero-cost position necessarily generates a return of $-x_t$. Define $\mu_t \equiv E_{t-1} x_t$, and let it represent, without loss of generality, the conditional expected return of the *current* zero-investment position. The disturbance term ε_t is *symmetric* and *unimodal*, with unbounded support, density $f(\varepsilon_t)$, and cumulative density $F(\varepsilon_t)$. $V(\cdot)$ denotes the maximum expected NPV of the strategy given the existing zero-investment position. Then:

Proposition 1. *For the decision problem in Eqs. (1) and (2) a unique $\mu^* < 0$ exists such that the investor maintains the current zero-investment position whenever $\mu_{t+1} > \mu^*$ and reverses the position whenever $\mu_{t+1} \leq \mu^*$. The resulting expected NPV is*

$$V(\mu_t) = \mu_t + \int_{(\mu^* - \rho\mu_t)/(\rho - \delta)}^{\infty} \left(\frac{V(\mu_{t+1})}{1+r} \right) dF(\varepsilon_t) + \int_{-\infty}^{(\mu^* - \rho\mu_t)/(\rho - \delta)} \left(\frac{V(-\mu_{t+1})}{1+r} - 2c \right) dF(\varepsilon_t) \tag{3}$$

subject to

$$\mu_{t+1} = \rho\mu_t + (\rho - \delta)\varepsilon_t \quad \text{with } E_{t-1}\varepsilon_t = 0, \quad \rho > \text{Max}(0, \delta). \tag{4}$$

Proof. See Appendix A.

Eq. (3) states that the maximum expected discounted investment return is equal to the expected return for the current position μ_t plus, either the once-discounted maximum expected NPV of the strategy when the current investment position is *maintained*, or the once-discounted maximum expected NPV of the strategy, net of adjustment costs incurred when the current investment position is *reversed*. The integral bounds indicate the critical

⁵Given the assumed risk neutrality, symmetry, and proportional transaction costs, intermediate positions, with investment in both assets or in neither asset, are never optimal.

(cutoff) value for the current disturbance term ε_t according to which the investment position is maintained or reversed (higher ε_t implies higher future expected return for the existing zero-investment position, so less incentive to switch). μ_t is a sufficient state variable for the individual investor maximizing the expected NPV from the zero-cost investment positions because Eq. (2) for the realized zero-investment return implies Eq. (4) for the conditional expected zero-investment return, and the latter is the pertinent variable for the risk-neutral expected value maximizer.

Given the benefit of switching, the difference between $V(-\mu)$ and $V(\mu)$ is monotonically decreasing in μ , there is exactly one critical value μ^* below which switching is optimal. Intuitively μ^* must be negative because it makes no sense to switch when μ_{t+1} is zero or positive. Our purpose is to provide in the following a specific characterization of μ^* for empirical purposes.

To examine the specific advantage from switching the asset position, take the difference between the value of long in one asset and short in the other after a switch, $V(-\mu)$, and the initial position, $V(\mu)$:

$$V(-\mu_t) - V(\mu_t) = -2\mu_t + B(\mu_t) + C(\mu_t). \tag{5}$$

Eq. (5) follows directly from applying Eq. (3) twice, for $-\mu_t$ and μ_t , and manipulating the integrals (as shown in Appendix A). The first term on the right-hand side indicates the direct benefit of switching from μ_t to $-\mu_t$. The second term, $B(\mu_t)$, is given as

$$B(\mu_t) = \int_{(\mu^* - \rho\mu_t)/(\rho - \delta)}^{(-\mu^* - \rho\mu_t)/(\rho - \delta)} \left(\frac{V(-\mu_{t+1}) - V(\mu_{t+1})}{1 + r} \right) dF(\varepsilon_t). \tag{6}$$

$B(\mu_t)$ represents the future benefit of switching now, given that neither of the two possible positions (the current available choices) would be switched in the next period. The integral bounds equal the critical values for ε_t at which a switch takes place (lower bound for switching next period when having switched in the previous period; upper bound for switching next period when not having switched previously). Use Eq. (4) to see that the difference $V(-\mu_{t+1}) - V(\mu_{t+1})$ on the right-hand side of Eq. (6) is evaluated from $+\mu^*$ to $-\mu^*$. The symmetry of this difference further implies that the sign depends only on the shape of the density function. In particular, taking $\mu^* < 0$ (which is always so, as we show shortly), if $\mu_t < 0$ then the range over which $V(-\mu_{t+1}) - V(\mu_{t+1})$ is negative (whenever $\mu_{t+1} > 0$) is weighted less than the range over which it is positive, because the density is symmetric and unimodal (vice versa if $\mu_t > 0$). Thus $B(\mu_t)$ is always strictly positive for $\mu_t < 0$, unless the density function is flat (as is the case for the uniform distribution). The third term in Eq. (5), $C(\mu_t)$, is given as

$$C(\mu_t) = 2c \int_{(\mu^* + \rho\mu_t)/(\rho - \delta)}^{(\mu^* - \rho\mu_t)/(\rho - \delta)} dF(\varepsilon_t) = 2c \left[F\left(\frac{\mu^* - \rho\mu_t}{\rho - \delta}\right) - F\left(\frac{\mu^* + \rho\mu_t}{\rho - \delta}\right) \right]. \tag{7}$$

It gives the future benefit of switching now, given that for one of the two possible positions it will be optimal to switch in the next period (note that it is never optimal for both positions to switch). In that case, both possible positions end up becoming identical one period later and the only difference is the transaction cost. $C(\mu_t)$ represents the expected difference in transaction cost expenditure for the two positions: the cost times the probability of not switching initially but switching anyway next period minus the

probability of switching initially and then switching back in the next period. Note that, from inspection of Eq. (7), $C(\mu_t)$ must be positive for $\mu_t < 0$ (and negative for $\mu_t > 0$).

To determine the critical value μ^* we evaluate Eq. (5) for $\mu_t = \mu^*$. First consider that μ^* is determined optimally. Since the integral bounds are chosen optimally to obtain the maximum expected NPV from the investment strategy, it must be that the derivative with respect to the bound in Eq. (3) is zero. Using Leibniz’s rule to obtain the first-order condition for μ^* in Eq. (3) gives

$$\frac{V(-\mu^*)}{1+r} - \frac{V(\mu^*)}{1+r} = 2c. \tag{8}$$

Eq. (8) reveals that the critical return is determined at the point where the total present value of expected benefit from reversing the investment position is exactly equal to the upfront transaction cost: at $\mu_{t+1} = \mu^* < 0$ the investor is indifferent between maintaining the current investment position with negative expected return μ^* next period and reversing the investment position that has an immediate transaction cost of $2c$ but a positive expected next-period return $-\mu^*$. It is directly clear from Eq. (8) that μ^* must be constant over time, as its notation without time subscript presumes.

Evaluate Eq. (5) at $\mu_t = \mu^*$ and use Eq. (8) to obtain⁶:

$$\mu^* = -c(1+r) + \frac{1}{2}[B(\mu^*) + C(\mu^*)]. \tag{9}$$

Intuitively, Eq. (9) sets the critical expected return μ^* equal to the benefit of switching positions at μ^* minus the up-front transaction cost (including interest) from switching. The benefit of switching at the expected return μ^* includes the cost saving of a reduction in the probability of switching next period, $C(\mu^*)$, as well as the expected immediate revenue from the improved asset position, $B(\mu^*)$. As $B(\mu^*) + C(\mu^*) > 0$ from Eqs. (6) and (7), given that $\mu^* < 0$, it follows that $-\mu^* < c(1+r)$. This is intuitive because, if the current expected return is greater than or equal to the transaction cost, a switch is clearly beneficial because the end-of-period gain in expected return $-2\mu^*$ immediately pays for the transaction cost $2c(1+r)$ while the switch also improves the potential for future profits.

2.2. Closed form solution for uniform innovations

It is difficult to obtain an explicit analytical solution for the optimal filter in Eq. (9) because, from Eq. (6), $B(\mu^*)$ depends on the value function, which is of unknown functional form. However, for the special case of a constant density over the relevant range (a uniform distribution), the $B(\mu^*)$ term simplifies substantially, as shown in the following, so that an explicit solution can be obtained.

Assume a uniform distribution for ε over the interval $[-z, z]$ with implied density $f(\varepsilon) = 1/(2z)$. To apply to this case, Proposition 1 requires minor modifications, which we omit for brevity, since the uniform distribution is bounded and not strictly unimodal.

⁶Given Eq. (8) or Eq. (9) characterizing the critical mean return and Eq. (3) stating the value function, it is possible to derive the general comparative statics results for the critical expected return and the probability of no transaction with respect to all of the parameters in the model of Eqs. (3) and (4). These results are available from the authors.

Eq. (6) evaluated at μ^* becomes:

$$B(\mu^*) = \frac{1}{2z(1+r)} \int_{\text{Max}\{-z, \mu^*(1-\rho)/(\rho-\delta)\}}^{\text{Min}\{z, -\mu^*(1+\rho)/(\rho-\delta)\}} (V[-\rho\mu_t - (\rho - \delta)\varepsilon_t] - V[\rho\mu_t + (\rho - \delta)\varepsilon_t]) d\varepsilon_t. \tag{10}$$

Eq. (7) evaluated at μ^* becomes⁷:

$$C(\mu^*) = \frac{c}{z} \int_{\text{Max}\{-z, \mu^*(1+\rho)/(\rho-\delta)\}}^{\text{Min}\{z, \mu^*(1-\rho)/(\rho-\delta)\}} d\varepsilon_t = \frac{c}{z} \left[\text{Min}\left(z, \frac{\mu^*(1-\rho)}{\rho-\delta}\right) - \text{Max}\left(-z, \frac{\mu^*(1+\rho)}{\rho-\delta}\right) \right]. \tag{11}$$

Proposition 2. *If ε is uniformly distributed over the interval $[-z, z]$ and if:*

$$z \geq \frac{[1+r(1+\rho)]c}{\rho-\delta}, \tag{12}$$

then the critical expected mean return in the model of Eqs. (1)–(2) is given as

$$\mu^* = \frac{-(1+r)c}{1+\rho c/[(\rho-\delta)z]}. \tag{13}$$

Proof. If $z \geq -(1+\rho)\mu^*/(\rho-\delta)$, then the bounds in $B(\mu^*)$ and $C(\mu^*)$ are interior. Hence $B(\mu^*) = 0$ given the constant density and the symmetry in Eq. (10), and $C(\mu^*) = -2c\rho\mu^*/[(\rho-\delta)z]$ from Eq. (11). Eq. (9) then implies Eq. (13). Eqs. (12) and (13) in turn imply the premise that $z \geq -(1+\rho)\mu^*/(\rho-\delta)$. □

Note that from Eq. (9) the assumption of the uniform distribution may lead to a more strongly negative value for μ^* because it causes $B(\mu^*)$ to be equal to its minimum value of zero. Eq. (12) guarantees that the extreme realizations for the return innovations are large enough in absolute value that, in the range $\mu_t \in [-\mu^*, \mu^*]$, a sufficiently large innovation can always occur for which a switch becomes immediately optimal for one of the two possible zero-investment positions. This assumption in the uniform distribution case replaces the assumption of unbounded support in the general case.

Eq. (13) states that, given $(\rho-\delta)$ fixed, the absolute value of the optimal filter as a fraction of the transaction cost (plus interest) depends negatively on the persistence of the mean of the return process, ρ : the trader should be willing to switch his position more readily toward a profitable opportunity if it is likely to persist longer. Considering Eq. (4) and the fact that for the uniform distribution $z = \sqrt{3}\sigma_\varepsilon$, the optimal filter depends positively on the variability of the mean of the return process, $(\rho-\delta)z$, scaled by the transactions cost, c : if the mean is highly variable compared to the transaction cost, then a trader should require a higher immediate expected return before switching since there is a higher chance that he may want to switch back soon.

It is instructive to evaluate a few special cases in Eq. (13). First, if we have a pure MA(1) process, $\rho = 0$, then the optimal filter becomes the naïve filter (plus interest), $\mu^* = -c(1+r)$. The reason is that there is no persistence in the profit opportunities

⁷However, $C(\mu^*) = 0$ when $\mu^*(1-\rho) < -(\rho-\delta)z$. This condition and the Min and Max operators in Eqs. (10) and (11) appear because $dF(\varepsilon) = 0$ outside of the domain $[-z, z]$ but $d\varepsilon$ is not. In the following we will provide a condition to avoid these additional cases.

because the current conditional expected return does not depend on past conditional expected returns. Second, if we have a pure AR(1) process, $\delta = 0$, then the optimal filter becomes $\mu^* = -c(1+r)/[1+(c/z)]$, which is closer to zero than in the MA(1) case, implying that a switch occurs more readily (because the conditional expected return is persistent). Interestingly, the filter in the AR(1) case does not depend on ρ . The explanation is that, as ρ increases, two offsetting effects occur: due to the $\rho\mu_t$ term in Eq. (4), the expected return becomes more persistent, making the critical expected return *less* negative, but due to the $\rho\varepsilon_t$ term in Eq. (4) the expected return also changes more rapidly, making the critical expected return *more* negative.

3. Empirical illustration for foreign exchange trading: optimal filter calculation

The model developed in the preceding section can be interpreted in three different ways. First, if we ignore the underlying x_t process in Eq. (2), the μ_t in Eq. (4) may be interpreted as an excess return that is fully known at time t . Thus, we are dealing with a case of pure arbitrage where the trader optimizes the after-transaction-cost return $V(\mu_t)$. Second, we may think of μ_t as the risk-adjusted expected return, the “alpha”, so that $V(\mu_t)$ represents the expected NPV adjusted for systematic risk. Similarly, μ_t could represent a particular expected utility level, which would also account for risk. Third, we can interpret μ_t as an expected return in a case where risk is relatively small or non-systematic. In this case, an appropriate risk correction can simply be applied to the ex post returns. A minor drawback is that the optimal filter has to be applied with unadjusted returns, but this is not a major issue if the risk adjustment is small and would anyway bias results away from finding positive trading returns.⁸

Empirically, it is difficult to find accurate data to examine the first interpretation, while the second interpretation requires employing a particular risk model. Accordingly, we adopt the third interpretation of the theory in considering uncovered interest speculation in the dollar–yen spot foreign exchange market.

3.1. A parametric moving average trading strategy

As discussed extensively in the literature (e.g., Cornell and Dietrich, 1978; Sweeney, 1986; Frankel and Froot, 1990; Levich and Thomas, 1993; Lee and Mathur, 1996; LeBaron, 1998, 1999; Qi and Wu, 2006; and others), profitable trading strategies in foreign exchange markets traditionally have employed moving-average (MA) technical trading rules.⁹ MA trading rules of size N work as follows: calculate the moving average using N lags of the exchange rate. Buy the currency if the current exchange rate exceeds this average; short-sell the currency if the current exchange rate falls short of this average.

Defining s_t as the log of the current-period spot exchange rate level (dollar price per yen) and Δs_t as the percentage appreciation of the yen, the implicit exchange rate forecasting

⁸Neely (2003) finds that optimizing rules based on an ex ante risk criterion provide substantially different results compared to results adjusted for risk ex post. He notes, however, that the ex ante risk adjustment implies higher (adjusted) returns from technical trading.

⁹The source of the excess returns from MA strategies in foreign exchange markets may be due to central bank intervention designed to smooth exchange rate fluctuations (e.g., Sweeney, 2000; Taylor, 1982). However, recent results by Neely and Weller (2001) and Neely (2002) argue against this perspective.

model behind the MA trading rule is

$$\Delta s_{t+1} = \lambda \left(s_t - \left[\left(\sum_{i=1}^N s_{t-i} \right) \right] / N \right) + \varepsilon_{t+1}, \quad E_t \varepsilon_{t+1} = 0, \quad \lambda > 0. \quad (14)$$

For any positive λ , Eq. (14) implies a positive expected exchange rate appreciation if the log of the current exchange rate exceeds the N -period MA. Empirically, we find λ to be positive in all our specifications. Hence, the decision rule based on Eq. (14) to buy (short) the currency if the expected appreciation is positive (negative) leads to a trading strategy equivalent to the MA trading strategy for any positive value of λ .¹⁰ In what follows, our strategy is to transform Eq. (14) to infer empirically a value for λ which, in addition to implying identical switch points as the MA rule, also provides a quantitative estimate of the expected gain from switching that may be compared directly to the transactions cost.

The true distribution of the exchange rate can be very complex and the simple MA process in Eq. (14) can only be an approximation of the true exchange rate process. We are motivated to use the MA rule because it is the most popular rule studied by researchers and used by practitioners. It is important to emphasize here that the primary purpose of this paper is not *per se* to search for the best exchange rate forecasting model to generate trading profitability, or to explain the potential profitability from certain trading strategies. But rather, the key point we want to make is that, given a data-generating process that exhibits some return predictability, an optimal transaction filter can be designed to maximize the after-cost expected profitability of a particular trading strategy. The optimal filter size is conditioned on a specific exchange rate forecasting model and can be easily and unambiguously computed using prior data. It is shown empirically below that the optimal filter can be significantly smaller than the naïve filter equal to the transaction cost. The optimal filter will in general outperform the naïve filters regardless of the specific return-generating process assumed.

Eq. (14) can straightforwardly be rewritten as an autoregressive process in the percentage change in the exchange rate, with the coefficients in the autoregression given by the Bartlett weights:

$$\Delta s_{t+1} = \lambda \left(\sum_{i=0}^{N-1} \frac{N-i}{N} \Delta s_{t-i} \right) + \varepsilon_{t+1}, \quad (15)$$

Typical studies on technical analysis of foreign exchange do not utilize information on interest rates in computing the moving averages and do not estimate a parametric model for forecasting. To be fully consistent with our theory, we want to treat the excess return as the variable to be forecasted in the forecasting equation. To do so, we add the interest rate differential to the percentage change in exchange rate, so that Eq. (15) becomes:

$$x_{t+1} = \lambda \left(\sum_{i=0}^{N-1} \frac{N-i}{N} x_{t-i} \right) + \varepsilon_{t+1}, \quad (16)$$

¹⁰Given Eq. (14), there is a clear link between the popular MA filters with a band and our filter. An exchange rate such that the moving average exchange rate is exceeded by x percent induces a switch. In our case, for a naïve filter equal to c , for instance, we need $\lambda x > 2c$ to induce a switch. So, for the numbers we find in our empirical section for the MA(21) process, our naïve filter corresponds to $x = 2c/\lambda = 0.1/0.025 = 4$, a 4% band for the ad hoc MA filter.

where $x_t \equiv \Delta s_t + r_{t-1}^{JP} - r_{t-1}^{US}$, r_t^{JP} is the daily Japanese interest rate, and r_t^{US} is the daily U.S. interest rate. In other words, x_t denotes the excess return from buying the Japanese yen and shorting the U.S. dollar (or the deviation from uncovered interest parity).¹¹

In the presence of transaction costs, the MA rule needs to be supplemented with a filter that indicates by how far the current spot rate must exceed (or fall short of) the MA in order to motivate a trade. The advantage of Eq. (16) is that it is parametric and, given an estimate for λ , can provide a quantitative measure of the filter based on comparing expected return to the transaction cost.

To obtain analytically the optimal filter for the MA criterion in the context of the model, it is necessary that Eq. (16) be translated to ARMA(1,1) format, for which the model provides the optimal filter (which is in closed form expression if the ε_t are uniformly distributed). LeBaron (1992), Taylor (1992), and others have shown that an ARMA(1,1) well replicates moving average trading rule results. The theoretical restriction on the ARMA(1,1) process in Eq. (2) that $\rho > \text{Max}(0, \delta)$ is satisfied for all our empirical specifications. Inverting the ARMA(1,1) process yields an alternative autoregression:

$$x_{t+1} = (\rho - \delta) \sum_{i=0}^{\infty} \delta^i x_{t-i} + \varepsilon_{t+1}. \quad (17)$$

Comparing Eqs. (16) and (17), it follows that Eq. (17) is a good approximation for Eq. (16) if we set $\rho - \delta = \lambda$ and if the $(N - i)/N$ terms are close to δ^i for all i . Taking a natural log approximation, and choosing δ to match the Bartlett weights:

$$-i/N \approx \ln[(N - i)/N] = i \ln(\delta) \Rightarrow \delta = \exp(-1/N). \quad (18)$$

Thus, for a given N , we run regression (16) to obtain an estimate of λ . Then from $\delta = \exp(-1/N)$ and $\rho = \lambda + \delta$, we can uniquely identify a δ and ρ that provide a good ARMA(1,1) proxy for an MA-based process with a relatively large N . In turn, the ARMA parameters allow us to calculate analytically the critical expected return μ^* governing the transaction choice.

3.2. Preliminary dollar–yen process estimates

Our data on the Japanese yen–U.S. dollar spot exchange rate cover the period from August 31, 1978 to May 3, 2003 with 6,195 daily observations. Daily exchange rate data for the Japanese yen are downloaded from the Federal Reserve's website. For interest rates we obtain the *Financial Times'* Euro-currency interest rates from *Datastream International*. The first third of the sample (2,065 observations) is used for model estimation. Out-of-sample forecasting starts on November 28, 1986 until the end of the sample (4,130 observations). We estimate the exchange rate dynamics in four ways: (1) an ARMA(1,1) process (Eq. (2)); (2) an AR(1) process (δ in Eq. (2) is set equal to zero); (3) a process consistent with an MA rule of 21 lags (21 trading days in a month), as is commonly considered with daily data (Eq. (16)); and (4) a process consistent with an MA rule of size 126 (half a year), around the size often used by traders, although results appear to depend little on the size of the MA process chosen (LeBaron, 1998, 1999). We choose these four models as our empirical illustration for the following reasons. Model 1, the ARMA(1,1), is

¹¹ Deviations from uncovered interest parity and their persistence are documented in previous research (e.g., Canova, 1991; Engel, 1996; Mark and Wu, 1998).

Table 1

Model parameters and implied optimal transaction cost filters.

This table reports parameter estimates for candidate forecasting models and for each model the implied optimal filter/transaction cost ratios. The forecasting models are ARMA(1,1): $x_{t+1} = \rho x_t + \varepsilon_{t+1} - \delta \varepsilon_t$; AR(1): $x_{t+1} = \rho x_t + \varepsilon_{t+1}$; and the transformed model of MA(N):

$$x_{t+1} = \lambda \sum_{i=0}^{N-1} \left[\frac{N-i}{N} x_{t-i} \right] + \varepsilon_{t+1},$$

with $N = 21$ and 126 , where $x_t \equiv \Delta s_t + r_{t-1}^{JP} - r_{t-1}^{US}$, s_t is the log of the U.S. dollar price of one Japanese yen, r_t^{JP} is the daily Japanese interest rate, and r_t^{US} is the daily U.S. interest rate. The full sample data cover the period from August 31, 1978 to May 3, 2003 with 6,195 daily observations. The parameters are estimated with the first third of the sample (2,065 observations). The implied optimal transaction cost filters are calculated under three differential distributional assumptions of ε_t : uniform where the optimal filter denoted by μ_U^* is calculated using Eq. (13); normal, where the optimal filter denoted by μ_N^* is estimated through Monte Carlo simulation with 500 replications; and bootstrap, where the optimal filter denoted by μ_B^* is estimated through bootstrapping with replacement with 500 replications. The round-trip transaction cost $2c = 0.1\%$. Numbers inside parentheses are t -ratios.

(1)	ρ (2)	δ (3)	λ (4)	σ_ε (5)	$-\mu_U^*/c$ (6)	$-\mu_N^*/c$ (7)	$-\mu_B^*/c$ (8)
ARMA(1,1) model	0.918 (19.232)	0.880 (15.665)		0.00658	0.32	0.34	0.32
AR(1) model	0.0548 (2.494)			0.00659	0.92	0.88	0.85
MA(21) model			0.0257 (3.846)	0.00656	0.23	0.24	0.23
MA(126) model			0.00818 (3.518)	0.00635	0.083	0.12	0.12

the exact model assumed in our theoretical derivation, while Model 2, the AR(1) model, is a more parsimonious specification that is an interesting special case. Models 3 and 4 are commonly employed by academia and practitioners.

Columns (1)–(5) of Table 1 show the results of the in-sample model estimation using the first third of our sample for each way of capturing the exchange rate dynamics. For the ARMA(1,1) process, we find $\rho = 0.918$ and $\delta = 0.880$, with a standard error of $\sigma_\varepsilon = 0.00658$; for the AR(1) process, we find low persistence with $\rho = 0.0548$ and a standard error of $\sigma_\varepsilon = 0.00659$. Thus, both processes provide similar accuracy although the parameters differ substantially.¹² While the data cannot tell us clearly whether the AR(1) or the ARMA(1,1) process is better at describing the exchange rate dynamics, we will see that the implications for optimal trading are substantially different.^{13,14} For the

¹²Previous more elaborate research on this issue by LeBaron (1992) and Taylor (1992) on the other hand finds that, while the ARMA(1,1) formulation does well, the AR(1) case is much poorer in replicating the key features of exchange rate series.

¹³Balvers and Mitchell (1997) raise a similar issue in the context of optimal portfolio choice under return predictability.

¹⁴Charles Engel pointed out to us the apparent puzzle that the ARMA(1,1) model $x_t - 0.918x_{t-1} = \varepsilon_t - 0.880\varepsilon_{t-1}$ implies a substantially different optimal filter from our AR(1) model, $x_t = 0.055x_{t-1} + \varepsilon_t$. This is surprising, given that the AR(1) model can be rewritten as an ARMA(2,1) model $x_t - 0.918x_{t-1} + 0.047x_{t-2} = \varepsilon_t - 0.863\varepsilon_{t-1}$, which is very similar to the ARMA(1,1) model. The AR(1) and ARMA(1,1) models are similar in that they unconditionally describe the data with roughly the same degree of precision as indicated by

representative moving average rule with 21 lags, we find for the slope in Eq. (16) that $\lambda = 0.0257$ and $\sigma_\varepsilon = 0.00656$. Since we have $N = 21$, we obtain from Eq. (18) that $\rho = 0.979$ and $\delta = 0.954$, which is not statistically distinguishable from the direct estimates from the ARMA(1,1) model. The MA(126) process yields $\lambda = 0.00818$ and $\sigma_\varepsilon = 0.00635$, implying by approximation that $\rho = 0.999$ and $\delta = 0.992$.

We assume one *round trip* transaction cost of $2c = 0.001$ (10 basis points) per switch throughout.¹⁵ Sweeney (1986) finds a transaction cost of 12.5 basis points for major foreign exchange markets, but more recent work by Bessembinder (1994), Melvin and Tan (1996), and Cheung and Wong (2000) finds bid–ask spreads for major exchange rates between 5 and 9 basis points. To account for transaction costs in addition to those imbedded in the bid–ask spread, related to broker fees and commissions and the lending–borrowing interest differential, we use 10 basis points as a realistic number for the dollar–yen market. The daily U.S. interest rate is on average over the first third of the sample equal to 0.000439%. This average rate is used as a proxy for the discount rate r in computing the optimal filter in Eq. (13).

The true distribution of the exchange rate can be quite complex, and we do not know a priori which distributional assumption is the best approximation. Therefore we choose to estimate the optimal filter μ^* using three methods. First, under the assumption that the error term ε_t is uniformly distributed, the optimal filter, denoted μ_U^* , can be analytically calculated using Eq. (13). Second, ε_t is assumed to follow a normal distribution. In this case, the result in Eq. (13) no longer holds, and we estimate the optimal filter, denoted μ_N^* , through Monte Carlo simulation. Last, we do not make an assumption about the distribution of ε_t and estimate the optimal filter, denoted μ_B^* , by bootstrapping the model residuals $\hat{\varepsilon}_t$ with replacement.

3.3. The optimal filter implied by the theory under the uniform distribution

Under the assumption of a uniform distribution, we can obtain z from the relation $z = \sqrt{3}\sigma_\varepsilon$. All the information now is there to allow us to calculate the optimal filter from Eq. (13) for the dollar–yen exchange rate. Column (6) of Table 1 provides the results. For the AR(1) case, we find that the ratio of the critical return to the transaction cost is $-\mu_U^*/c = 0.92$.¹⁶ Hence, in this case the optimal filter is not very different from a naïve filter that equals the transaction cost c . The main reason is that, from Eq. (4), the

(footnote continued)

σ_ε in our Table 1. The reason they imply quite different filters is because the optimal filter depends on the persistence of the *conditional expected* return. Eq. (4) says that the conditional expected return is an AR(1) process, which becomes $\mu_t = 0.918\mu_{t-1} + (0.918 - 0.880)\varepsilon_{t-1} = 0.918\mu_{t-1} + 0.038\varepsilon_{t-1}$ for our empirical ARMA(1,1) model, while our empirical AR(1) model (or the equivalent ARMA(2,1)) implies $\mu_t = 0.055\mu_{t-1} + 0.055\varepsilon_{t-1}$. We can see that the *conditional expected* return from the ARMA(1,1) model is much more persistent than that from the AR(1) model. Therefore, we apply a smaller filter for the ARMA(1,1) model since even a small positive expected one-period return is enough to make up for the transaction cost because the new position is likely to remain optimal for a number of future periods.

¹⁵Note that c in the model represents the cost of closing out a zero-investment position. For most assets this requires both a sale and a purchase implying a round trip transaction cost. However, in the case of foreign exchange, we directly purchase one currency with the other, implying only a one-way transaction cost. Reversing the position requires double that transaction cost $2c$, which is one round trip.

¹⁶This ratio is the immediate gain of switching to $-\mu^*$ from μ^* (the gain of $-2\mu^*$) divided by the transaction cost $2c$.

persistence in the mean return is small at $\rho = 0.0548$ so that, no matter what the current holdings are, there is not much difference in future probabilities of trading.

For the 1-month MA process, the parameters backed out from the MA(21) model yield $-\mu_U^*/c = 0.23$. Note that inequality (12) is violated, as is necessary here when $-\mu_U^*/c < 0.50$, implying that the analytical value obtained from Eq. (13) is no longer accurate and must be viewed as a good approximation; hence, it is more precise to state that $-\mu_U^*/c \approx 0.23$. Intuitively, the slow adjustment in the conditional mean for these parameter values implies that, in some cases, even the most extreme realization of the exchange rate innovation would not be sufficient to induce switching. Hence, one would be certain of avoiding transaction costs for at least one period (and likely more) by buying/keeping the exchange with the positive expected return. This explains the low value of the critical expected return relative to the transaction cost.

The 6-month MA process, MA(126), yields the smallest filter, $-\mu_U^*/c = 0.083$. One reason is the high persistence of expected return (the implied persistence parameter $\rho = 0.999$). Another is the fact that by nature the long MA process is very smooth so changes in the mean occur very slowly so that the number of transactions is small, even when there is no filter. This is a possible reason for the popularity of this particular trading rule with practitioners.

For our main specification, the ARMA(1,1) case with $\rho = 0.918$ and $\delta = 0.880$, we find that the optimal filter is $-\mu_U^*/c = 0.32$. The reason that this number is so much lower than under the AR(1) case is clear from Eq. (4). The persistence is not only high now with $\rho = 0.918$ but it is also high relative to the innovation in the conditional mean, given by $(\rho - \delta)\sigma_\varepsilon = 0.038\sigma_\varepsilon$. Hence, it is highly likely that the exchange position (dollar or yen) with the currently positive expected return is going to be unchanged in the nearby future.

3.4. The optimal filter obtained numerically

As a check on the dependence of the results on the uniform distribution, we also find the optimal filter numerically using a Monte Carlo approach, assuming normality and a bootstrapping approach.

For each Monte Carlo trial, we simulate expected returns μ_t using Eq. (16) with parameters estimated from the first third of the sample. We then choose the filter μ_N^* , which maximizes the after-cost average excess return. This process is replicated 500 times. Column (7) of Table 1 reports the median value of the optimal filter to transaction cost ratio, $-\mu_N^*/c$, over the 500 Monte Carlo trials. For each model, the ratio $-\mu_N^*/c$ is quite close to the optimal ratio implied under the uniform distribution, $-\mu_U^*/c$, with the difference between them never exceeding 5% of the transaction cost.

The actual distribution of ε_t may be neither uniform nor normal. In this case, we re-sample with replacement the fitted residuals $\hat{\varepsilon}_t$ of Eq. (16) and use model parameters to generate expected return observations μ_t . Similar to the Monte Carlo experiment, for each *bootstrapping* trial, the optimal filter is chosen to be the one that maximizes the after-cost average excess return. Column (8) of Table 1 reports the median estimate of the optimal filter to cost ratio $-\mu_B^*/c$ over 500 bootstrapping replications. Encouragingly, the optimal filters, for the theoretical uniform distribution case and the numerical normal and bootstrapping cases, are quite similar for each of the returns processes. Thus, the optimal filter value is robust to distributional assumptions.

Applying the optimal filters to the data in trading on deviations from uncovered interest parity, we expect straightforwardly that the optimal filters will outperform naïve filters. In particular, we expect that the optimal filter does better than the naïve “0” filter that is used implicitly when transaction costs are ignored for trading decisions (but not for calculating returns), because it saves on transaction costs; and does better than the naïve “ c ” filter that is employed when trading costs are considered myopically, because it does not filter out as many profitable transactions.

3.5. The filter obtained empirically

An alternative, non-theoretical procedure would establish potential trading strategy profits after transaction costs based on a prior data sample and for a grid of ratios of the (negative of the) critical expected return to the transaction cost. We then smooth the graph of the after-transaction-cost trading strategy profits over the grid and choose the ratio, $-\mu_E^*/c$, that maximizes the ex post smoothed profits for the *prior* sample and apply this filter to out of sample data. This empirical procedure uses the in-sample part of the data more intensively, not just to estimate the time-series parameters but also the empirically optimal filter. However, the chosen filter depends on the smoothing procedure, and the intensive use of the in-sample data may hurt if that sample is not representative enough of the full data.

As before, we use the first third of the sample as the prior sample. We use the first half of the prior sample to obtain the initial model estimates and then roll forward to the end of the prior sample, calculating the after-cost profits for a grid of values of $-\mu/c$ over the feasible range $[0, 1.2]$ in increments of 0.02. The profit function is then smoothed using a fifth-order moving average (the average of 2 leads, the current value, and 2 lags) filter and we pick the value of $-\mu/c$ maximizing the smoothed profits as the filter $-\mu_E^*/c$, to be applied out of sample.

The filters, $-\mu_E^*/c$, obtained in this way are shown in Figs. 1–4 (we do not report the accompanying results in Table 2 to save space).¹⁷ These filters are relatively close to the theoretical filters for the ARMA(1,1) case and the two MA cases, and generate approximately the same after-cost trading profits for these cases (albeit a bit lower for the MA(21) case). The empirical filter is quite different for the AR(1) case, around 0.4 as compared to 0.9 for the theoretical cases, and generates zero after-cost profits. However, the choice of empirical filter depends crucially on the smoothing method. For instance in this AR(1) case, more smoothing generates an empirical filter above 0.8.

3.6. Which optimal filter to use in practice?

The above results demonstrate that the distributional assumptions make very little difference in practice in determining the optimal filter. The robustness of our theoretical result in Proposition 2 with respect to the choice of innovation distribution is reassuring since a uniform distribution is not exactly a good proxy for many return distributions in practice. A reason for the apparent irrelevance of this assumption may be that, in this application, we implicitly have that the first three distribution moments are the same for

¹⁷The smoothed and unsmoothed graphs of the after-cost profits for the four cases (ARMA(1,1), AR(1), MA(21), and MA(126)) are available from the authors.

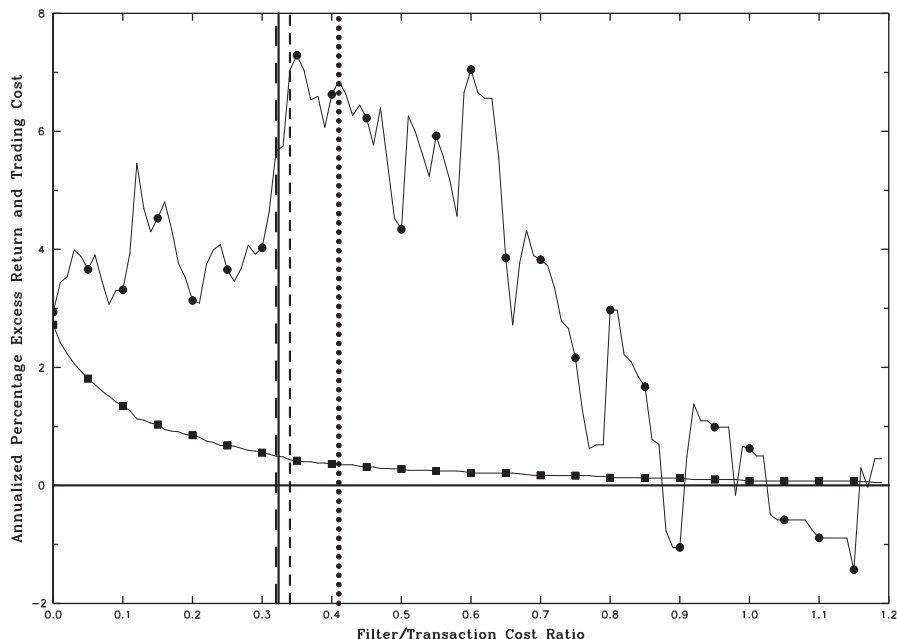


Fig. 1. The effect of filter on after-cost returns for the ARMA(1,1) model. *Note:* The solid line with round symbols displays the after-cost excess return and the one with square symbols displays the trading cost. The vertical lines represent different filters: the solid line is the theoretical optimal filter under uniform distribution $-\mu_U^*/c$; the short dashed line is the optimal filter under normal distribution $-\mu_N^*/c$; the long dashed line is the optimal filter under bootstrap distribution $-\mu_B^*/c$; and the dotted line is the ad hoc filter that maximizes the in-sample smoothed after-cost excess return using the first third of the sample, $-\mu_E^*/c$.

any innovation distribution. This occurs because the mean of the distribution is set to zero in each case, the variance is calibrated to be the same in each case, and the skewness is similar in each case because of the symmetry in the returns of zero-investment positions relative to their reverse.

The effective similarity of different return distributions in this case together with the obtained similarity in the results in Table 2 for the different distributions both argue for using Proposition 2, which holds for the uniform distribution only, in practice. Calculation of the optimal filter can be performed directly from Eq. (13) and does not require simulations. Given the transaction cost and riskless interest rate, only a prior estimate of the ARMA(1,1) process is needed, yielding the autoregressive parameter, ρ , the moving average parameter, δ , and the estimate of the innovation variance, σ_ε^2 (the latter determining z from $z = 3^{1/2}\sigma_\varepsilon$). Note also that the alternative models, the AR(1) and the MA models, are simply special cases of our model that we have employed to illustrate the implications of particular trading rules often used in foreign exchange markets. Based also on the results of LeBaron (1992) and Taylor (1992), who find that this specification is preferred with data prior to our holdout sample, it should be clear that the ARMA(1,1) process assumed in the theoretical model is the preferred specification for application. Accordingly, for given prior and holdout samples, the optimal filter can be obtained simply and uniquely.

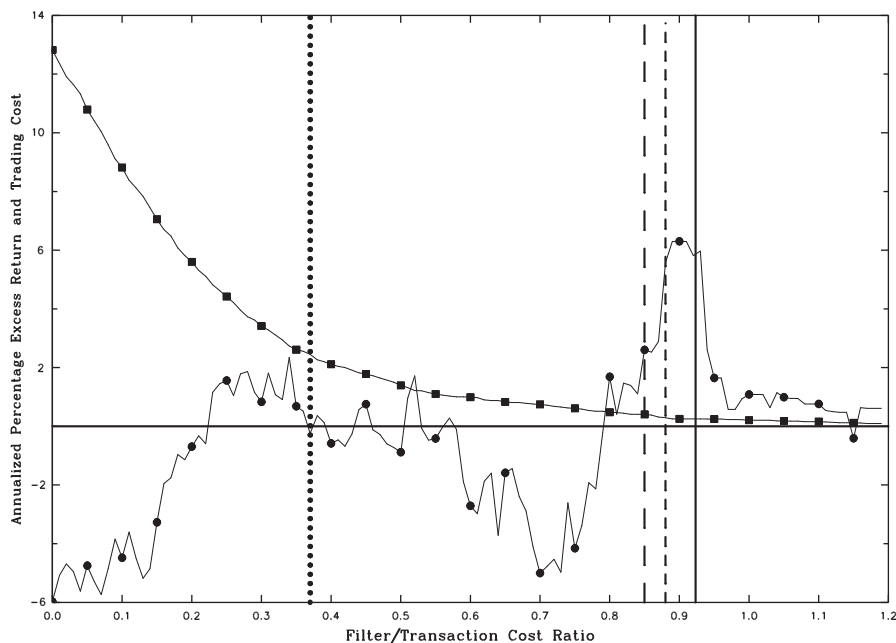


Fig. 2. The effect of filter on after-cost returns for the AR(1) model. *Note:* The solid line with round symbols displays the after-cost excess return and the one with square symbols displays the trading cost. The vertical lines represent different filters: the solid line is the theoretical optimal filter under uniform distribution $-\mu_U^*/c$; the short dashed line is the optimal filter under normal distribution $-\mu_N^*/c$; the long dashed line is the optimal filter under bootstrap distribution $-\mu_B^*/c$; and the dotted line is the ad hoc filter that maximizes the in-sample smoothed after-cost excess return using the first third of the sample, $-\mu_E^*/c$.

4. Out-of-sample optimal switching strategy results

4.1. Description of the strategy

We start our first-day forecast on November 28, 1986 (after the first third of the sample).¹⁸ For each of the four exchange rate return specifications, we estimate the model parameters using all observations for the first third of the sample (up to November 27, 1986) and make the first forecast (for November 28). If the forecasted return (recall that the return is defined as the difference between the return from holding the Japanese yen, which is the percentage exchange rate change plus the one-day Japanese interest rate, and the return from holding the U.S. dollar, which is the one-day U.S. interest rate) is positive, we take a long position in the yen, and simultaneously take a short position in the dollar. Conversely, if the forecasted return is negative, we take a long position in the dollar and a short position in the yen. The difference in returns between the long and short positions represents the return from a zero-cost investment strategy. While daily data are employed in this study, bid–ask bounce is not an issue here since the exchange rate data give the last observation on the *midpoint* of the bid and ask prices; further, due to the heavy trading

¹⁸Our results appear to be quite robust to the starting point of the forecast period: results for each of the four models are very similar if we start the forecast period at $\frac{1}{4}$ or $\frac{1}{2}$ of the sample instead of at $\frac{1}{3}$.

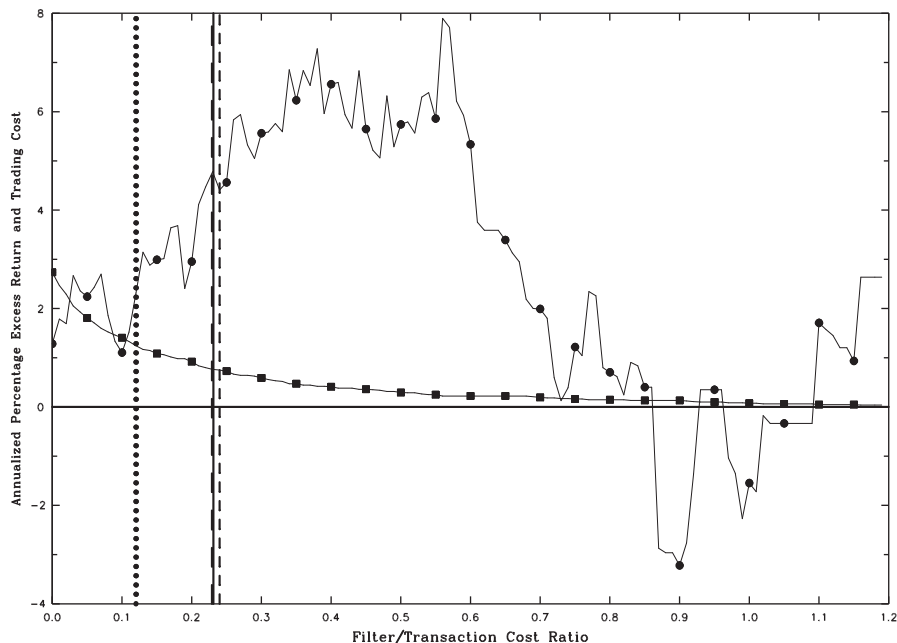


Fig. 3. The effect of filter on after-cost returns for the MA(21) model. *Note:* The solid line with round symbols displays the after-cost excess return and the one with square symbols displays the trading cost. The vertical lines represent different filters: the solid line is the theoretical optimal filter under uniform distribution $-\mu_U^*/c$; the short dashed line is the optimal filter under normal distribution $-\mu_N^*/c$; the long dashed line is the optimal filter under bootstrap distribution $-\mu_B^*/c$; and the dotted line is the ad hoc filter that maximizes the in-sample smoothed after-cost excess return using the first third of the sample, $-\mu_E^*/c$.

volume of the yen, non-synchronous trading issue is not relevant. Additionally, while the daily interest observations do not coincide exactly with the exchange rate observations, the high volatility of the exchange rate relative to the interest rates implies that any bias due to a timing mismatch is probably negligible.

From the second forecasted day (November 29) until the end of the sample, our strategy works as follows. For each day, we use all available observations to estimate the model parameters and forecast the excess return for the following day. If either of the following two conditions occurs, a transaction will take place. (1) If the forecasted excess return is positive, its magnitude is larger than the transaction cost filter, and we currently have a long position in the dollar (and a short position in the yen), then we reverse our position by taking a long position in the yen and a short position in the dollar for the following day. This counts as one trade involving two one-way transaction costs. (2) If the forecasted excess return is negative, its magnitude is larger than the transaction cost filter, and the current holdings are long in the yen and short in the dollar, then we reverse our position by taking a long position in the dollar and a short position in the yen. This counts as one transaction and again involves two one-way transaction costs. If neither of the above two conditions applies, no trade takes place. The current holdings (both long and short) carry over to the following day and no transaction costs are incurred.

We compute the average excess return for the zero-cost investment strategy and the associated t -ratio for the out-of-sample forecasting period. We document the before-cost

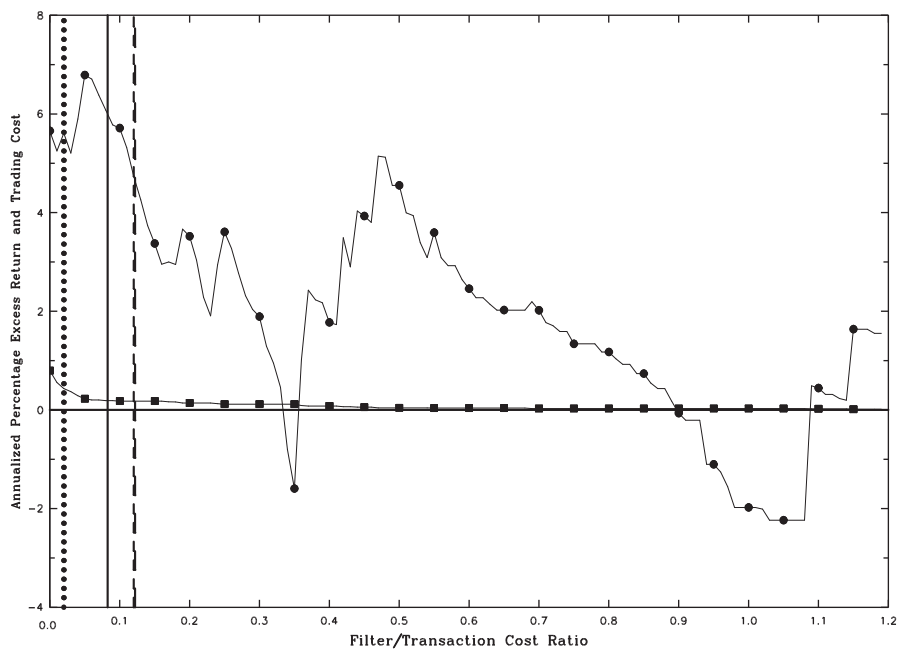


Fig. 4. The effect of the filter on after-cost returns for the MA(126) model. *Note:* The solid line with round symbols displays the after-cost excess return and the one with square symbols displays the trading cost. The vertical lines represent different filters: the solid line is the theoretical optimal filter under uniform distribution $-\mu_U^*/c$; the short dashed line is the optimal filter under normal distribution $-\mu_N^*/c$; the long dashed line is the optimal filter under bootstrap distribution $-\mu_B^*/c$; and the dotted line is the ad hoc filter that maximizes the in-sample smoothed after-cost excess return using the first third of the sample $-\mu_E^*/c$.

and after-cost excess return rates for the case *without* filter, and the after-cost excess return rates for the cases *with* transaction cost filter. For perspective, the simple buy-and-hold strategy of holding the yen and shorting the dollar over the whole out-of-sample period yields an annualized return of -0.00953 (the reverse strategy of holding the dollar and shorting the yen yields therefore $+0.00953$), $<1\%$. This return is not statistically distinguishable from zero (t -ratio = 0.348).

4.2. Basic results

Table 2 reports the results for the four forecasting models. Each model is discussed separately below. For the ARMA(1,1) model, the before-cost excess return in the case without a filter is 5.7% per annum, which is statistically significant at the 5% level. The strategy, however, requires 444 trades over 4,130 trading days and accounting for the round-trip costs of 10 basis points reduces the after-cost excess return to 3.0%, which is no longer statistically significant. The naïve filter equal to the 5 basis points one-way transaction cost “ c ” dramatically reduces the number of trades to 12, resulting in a lower excess return of 0.6%, which is statistically insignificant. In contrast, the optimal filter, μ_U^* , captures many of the profitable trades and yields an excess return of 5.5%, which is significant at the 5% level. The filter under the bootstrapped distribution μ_B^* produces

Table 2

Effects of transaction costs on trading performance in foreign exchange.

This table reports trading performance in the Japanese yen with a round-trip transaction cost $2c = 0.1\%$. The data cover the period from August 31, 1978 to May 3, 2003 with 6,195 daily observations. The first third of the sample (2,065 observations) is used for model estimation. The parameter estimates are used to calculate the optimal transaction cost filters. Out-of-sample forecasting starts on November 28, 1986 until the end of the sample (4,130 observations). Columns 1–5 display results where no transaction cost filter is imposed. Columns 2–3 report the *before-cost* excess returns (from the zero-cost investment strategies) and *t*-ratios, whereas Columns 4–5 report the *after-cost* excess returns and *t*-ratios. Similarly, Columns 6–8 report the results (after-cost) when a naïve filter equal to the actual transaction cost c is imposed. Columns 9–11, 12–14, and 15–17, show the results when the optimal filters, μ_U^* , μ_N^* and μ_B^* , are imposed, respectively. All returns are annualized.

Model	Without transaction cost filter					With naïve filter “c”			With optimal filter μ_U^*			With optimal filter μ_N^*			With optimal filter μ_B^*		
	Before cost		After cost			After cost			After cost			After cost			After cost		
	# of trades	Excess return	<i>t</i> -Ratio	Excess return	<i>t</i> -ratio	# of trades	Excess return	<i>t</i> -Ratio	# of trades	Excess return	<i>t</i> -Ratio	# of trades	Excess return	<i>t</i> -Ratio	# of trades	Excess return	<i>t</i> -Ratio
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	
ARMA(1,1)	444	0.057	2.054	0.030	1.073	12	0.006	0.225	81	0.055	2.007	71	0.070	2.549	81	0.057	2.050
AR(1)	2100	0.069	2.486	−0.059	−2.144	34	0.011	0.390	42	0.058	2.105	48	0.056	2.017	66	0.026	0.940
MA(21)	448	0.040	1.453	0.013	0.463	12	−0.015	−0.560	125	0.044	1.591	123	0.044	1.596	125	0.048	1.728
MA(126)	131	0.065	2.338	0.057	2.047	5	−0.020	−0.714	31	0.059	2.143	29	0.047	1.715	29	0.047	1.715

nearly the same results as μ_U^* , whereas the filter under the normality assumption, μ_N^* , yields an even higher return of 7.0%, which is significant at the 1% level.

For the AR(1) model, the strategy without filter involves 2,100 switches over 4,130 trading days (over 50% of the time). In the absence of transaction costs, the strategy produces an annualized excess return of 6.9% with a t -ratio of 2.486, which is statistically significant at the 5% level. However, the transaction cost completely wipes out the profits, resulting in a negative excess return of 5.9%. A naïve filter equal to the actual round-trip transaction cost of 10 basis points dramatically reduces the number of transactions to 34, and yields an insignificant excess return of 1.1% per annum. While it is somewhat useful, this naïve filter may be too conservative because it does not exploit the information on the persistence of expected return in the exchange rate, thereby missing a number of profitable trades. The strategy with the optimal filter μ_U^* under the assumption of a uniform distribution captures just that opportunity. It produces 42 trades and yields a higher excess return of 5.8%, which is significant at the 5% level. Similarly, the optimal filter under the normality assumption μ_N^* produces an average excess return of 5.6% per annum, also significant at the 5% level. The bootstrapped filter μ_B^* yields an insignificant excess return of 2.6%.

For the MA(21) model, our strategy with the optimal filters again generates higher excess returns than the alternatives, although none of the excess returns are statistically significant at the 5% level.

Finally, the long MA(126) process provides a very smooth forecast of expected returns. While the strategy without filter yields an after-cost return of 5.7%, which is significant at the 5% level, the naïve filter equal to c skips too many profitable trades, resulting in a negative return of 2%. The optimal filter μ_U^* , while very small relative to c , is capable of filtering many days with low expected returns and capturing those days when expected returns are substantial. This filter produces an expected return of 5.9%, which is significant at the 5% level. The other two filters, μ_N^* and μ_B^* , yield somewhat smaller returns (4.7%), which are statistically significant only at the 10% level.¹⁹

4.3. Filter size and after-cost excess returns

Figs. 1–4 display the trading strategy returns (after cost) and trading costs for the four return processes as a function of the filter value. As expected, the trading cost declines monotonically as the filter value rises. The after-cost excess return lines illustrate that in all cases the ex ante optimal filters are reasonably close to the ex post optimum (with the empirical filter in the AR(1) case being the only exception). Since the actual data are just one random draw from the unobserved true process, this is all one should expect of a good model. Except for the AR(1) case, the trading strategy returns display the hump-shaped pattern expected for the after-cost returns.

A striking feature of these four figures is that, even though the optimal filters differ radically across the four cases, the empirical (ex post) maximum filter value is quite close to the (ex ante) optimal filter in all four cases. While each case approximates the true data process to a certain extent, it is not surprising that the ARMA(1,1) process provides the

¹⁹The trading strategies for each of the forecasting models imply a reasonably even choice of each currency. For example, with the optimal filter μ_U^* , the fraction of long yen and short U.S. dollar is: ARMA(1,1) 1830/4130, AR(1) 2555/4130, MA(21) 1987/4130, and MA(216) 1875/4130.

best overall fit as it is well-known to be a parsimonious description of general ARMA(p, q) processes. The strong performance of the ARMA(1,1) process and the poorer performance of the AR(1) process is consistent with the results of LeBaron (1992) and Taylor (1992) that ARMA(1,1) processes are far better at capturing the key features of exchange rate series.^{20,21}

4.4. Correction for potential data snooping

The previous subsections report the trading strategy results by examining multiple model specifications. Overall, we explore four data-generating processes (ARMA(1,1), AR(1), MA(21), and MA(126)) with three distributions (uniform, normal, and bootstrap), yielding a total of 12 different specifications. It would be undue and incorrect for a researcher to report only the best-performing specification or draw inferences based solely on the best-performing specification, as doing so leads to a so-called data-snooping bias.

Traditionally, bounds on the p -value for tests of the null when multiple specifications are investigated can be constructed using the Bonferroni inequality [see, for example, Lehmann and Modest, 1987 for an application]. White (2000) develops a novel procedure, called Reality Check, to measure and correct for data-snooping biases. The idea is to generate the empirical distribution from the full set of specifications that generates the best-performing outcome and to draw inferences from this distribution. We apply this methodology to check for the robustness of our results.

White's (2000) test evaluates the distribution of a performance measure, accounting for the full set of specifications that lead to the best-performing specification. In our case, the performance measure is the after-cost excess return. The null hypothesis to be tested is that the best specification is no better than the benchmark (which is zero excess return), i.e.,

$$H_0 : \max_{k=1, \dots, L} \{E(R_k)\} \leq 0, \quad (19)$$

where k denotes a specification, L is the total number of specifications, the expectation is evaluated with the average

$$\bar{R}_k = \frac{1}{T} \sum_{t=1}^T R_{k,t}.$$

$R_{k,t}$ is the after-cost excess return of specification k at time t , and T is the number of forecasting periods. Rejection of the null will lead to the conclusion that the best specification achieves performance superior to the benchmark.

Following White (2000), we test the null hypothesis H_0 by applying the stationary bootstrap method of Politis and Romano (1994) to the observed values of $R_{k,t}$ as follows.

²⁰Figures available from the authors provide a breakdown of the effect of the optimal filters on trading frequency. In each of the models, the optimal filters, the μ_U^* , reduce trading frequency considerably but the trades remain quite evenly distributed over time. For instance, for the ARMA(1,1) model, a minimum of two trades and a maximum of eight trades occurs in each (full) year under the optimal filter trading strategy.

²¹A table available from the authors provides risk-adjusted trading rule returns. We correct the ex post trading rule returns from each of the four forecast models for CAPM market risk using the MSCI World market index and the euro interest rate as the risk-free rate (results using the U.S. S&P 500 value-weighted market index are similar). In all cases the market risk sensitivities of the zero-cost investment positions are near zero. Thus, the risk-adjusted returns, the "alphas," are very close to the unadjusted returns.

Step 1: We resample the realized excess return series $R_{k,t}$, one block of observations at a time with replacement, and denote the resulting series by $R_{k,t}^*$. This process is repeated N times.

Step 2: For each replication i , we compute the sample average of the bootstrapped returns, denoted by

$$\bar{R}_{k,i}^* \equiv \frac{1}{T} \sum_{t=1}^T R_{k,t}^*, \quad i = 1, \dots, N.$$

Step 3: We construct the following statistics:

$$\bar{V} = \max_{k=1, \dots, L} \{\sqrt{T}(\bar{R}_k)\}, \quad (20)$$

$$\bar{V}_i^* = \max_{k=1, \dots, L} \{\sqrt{T}(\bar{R}_{k,i}^* - \bar{R}_k)\}, \quad i = 1, \dots, N. \quad (21)$$

Step 4: White's Reality Check p -value is obtained by comparing \bar{V} to the percentiles of \bar{V}_i^* .

Our out-of-sample forecasting period contains 4,130 observations, and we experiment with a total of 12 specifications; therefore, $T = 4,130$ and $L = 12$. We choose $N = 10,000$. With the smoothing parameter set equal to 0.5,²² we find a White p -value of 0.034. Our results are robust to the value of the smoothing parameter. We obtain White p -values of 0.041, 0.030, and 0.032, for smoothing parameter values of 0.1, 0.9, and 1, respectively. These results imply that the null hypothesis that our trading strategy does not produce an after-cost excess return can be rejected at the 5% significance level even after properly accounting for the potential data-snooping bias. The robustness of these results should, however, be interpreted with caution because our Reality Check simulation for data-snooping bias correction is based on the 12 trading rules experimented with in this paper but does not account for all the rules studied in the literature.

5. Conclusion

If transitory profitable trading opportunities exist, transaction filters are used in practice to mitigate trading costs; but the filter size is difficult to determine a priori. This paper uses a dynamic programming framework to design a filter that is optimal in the sense of maximizing expected returns after transaction costs. The optimal filter size depends negatively on the degree of persistence of the profitable trading opportunities, positively on transaction costs, and positively on the standard deviation of shocks.

We apply our theoretical results to foreign exchange trading by parameterizing the moving average strategy often employed in foreign exchange markets. The parameterization implies the same decisions as the moving average rule in the absence of transaction costs, but has the advantage of translating the buy/sell signal into the same units as the transaction costs so that the optimal filter can be calculated.

²²The smoothing parameter, which is greater than 0 and no greater than 1, controls the block length in the bootstrap re-sampling. A larger value is appropriate for data with little dependence and a smaller value is appropriate for data with more dependence [see Politis and Romano, 1994; White, 2000].

Application to daily dollar–yen trading demonstrates that the optimal filter is not solely of academic interest but may differ to an economically significant extent from a naïve filter equal to the transaction cost. This depends importantly on the time series process that we assume for the exchange rate dynamics. In particular, we find that for an AR(1) process, the optimal filter is close to the naïve transaction cost filter, but for the ARMA(1,1) process, the optimal filter is only around 30% of the naïve transaction cost filter, and for the more stable MA processes, the optimal filter is smaller still as a fraction of the transaction cost. Impressively, the ex ante optimal filters under the assumptions of uniform, normal, and bootstrap distributions are all very close to one another and all are quite close to the ex post after-cost return-maximizing level.

We confirm that simple daily moving average foreign exchange trading generates positive returns that disappear after accounting for transaction costs. However, when the optimal filter is used, returns after transaction costs remain positive and are higher than for naïve filters. This result has implications beyond foreign exchange markets. It cautions against dismissing abnormal returns as due to transactions costs, merely because the after-cost return is negative or insignificant. For instance, Lesmond et al. (2004) argue convincingly that momentum profits disappear when actual transaction costs are properly considered, even after accounting for the proportion of securities held over in each period. But their after-cost returns are akin to those for our suboptimal zero filter strategy. It would be interesting to see what outcome would arise if an optimal filter were used.

In our sample, the trading strategy returns, gross of transaction costs, are significantly positive, but no longer significant after transaction costs are subtracted. But if we optimally eliminate trades that do not make up for their transaction cost, then the after-cost profits are only slightly lower than the gross profits from unrestricted trading and are statistically significant. They are also economically significant, around 0.5% per month after transaction costs, which raises the issue of market efficiency. The profits are of similar magnitude as the momentum profits after transaction costs and may in fact be closely related to the momentum phenomenon. However, given the lower total variance of the trading strategy returns,²³ it is even more difficult here, compared to the momentum case for equity returns, to argue that an unobserved systematic risk is responsible. So the “anomaly” may be exploitable and, in the absence of a risk explanation, could suggest market inefficiency.

Apart from the practical advantages of using the optimal filter, there is also a methodological advantage: in studies attempting to calculate abnormal returns from particular trading strategies in which transaction costs are important, there is no guideline as to what filter to use in dealing with transaction costs. Lesmond et al. (2004, p. 370) note: “Although we observe that trading costs are of similar magnitude to the relative strength returns for the specific strategies we consider, there is an infinite number of momentum-oriented strategies to evaluate, so we cannot reject the existence of trading profits for all strategies”. Rather than allowing the data-mining problem that is likely to arise when a variety of filter sizes are applied, our approach here provides a unique filter in Eq. (13) that can be unambiguously obtained in advance from observable variables.

²³For example, from Jegadeesh and Titman (1993), the standard deviation of momentum excess return for the 6-month sorting and 6-month holding strategy is 5.4% per month. Based on our theoretical optimal filter, μ_U^* , the standard deviation of after-cost return for the ARMA(1,1) specification is 0.7% per day, or 3.2% per month, which is lower than that from the momentum strategy.

Acknowledgments

We would like to thank Gershon Alperovich, Ake Blomqvist, Ivan Brick, Justin Chan, Sris Chatterjee, N.K. Chidambaran, Chong Tze Chua, Charles Engel, Aditya Goenka, Sam Henkel, Dong Hong, Lixin Huang, Qiang Kang, James Lothian, Paul McNelis, Salih Neftci, Darius Palia, Andrew Rose, Menahem Spiegel, John Wald, Yonggan Zhao, the editor (Bruce Lehmann), the anonymous referee, and seminar participants at the Financial Management Association International, the Pacific Basin Finance, Economics and Accounting Conference, the China International Conference in Finance, Fordham University, Nanyang Technological University, the National University of Singapore, Peking University, Rutgers University, Singapore Management University and the University of Hong Kong for helpful conversations and comments. Yangru Wu would like to thank the Whitcomb Center at the Rutgers Business School and the Office of Research of Singapore Management University for financial support. Part of this work was completed while Yangru Wu visited Singapore Management University and the Central University of Finance and Economics. He would like to thank these institutions for their hospitality. We are solely responsible for any remaining errors.

Appendix A

A.1. Proof of Proposition 1

In Eq. (1), consider that $V_2(x_{t-1}, \varepsilon_{t-1}) = V_1(-x_{t-1}, -\varepsilon_{t-1})$ by virtue of the fact that, for the symmetric distribution of innovations, the expected returns of one position are equivalent to the expected returns of the reverse position if the return variables in Eq. (2) are the negative. Define $\mu_t \equiv E_{t-1}x_t$. Then, given Eq. (2), we have $\mu_{t+1} \equiv \rho x_t - \delta \varepsilon_t$ and thus we can without loss of generality redefine the state variables so that $V_1(x_t, \varepsilon_t) \equiv V(\mu_{t+1}, \varepsilon_t)$. Then Eq. (1) becomes:

$$V(\mu_t, \varepsilon_{t-1}) = \mu_t + E_{t-1} \left[\text{Max} \left(\frac{V(\mu_{t+1}, \varepsilon_t)}{1+r}, \frac{V(-\mu_{t+1}, -\varepsilon_t)}{1+r} - 2c \right) \right]. \tag{A.1}$$

To derive Eq. (4), consider Eq. (2), $x_t - \rho x_{t-1} = \varepsilon_t - \delta \varepsilon_{t-1}$. Taking conditional expectations on both sides yields $x_t - \rho x_{t-1} - \delta \varepsilon_{t-1}$. Now lag the first equation by one period and take conditional expectations at time $t-2$. Simple algebra yields $x_{t-1} = \mu_{t-1} + \varepsilon_{t-1}$. Combining the last two equations produces Eq. (4), which says that the conditional expected return of the ARMA(1,1) model is an AR(1) process. It is further clear, by inspection of the decision problem now summarized by Eqs. (A.1) and (4), that, once the state variable μ_t is considered, there is no additional role for the state variable ε_{t-1} . Thus, we obtain $V(\mu_t, \varepsilon_{t-1}) \equiv V(\mu_t)$ and (A.1) becomes:

$$V(\mu_t) = \mu_t + E_{t-1} \left[\text{Max} \left(\frac{V(\mu_{t+1})}{1+r}, \frac{V(-\mu_{t+1})}{1+r} - 2c \right) \right]. \tag{A.2}$$

By the symmetry of the density function, we also have

$$V(-\mu_t) = -\mu_t + E_{t-1} \left[\text{Max} \left(\frac{V(-\mu_{t+1})}{1+r}, \frac{V(\mu_{t+1})}{1+r} - 2c \right) \right]. \tag{A.2'}$$

From (A.2) and (A.2'), we have

$$\begin{aligned}
 &V(-\mu_t) - V(\mu_t) \\
 &= -2\mu_t + \int_{\{\varepsilon_t: -2c \leq (V(-\mu_{t+1}) - V(\mu_{t+1})) / (1+r) \leq 2c\}} \frac{V(-\mu_{t+1}) - V(\mu_{t+1})}{1+r} dF(\varepsilon_t) \\
 &\quad + 2c \left[\int_{\{\varepsilon_t: (V(-\mu_{t+1}) - V(\mu_{t+1})) / (1+r) > 2c\}} dF(\varepsilon_t) - \int_{\{\varepsilon_t: -2c > (V(-\mu_{t+1}) - V(\mu_{t+1})) / (1+r)\}} dF(\varepsilon_t) \right],
 \end{aligned} \tag{A.3}$$

which evaluates the Max contingent on the possible realizations of the innovation ε_t . It can be shown by induction that $V(-\mu_t) - V(\mu_t)$ is strictly decreasing in μ_t (it is true for some terminal time T ; then, if it holds for some time $t+1$, and given that $\rho > 0$ in Eq. (4), the second term on the right-hand side of (A.3) is also decreasing in μ_t and we can establish the result for time t . Then let T go to infinity). Further, a unit decrease in μ_t increases $V(-\mu_t) - V(\mu_t)$ by at least 2 (the direct gain from switching). Hence, since the innovation ε_t has unbounded support, then the distribution of μ_{t+1} is unbounded and there exists exactly one finite value μ^* at which $(V(-\mu^*) - V(\mu^*)) / (1+r) = 2c$. Accordingly, the investor maintains the current zero-investment position whenever $\mu_{t+1} > \mu^*$ and reverses the current position whenever $\mu_{t+1} \leq \mu^*$. It is then straightforward to convert (A.2) to Eq. (3).

A.2. Derivation of Eq. (5)

From Eq. (3), $V(-\mu_t)$ is given as

$$\begin{aligned}
 V(-\mu_t) &= -\mu_t + \int_{(\mu^* + \rho\mu_t) / (\rho - \delta)}^{\infty} \left(\frac{V[-\rho\mu_t + (\rho - \delta)\varepsilon_t]}{1+r} \right) dF(\varepsilon_t) \\
 &\quad + \int_{-\infty}^{(\mu^* + \rho\mu_t) / (\rho - \delta)} \left(\frac{V[\rho\mu_t - (\rho - \delta)\varepsilon_t]}{1+r} - 2c \right) dF(\varepsilon_t).
 \end{aligned} \tag{A.4}$$

Using the symmetry property of the density function and a change in variable from ε_t to $-\varepsilon_t$, we can write Eq. (A.4) as follows:

$$V(-\mu_t) = -\mu_t + \int_{-\infty}^{(-\mu^* - \rho\mu_t) / (\rho - \delta)} \left(\frac{V(-\mu_{t+1})}{1+r} \right) dF(\varepsilon_t) + \int_{(-\mu^* - \rho\mu_t) / (\rho - \delta)}^{\infty} \left(\frac{V(\mu_{t+1})}{1+r} - 2c \right) dF(\varepsilon_t). \tag{A.5}$$

Then, from Eqs. (3) and (A.5), we obtain

$$V(-\mu_t) - V(\mu_t) = -2\mu_t + B(\mu_t) + C(\mu_t),$$

with

$$B(\mu_t) = \int_{(\mu^* - \rho\mu_t) / (\rho - \delta)}^{(-\mu^* - \rho\mu_t) / (\rho - \delta)} \left(\frac{V(-\mu_{t+1}) - V(\mu_{t+1})}{1+r} \right) dF(\varepsilon_t), \tag{A.6}$$

$$C(\mu_t) = -2c \left(\int_{(\mu^* - \rho\mu_t) / (\rho - \delta)}^{\infty} dF(\varepsilon_t) - \int_{-\infty}^{(\mu^* - \rho\mu_t) / (\rho - \delta)} dF(\varepsilon_t) \right) = 2c \int_{(\mu^* + \rho\mu_t) / (\rho - \delta)}^{(\mu^* - \rho\mu_t) / (\rho - \delta)} dF(\varepsilon_t), \tag{A7}$$

where the last equality follows from the symmetry of the density function. Thus,

$$\begin{aligned}
 V(-\mu_t) - V(\mu_t) = & -2\mu_t + \int_{(\mu^* - \rho\mu_t)/(\rho - \delta)}^{(-\mu^* - \rho\mu_t)/(\rho - \delta)} \left(\frac{V(-\mu_{t+1}) - V(\mu_{t+1})}{1 + r} \right) dF(\varepsilon_t) \\
 & + 2c \int_{(\mu^* + \rho\mu_t)/(\rho - \delta)}^{(\mu^* - \rho\mu_t)/(\rho - \delta)} dF(\varepsilon_t), \tag{A.8}
 \end{aligned}$$

which is Eq. (5) given the definitions in Eqs. (6) and (7).

References

- Alexander, S.S., 1961. Price movements in speculative markets: trends or random walks. *Industrial Management Review* 2, 7–26.
- Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics* 52, 47–78.
- Balvers, R.J., Mitchell, D.W., 1997. Autocorrelated returns and optimal intertemporal portfolio choice. *Management Science* 43, 1537–1551.
- Balvers, R.J., Wu, Y., Gilliland, E., 2000. Mean reversion across national stock markets and parametric contrarian investment strategies. *Journal of Finance* 55, 745–772.
- Bessembinder, H., 1994. Bid-ask spreads in the interbank foreign exchange markets. *Journal of Financial Economics* 35, 317–348.
- Canova, F., 1991. An empirical analysis of ex ante profits from forward speculation in foreign exchange markets. *Review of Economics and Statistics* 73, 489–496.
- Cheung, Y.W., Wong, C., 2000. A survey of market practitioners' view on exchange rate dynamics. *Journal of International Economics* 51, 401–419.
- Cooper, M., 1999. Filter rules based on price and volume in individual security overreaction. *Review of Financial Studies* 12, 901–935.
- Cornell, W.B., Dietrich, J.K., 1978. The efficiency of the market for foreign exchange under floating exchange rates. *Review of Economics and Statistics* 60, 111–120.
- Engel, C., 1996. The forward discount anomaly and the risk premium: a survey of recent evidence. *Journal of Empirical Finance* 3, 123–191.
- Fama, E.F., Blume, M., 1966. Filter rules and stock market trading profits. *Journal of Business* 39, 226–241.
- Frankel, J.A., Froot, K.A., 1990. Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review* 80, 181–185.
- Gencay, R., 1999. Linear, nonlinear and essential foreign exchange prediction with simple technical rules. *Journal of International Economics* 47, 91–107.
- Gencay, R., Ballocci, G., Dacorogna, M., Olsen, R., Pictet, O., 2002. Real-time trading models and the statistical properties of foreign exchange rates. *International Economic Review* 43, 463–491.
- Grundy, B., Martin, J.S., 2001. Understanding the nature and the risks and the sources of the rewards to momentum investing. *Review of Financial Studies* 14, 29–78.
- Hanna, J.D., Ready, M.J., 2005. Profitable predictability in the cross-section of stock returns. *Journal of Financial Economics* 78, 463–505.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48, 65–91.
- Knez, P.J., Ready, M.J., 1996. Estimating the profits from trading strategies. *Review of Financial Studies* 9, 1121–1163.
- LeBaron, B., 1992. Do moving average trading rule results imply nonlinearities in foreign exchange markets? Technical Report, University of Wisconsin–Madison.
- LeBaron, B., 1998. Technical trading rules and regime shifts in foreign exchange. In: Acar, F., Satchell, S. (Eds.), *Advanced Trading Rules*. Butterworth-Heinemann, Woburn, MA, pp. 5–40.
- LeBaron, B., 1999. Technical trading rule profitability and foreign exchange intervention. *Journal of International Economics* 49, 125–143.
- Lee, C.I., Mathur, I., 1996. Trading rule profits in European currency spot cross rates. *Journal of Banking and Finance* 20, 949–962.

- Lehmann, B.N., 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1–28.
- Lehmann, B., Modest, D.M., 1987. Mutual fund performance evaluation: a comparison of benchmarks and benchmark comparisons. *Journal of Finance* 42, 233–265.
- Lesmond, D.A., Schill, M.J., Zhou, C., 2004. The illusory nature of momentum profits. *Journal of Financial Economics* 71, 349–380.
- Levich, R.M., 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.
- Liu, H., 2004. Optimal consumption and investment with transaction costs and multiple risky assets. *Journal of Finance* 59, 289–338.
- Liu, H., Loewenstein, M., 2002. Optimal portfolio selection with transaction costs and finite horizons. *Review of Financial Studies* 15, 805–835.
- Lynch, A.W., Balduzzi, P., 2000. Predictability and transaction costs: the impact on rebalancing rules and behavior. *Journal of Finance* 55, 2285–2310.
- Lynch, A.W., Tan, S., 2009. Multiple risky assets, transaction costs and return predictability: allocation rules and implications for US investors. *Journal of Financial and Quantitative Analysis*, forthcoming.
- Mark, N.C., Wu, Y., 1998. Rethinking deviations from uncovered interest parity: the role of covariance risk and noise. *The Economic Journal* 108, 1686–1706.
- Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis* 39, 407–429.
- Melvin, M., Tan, K.H., 1996. Foreign exchange market bid–ask spreads and the market price of social unrest. *Oxford Economic Papers* 48, 329–341.
- Neely, C., 2002. The temporal pattern of trading rule returns and Central Bank intervention: intervention does not generate technical trading rule profits. *Journal of International Economics* 58, 211–232.
- Neely, C., 2003. Risk-adjusted, ex ante, optimal technical trading rules in equity markets. *International Review of Economics and Finance* 12, 69–87.
- Neely, C., Weller, P., 2001. Technical analysis and Central Bank intervention. *Journal of International Money and Finance* 20, 949–970.
- Neely, C., Weller, P., 2003. Intraday technical trading in the foreign exchange market. *Journal of International Money and Finance* 22, 223–237.
- Politis, D., Romano, J., 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.
- Qi, M., Wu, Y., 2006. Technical trading-rule profitability, data snooping, and reality check: evidence from the foreign exchange market. *Journal of Money, Credit and Banking* 30, 2135–2158.
- Sweeney, R.J., 1986. Beating the foreign exchange market. *Journal of Finance* 41, 163–182.
- Sweeney, R.J., 2000. Does the fed beat the foreign-exchange market? *Journal of Banking and Finance* 24, 665–694.
- Taylor, D., 1982. Official intervention in the foreign exchange market, or, bet against the central bank. *Journal of Political Economy* 90, 356–368.
- Taylor, S.J., 1992. Rewards available to currency futures speculators: compensation for risk or evidence of inefficient pricing? *Economic Record* 68 (Suppl.), 105–116.
- White, H., 2000. A reality check for data snooping. *Econometrica* 68, 1097–1126.