

**OPTIMAL TRANSACTION FILTERS UNDER TRANSITORY TRADING
OPPORTUNITIES:
Theory and Empirical Illustration***

RONALD J. BALVERS
Division of Economics and Finance
College of Business
West Virginia University
Morgantown, WV 26506

rbalvers@wvu.edu
Phone: (304) 293-7880
Fax: (304) 293-2233

YANGRU WU
Department of Finance & Economics
Rutgers Business School-Newark & New Brunswick
Rutgers University
Newark, NJ 07102-1820

yangruwu@andromeda.rutgers.edu
Phone: (973) 353-1146
Fax: (973) 353-1233

December 6, 2004

ABSTRACT

If transitory profitable trading opportunities exist, filter rules are used to mitigate transaction costs. We use a dynamic programming framework to design an optimal filter which 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. Applying our theory to daily dollar-yen exchange trading, we find that 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 accounting for transaction costs. However, when the optimal filter is used, returns after costs remain positive and are higher than for naïve filters.

Keywords: Transaction Costs, Filter Rules, Trading Strategies, Foreign Exchange
JEL Codes: G10, G15, G11.

* We would like to thank Gershon Alperovich, Ake Blomqvist, Ivan Brick, Justin Chan, Sris Chatterjee, N. K. Chidambaran, Chong Tze Chua, Aditya Goenka, Dong Hong, Lixin Huang, Qiang Kang, James Lothian, Paul McNelis, Salih Nefci, Darius Palia, Andrew Rose, Menahem Spiegel, John Wald and seminar participants at the FMA, the China International Conference in Finance, Fordham 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 Research Council of Rutgers University and Rutgers Business School for financial support. Part of this work was completed while Yangru Wu visited the Hong Kong Institute for Monetary Research. He thanks the Hong Kong Institute for Monetary Research for its hospitality. We are solely responsible for any remaining errors.

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 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 arbitraging of microstructure inefficiencies that require frequent and timely transactions, which may be largely riskless. An example is uncovered interest speculation in currency market where a trader takes either one side of the market or the reverse. Alternatively, a trader arbitrages differences between a stock's return and that of one of its derivatives: going long on the arbitrage position or reversing the position and going short. The model also applies to momentum or contrarian trading when these phenomena may reflect inefficiencies due to cognitive biases.

Starting with Alexander (1961) and Fama and Blume (1966), trading rules often involve a "filter" that prompts one to trade only if a realization exceeds a benchmark by a certain percentage. For instance in foreign exchange trading an x% filter might stipulate that a currency be purchased if its current exchange value exceeds some moving average of its past values by more than the x% "band". Presumably, such a filter reduces trading and thus transaction costs. The idea is that if the trade indicator is "weak" the expected return from the transaction may not compensate for the transaction cost. Knez and Ready (1996) and Cooper (1999), for instance, explore different filters and find that the after-transaction-cost returns 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 would be reasonable, since the buy/sell signal and the transaction cost are not in the same units. This also implies that there is no discipline against data mining for researchers: a large number of filters with different bands can be tried until the strategy generates

¹For instance, Grundy and Martin (2001) express doubt that the anomalous momentum profits survive transaction costs, and Hanna and Ready (2001) find that the momentum profits are substantially reduced when transactions costs are accounted for. Lesmond, Schill, and Zhou (2004) conclude more strongly that momentum profits with transactions costs are illusory.

positive net returns.

The purpose of this paper is to design an optimal filter that maximizes the expected return net of transaction cost. To accomplish this we employ a “parametric” approach (see for instance Balvers, Wu, and Gilliland, 2000) that allows the trading signal and the transaction cost 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 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 under the assumption of changes in expected returns being uniformly distributed.

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 (2002).² 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

² There 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 transactions costs. Their approach, however, is a complement to ours in that they integrate risk into the optimal switching choice while accounting for transactions costs after the fact, whereas our approach integrates transactions costs into the optimal switching choice while accounting for risk after the fact.

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, Lynch and Balduzzi, and Lynch and Tan, 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 framework implies that trading strategy returns using the optimal filter should be smaller than the returns when transaction costs are ignored but larger than for any other filter strategy with transaction costs, in particular those for which the filter is zero percent or equal to the transaction cost. We apply our optimal filter to a natural case for our model: daily foreign exchange trading in the yen/dollar market. As is well-known [see for instance Cornell and Dietrich (1978), Sweeney (1986), LeBaron (1998), and Qi and Wu (2004)], simple moving-average trading rules generate positive expected returns (with or without risk adjustment) in the foreign exchange market. However, for daily trading, returns net of transaction costs are negative or insignificant if no filter is applied.

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. Furthermore, 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 moreover 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 transactions cost, without leading to excessive data mining. The results also hint that in some cases conclusions of abnormal profits disappearing after accounting for

transaction costs may be worth revisiting.

The next section 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 II, 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 AR(1), an ARMA(1,1), 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. Section III then conducts the out-of-sample test with the final two-thirds of our 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 AR(1), ARMA(1,1), and MA returns cases. Section IV concludes the paper.

I. The Theoretical Model

A. Autoregressive Conditional Returns and Two Risky Assets

An individual investor attempts to maximize 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. The two assets are treated symmetrically in that shorting asset 1 is equivalent to holding asset 2 and vice versa. 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 asset. This implies that any profits or losses do not affect future investment positions.

The problem can be solved by using an autonomous stochastic dynamic programming approach with state variables indicating: (1) the asset currently held (asset 1 or 2), and (2) the variables summarizing the state of the forecast variable, x , that characterizes the persistent part in the returns of assets 1 and 2; x is assumed to follow an ARMA(1,1) process as a parsimonious parameterization of mild return predictability. The solution is characterized in general and is given in closed form when the shock to expected return is

uniformly distributed.

The Bellman equation for this decision problem is as follows:

$$V_i(x_{t-1}, \varepsilon_{t-1}) = E_{t-1} r_i(x_t, \eta_{it}) + E_{t-1} \text{Max} \left[\frac{1}{1+r} V_i(x_t, \varepsilon_t), \frac{1}{1+r} V_j(x_t, \varepsilon_t) - c \right],$$

$$\text{for } i, j \in \{1, 2\}, j \neq i \quad (1)$$

$$\text{subject to: } x_t = \rho x_{t-1} - \delta \varepsilon_{t-1} + \varepsilon_t, \quad E_{t-1} \varepsilon_t = 0, \rho \neq \delta, \quad (2)$$

where r is the discount rate. The disturbance term ε_t is assumed to have a symmetric and unimodal probability density function denoted by $f(\varepsilon_t)$. Each period, the investor chooses whether to invest a fixed amount (one unit) in asset 1 or asset 2 (i equals 1 or 2). A proportional transaction cost c is incurred whenever there is a switch from one asset to the other.³ The return of asset i is given by $r_i(x_t, \eta_{it})$, with a distribution that is time-varying based on the autoregressive state variable x_t described by equation (2); η_{it} is white noise. Thus, the value function at time $t-1$ depends on the state as given by the holdings of asset i and the variables describing the distribution of x_t , namely x_{t-1} and ε_{t-1} . It equals the expected return depending on current asset holdings plus the discounted expected value in the next period which depends on the return state variable x_t as well as the choice of holdings after x_t is revealed and minus the up-front adjustment cost that is incurred when asset holdings are switched from i to j . Wealth is not a state variable. Although the strategy requires an ability to borrow and corresponding margin, the scale of the strategy is unrestricted; so it makes sense to preserve the relative simplicity of the decision problem by assuming that the scale of investment is low enough that wealth/margin does not become an issue.

The optimal asset choice depends purely on the difference in returns between the two assets rather than on their common levels. Hence, we can specify parsimoniously and with little loss of generality:

$$E_{t-1}[r_2(x_t, \eta_{2t})] - E_{t-1}[r_1(x_t, \eta_{1t})] = E_{t-1} x_t \equiv \mu_t. \quad (3)$$

The implied general restrictions on the $r_i(x_t, \eta_{it})$ are that $r_1(x_t, \eta_{1t}) = -\frac{1}{2} x_t + g(x_t) + \eta_{1t}$ and

³ Given the assumed risk neutrality, symmetry, and proportional transaction costs, intermediate positions, with investment in both assets or neither asset, are never optimal.

$r_2(x_t, \eta_{2t}) = \frac{1}{2}x_t + g(x_t) + \eta_{2t}$ with $g(x)$ time-invariant but otherwise unrestricted. To rule out unbounded profits or losses, it is sufficient to assume that μ_t is finite for all t .

It is now convenient to define the “difference-in-value” function:

$$V(\mu_t) \equiv V_2(x_{t-1}, \varepsilon_{t-1}) - V_1(x_{t-1}, \varepsilon_{t-1}), \quad \mu_t = \rho x_{t-1} - \delta \varepsilon_{t-1}, \quad (4)$$

where the conditional expected excess return μ_t is defined in equation (3) and its expression is obtained from equation (2). Note that due to risk neutrality the “difference-in-value” function gives the “value of the difference”, i.e., the optimized value from a long-short strategy. The Appendix shows formally that the one state variable μ_t is sufficient to capture the dynamics related to the difference-in-value function.

The Appendix also shows that equation (4) for the difference between holding asset 2 and holding asset 1 can be written as:

$$V(\mu_t) = \mu_t + \int_{\varepsilon_{2t}^*}^{\varepsilon_{1t}^*} \frac{1}{1+r} V(\mu_{t+1}) f(\varepsilon_t) d\varepsilon_t + c \left[\int_{\varepsilon_{1t}^*}^{\infty} f(\varepsilon_t) d\varepsilon_t - \int_{-\infty}^{\varepsilon_{2t}^*} f(\varepsilon_t) d\varepsilon_t \right], \quad (5)$$

$$\text{with } \mu_{t+1} = \rho \mu_t + (\rho - \delta) \varepsilon_t, \quad (6)$$

as follows from equation (2). It may be clear from the set-up of the problem and more directly from equation (5) that the optimal investment choice conforms to an “S-s” type policy where a switch occurs if and only if the expected return difference lies outside the upper bound S and lower bound s, where $s = -S$. Our purpose in this paper, however, is to obtain a specific characterization of S for empirical purposes, which we provide in the following and which motivates our specific autocorrelation and symmetry assumptions in equations (2) and (3) above.

The optimal choice of the critical values ε_{1t}^* and ε_{2t}^* is found by differentiating the right-hand side of equation (5) using Leibniz’s rule, and is given by

$$\frac{1}{1+r} V[\rho \mu_t + (\rho - \delta) \varepsilon_{1t}^*] = c, \text{ and} \quad (7)$$

$$\frac{1}{1+r} V[\rho \mu_t + (\rho - \delta) \varepsilon_{2t}^*] = -c. \quad (8)$$

Intuitively, in equation (7), the investor is indifferent between switching assets and continuing to hold asset

1 if the discounted gain in value from switching is exactly equal to the up-front transaction cost. Equation (8) has a similar interpretation. Inverting the difference-in-value function yields for the critical values $\varepsilon_{1t}^* = \{V^{-1}[c(1+r)] - \rho \mu_t\} / (\rho - \delta)$ and $\varepsilon_{2t}^* = \{V^{-1}[-c(1+r)] - \rho \mu_t\} / (\rho - \delta)$.

Given equations (1) and (4) and the symmetry of the error density we have

$$V(\mu_t) = -V(-\mu_t). \quad (9)$$

Thus, given equation (7), (8), and (9), we have

$$\varepsilon_{1t}^* = -\varepsilon_{2t}^* - \frac{2\rho\mu_t}{\rho - \delta}. \quad (10)$$

Defining the critical mean returns, μ_1^* and μ_2^* , as $\frac{1}{1+r}V(\mu_1^*) = c$ and $\frac{1}{1+r}V(\mu_2^*) = -c$, we have $\mu^* \equiv \mu_1^* = -\mu_2^* = V^{-1}[c(1+r)]$. Hence, although the critical shocks ε_{1t}^* and ε_{2t}^* are time dependent, the critical mean excess returns are constant over time. Given that $V(0) = 0$ and $V'(\mu_t) > 1$:

$$0 < \mu^* < c(1+r). \quad (11)$$

Summarizing and expanding the above results, the Appendix proves the following:

Proposition 1. *For the decision problem in equations (1) - (3) and the definitions of $V(\mu)$ and μ^* above and given that $f(\varepsilon_t)$ has unbounded support:*

(a) *For $\rho > 0$: $V'(\mu) > 1 \forall \mu$, and $V''(\mu) \geq (\leq) 0 \leftrightarrow (\rho - \delta)\mu \leq (\geq) 0$.*

(b) *$V(\mu) = -V(-\mu) \forall \mu$.*

(c) *\exists unique μ^* : $0 < \mu^* < c(1+r)$.*

(d) *$\mu^* = \rho \mu_t + (\rho - \delta) \varepsilon_{1t}^*$, $\varepsilon_{1t}^* = -\varepsilon_{2t}^* - \frac{2\rho\mu_t}{\rho - \delta}$, $\mu^* = V^{-1}[c(1+r)]$.*

(e) *The investor buys asset 2 and shorts asset 1 if $\mu_{t+1} > \mu^*$ and vice versa if $\mu_{t+1} < -\mu^*$. If*

$\mu^ \leq \mu_{t+1} \leq -\mu^*$, the current position is maintained.*

Proof. See Appendix. \square

Result (a) implies that, if profit opportunities have positive persistence, then the value function has a slope that is everywhere larger than one and is concave for positive $(\rho - \delta)\mu_t$ and convex for negative $(\rho - \delta)\mu_t$. The slope exceeds one because an increase in the expected return difference has an immediate one-to-one effect on the differential between holding asset 2 and holding asset 1 plus an additional future positive effect due to the persistence of the expected return differential that applies if there is no switching which happens with positive probability. The concavity (convexity) for $(\rho - \delta)\mu_t \geq (\leq) 0$ arises because the higher the expected return differential is, the more incentive there is for the trader to switch from the asset with the lower mean to that with the higher mean at the fixed transaction cost, thus mitigating the expected return difference between the two positions.

Using an induction argument and the symmetry of $f(\varepsilon_t)$ in equation (2), result (b) follows. The reason is the symmetry of the two assets: holding asset 1 given mean return difference of μ_t is no different from holding asset 2 given mean return difference of $-\mu_t$.

Given the larger than unitary slope of the excess value function it follows that there is exactly one bounded value of μ^* so that equation (7) holds. Result (c) also states that μ^* must be positive. This is so because $V(0) = 0$ from result (b) so that the positive slope of $V(\mu)$ implies that equation (7) can only hold at $\mu^* > 0$; similarly, the larger than unitary slope of $V(\mu)$ implies that equation (7) can only hold at $\mu^* < c(1+r)$.

Redefine the last two terms on the right-hand side of equation (5) for convenience, using Proposition 1d:

$$V(\mu_t) = \mu_t + B(\mu_t, \mu^*) + C(\mu_t, \mu^*) \quad (12)$$

By converting from the left tail of the distribution to the right tail, we can write:

$$C(\mu_t, \mu^*) = c \left[\int_{\varepsilon_{1t}^*}^{-\varepsilon_{2t}^*} f(\varepsilon_t) d\varepsilon_t \right] = c \left[\int_{\frac{\mu^* - \rho \mu_t}{\rho - \delta}}^{\frac{\mu^* + \rho \mu_t}{\rho - \delta}} f(\varepsilon_t) d\varepsilon_t \right]. \quad (13)$$

The term $C(\mu, \mu^*)$ is interpreted as the expected transaction costs due to switching from asset 1 to asset

2 minus the expected transaction cost due to switching from asset 2 to asset 1.

The term $B(\mu, \mu^*)$ indicates the average difference in next period's value from holding asset 2 compared to holding asset 1, in the region where switching is not optimal, so that any expected return differential persists:

$$B(\mu_t, \mu^*) = \frac{1}{1+r} \int_{\frac{-\mu^* - \rho \mu_t}{\rho - \delta}}^{\frac{\mu^* - \rho \mu_t}{\rho - \delta}} V[\rho \mu_t + (\rho - \delta) \varepsilon_t] f(\varepsilon_t) d\varepsilon_t \quad (14)$$

$V(\mu)$ is evaluated from $-\mu^*$ to $+\mu^*$. Thus, given the symmetry of $V(\mu)$ in Figure 1, the sign of $B(\mu, \mu^*)$ depends on the shape of the density function. For $\mu_t > 0$ the positive part of $V(\mu)$ is weighted more than the negative part if the density is symmetric and unimodal (and vice versa for $\mu_t < 0$). Thus, $B(\mu, \mu^*)$ is positive for $\mu_t > 0$ unless the density function is flat (as is the case for the uniform distribution) in which case $B(\mu, \mu^*)$ is zero.

Evaluating $V(\mu)$ at the critical expected return difference μ^* , using equations (7) and (12) gives:

$$\mu^* = c(1+r) - B(\mu^*, \mu^*) - C(\mu^*, \mu^*). \quad (15)$$

Intuitively, equation (15) indicates the expected excess return from holding asset 2 instead of asset 1 at which the investor is indifferent between staying with asset 1 or switching to asset 2, μ^* , to be equal to the up-front transaction cost with interest plus the expected transaction cost at state μ^* of switching back in the next period from asset 2 to 1 minus the expected transaction cost of not switching initially but switching anyway in the next period.⁴

⁴ Given equation (15) or equation (7) characterizing the critical mean return differential and equation (5) stating the difference-in-value function, it is possible to derive the general comparative statics results for the critical mean excess return and the probability of no transaction with respect to all of the parameters in the model of equations (1) - (3). These results are available from the authors.

B. Closed Form Solution for Uniform Innovations

It is difficult to obtain an explicit analytical solution for the optimal filter in equation (15) because, from equation (14), $B(\mu, \mu^*)$ 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 $B(\mu, \mu^*)$ 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 $f(\varepsilon)=1/(2z)$. Proposition 1 requires minor modifications to apply in this case, which we omit for brevity, since the uniform distribution is bounded and not strictly unimodal. Equation (13) becomes:

$$C(\mu_t, \mu^*) = \int_{\text{Max}(-z, \frac{\mu^* - \rho \mu_t}{\rho - \delta})}^{\text{Min}(z, \frac{\mu^* + \rho \mu_t}{\rho - \delta})} c f(\varepsilon_t) d\varepsilon_t = \frac{c}{2z} [\text{Min}(z, \frac{\mu^* + \rho \mu_t}{\rho - \delta}) - \text{Max}(-z, \frac{\mu^* - \rho \mu_t}{\rho - \delta})] \quad (16)$$

Equation (14) becomes:

$$B(\mu_t, \mu^*) = \frac{1}{2z(1+r)} \int_{\text{Max}(-z, \frac{\mu^* - \rho \mu_t}{\rho - \delta})}^{\text{Min}(z, \frac{\mu^* + \rho \mu_t}{\rho - \delta})} V[\rho \mu_t + (\rho - \delta) \varepsilon_t] d\varepsilon_t \quad (17)$$

Note that the Min and Max operators appear because $f(\varepsilon) = 0$ outside of the domain $[-z, z]$.

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

$$z \geq \frac{[1+r(1+\rho)]c}{\rho - \delta}, \quad (18)$$

then the critical mean return differential in the model of equations (1) - (3) is given by

$$\mu^* = \frac{(1+r)c}{1 + \{\rho c / [(\rho - \delta)z]\}}. \quad (19)$$

Proof. If $z \geq (1+\rho)\mu^*/(\rho - \delta)$, then the bounds in $B(\mu^*, \mu^*)$ and $C(\mu^*, \mu^*)$ are interior. Hence:

$B(\mu^*, \mu^*) = 0$ by Proposition 1b and equation (17), and $C(\mu^*, \mu^*) = c \rho \mu^* / [(\rho - \delta)z]$ from equation

(16). Equation (15) then implies equation (19). Equations (18) and (19) in turn imply the premise that $z \geq (1 + \rho) \mu^* / (\rho - \delta)$. \square

Note that the assumption of the uniform distribution leads to relatively high values for μ^* because it causes $B(\mu^*, \mu^*)$ to be equal to its minimum value. Equation (19) states that given $(\rho - \delta)$ fixed, 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 equation (6) 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.

II. 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 equation (2), the μ_t in equation (6) 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 excess return $V(\mu_t)$. In a fully efficient market we may then expect $V(\mu_t) = 0$.⁵ Second, we may think of μ_t as the risk-adjusted expected excess return, the “alpha”, so that positive $V(\mu_t)$ represents an average excess return adjusted for systematic risk. Similarly, μ_t could represent a particular expected utility level, which would also account for risk. Thirdly, 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

⁵ In fact, imposing this condition would give an alternative approach to estimating the transaction cost c if we assume that, in an efficient market, the after-transaction-cost excess returns would be nonnegative (zero) only when using the optimal strategy.

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.

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.

A. A Parametric Moving Average Trading Strategy

As discussed extensively in the literature, see for instance Cornell and Dietrich (1978), Frankel and Froot (1990), LeBaron (1998, 1999), Lee and Mathur (1996), Levich and Thomas (1993), Qi and Wu (2004), Sweeney (1986), and others, profitable trading strategies in foreign exchange markets traditionally have employed moving-average (MA) technical trading rules.⁶ 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 model behind the MA trading rule is

$$\Delta s_{t+1} = \lambda \left(s_t - \left[\left(\sum_{i=1}^N s_{t-i} \right) / N \right] \right) + \varepsilon_{t+1}, \quad E_t \varepsilon_{t+1} = 0, \quad \lambda > 0. \quad (20)$$

For any positive λ , equation (20) 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 equation (20) 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 .⁷ In what follows, our strategy is to transform equation (20) to infer empirically a value

⁶ The 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. See for instance Sweeney (2000) and Taylor (1982).

⁷ Given equation (20) 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 $x\%$ induces a switch. In our case, for a naïve filter equal to c ,

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 equation (20) 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 which exhibits mild 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 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.

Equation (20) can straightforwardly be rewritten as an autoregressive process in the percentage change in exchange rate:

$$\Delta s_{t+1} = \lambda \left(\sum_{i=0}^{N-1} \frac{N-i}{N} \Delta s_{t-i} \right) + \varepsilon_{t+1}, \quad (21)$$

where the coefficients in the autoregression are given by the Bartlett weights.

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 Equation (21) becomes:

for instance, we need $\lambda x > c$ to induce a switch. So, for the numbers we find in our empirical section for the MA(21) process, our naïve filter corresponds to $x = c/\lambda = 0.1/0.025=4\%$, a 4% band for the *ad hoc* filter.

$$x_{t+1} = \lambda \left(\sum_{i=0}^{N-1} \frac{N-i}{N} x_{t-i} \right) + \varepsilon_{t+1}, \quad (22)$$

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).⁸

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 equation (22) 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 equation (22) 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. For the remainder of the paper, we assume the parameter restriction $\rho > \delta \geq 0$ for the ARMA(1,1) process (2). This restriction is satisfied for all our empirical specifications.

Inverting the ARMA(1,1) process of equation (2) yields an alternative autoregression:

$$x_{t+1} = (\rho - \delta) \sum_{i=0}^{\infty} \delta^i x_{t-i} + \varepsilon_{t+1} \quad (23)$$

Comparing equations (22) and (23) it follows that equation (23) is a good approximation for equation (22) if we set $\rho - \delta = \lambda$ and if the $(N-i)/N$ terms are close to the δ^i for all i . Taking a natural log approximation, and choosing δ to match the Bartlett weights:

$$-i/N \approx \ln[(N-i)/N] = i \ln(\delta) \quad \Rightarrow \quad \delta = \exp(-1/N). \quad (24)$$

Thus, for a given N , we run regression (22) to obtain an estimate of λ . Then from $\delta = \exp(-1/N)$ and

⁸ Deviations from uncovered interest parity and their persistence are documented in previous research. See, for example, Canova (1991) and Mark and Wu (1998).

$\rho = \lambda + \delta$, 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.

B. 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 6195 daily observations. Daily exchange rate data for the Japanese yen are downloaded from the Federal Reserve’s webpage. For interest rates, we obtain *Financial Times*’ Euro-currency interest rates from *Datastream International*. The first 1/3 of the sample (2065 observations) is used for model estimation. Out-of-sample forecasting starts on November 28, 1986 until the end of the sample (4130 observations). We estimate the exchange rate dynamics in four ways: 1. an AR(1) process (δ in equation 2 is set equal to zero); 2. an ARMA(1,1) process (equation 2); 3. a process consistent with an MA rule of 21 lags (21 trading days in a month), as is commonly considered with daily data (equation 22); 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. Models 3 and 4 are commonly employed by academia and practitioners. The ARMA(1,1) is the exact model assumed in our theoretical derivation, while the AR(1) model is a more parsimonious specification and is a natural candidate. It is important to reiterate that the primary purpose of this paper is not to search extensively for the best possible models to predict exchange rates. Rather our main focus is to provide an *ex ante* model-specific optimal filter for some popular models commonly used in academia and by practitioners and to compare the performance of our theoretical optimal filter with other *ad hoc* filters.

Columns (1)-(5) of Table I show the results of the in-sample model estimation using the first 1/3 of our sample for each way of capturing the exchange rate dynamics. For the AR(1) process we find low persistence with $\rho = 0.0548$ and a standard error of $\sigma_e = 0.00659$; for the ARMA(1,1) process we find

$\rho = 0.918$ and $\delta = 0.880$ with a standard error of $\sigma_\varepsilon = 0.00658$. Thus, both processes provide similar accuracy although the parameters differ substantially.⁹ 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.¹⁰ For the representative moving average rule with 21 lags we find for the slope in equation (22) that $\lambda = 0.0257$ and $\sigma_\varepsilon = 0.00656$. Since we have $N = 21$ we obtain from equation (24) 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 a round trip transaction cost of $c = 0.001$ (10 basis points) throughout. 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 1/3 of the sample equal to 0.000439 percent. This average interest rate is used as a proxy for the discount rate r in computing the optimal filter in Equation (19).

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. Firstly, under the assumption that the error term ε_t is uniformly distributed the optimal filter, denoted by μ_U^* , can be analytically calculated using equation (19). Secondly, ε_t is assumed to follow a normal distribution. In this case, the result in equation (19) no longer holds, and we

⁹ Previous 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.

¹⁰ Balvers and Mitchell (1997) raise this issue in the context of optimal portfolio choice under return predictability.

estimate the optimal filter, denoted by μ_N^* , through Monte-Carlo simulation. Lastly, we do not make an assumption about the distribution of ε_t and estimate the optimal filter, denoted by μ_B^* , by bootstrapping the model residuals $\hat{\varepsilon}_t$ with replacement.

C. 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 equation (19) for the dollar-yen exchange rate. Column (6) of Table I provides the results. For the AR(1) case we find that the ratio of the critical return to the transaction cost is $\mu_U^* / c = 0.92$. 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 equation (6), the 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.

On the other hand, for the ARMA(1,1) case with $\rho = 0.918$ and $\delta = 0.880$, we find that the optimal filter is dramatically different at $\mu_U^* / c = 0.32$. The reason that this number is so much lower than under the AR(1) case is clear from equation (6). The persistence is not only high now with $\rho = 0.918$ but it is also high relative to the innovation in the conditional mean that is $(\rho - \delta) \sigma_\varepsilon = 0.038 \sigma_\varepsilon$. 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.

For the 1-month MA process, the parameters backed out from the MA(21) model yield $\mu_U^* / c = 0.23$. Note that inequality (18) is violated, as is necessary when $\mu_U^* / c < 0.50$, implying that the analytical value obtained from equation (19) 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 undoubtedly a reason for the popularity of this particular trading rule with practitioners.

D. 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 Equation (25) with parameters estimated from the first 1/3 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 I reports the median value of the optimal filter to transaction cost ratio, μ_N^* / c , over the 500 Monte-Carlo trials. For each model, the ratio μ_N^* / c is quite close to the optimal ratio implied under the uniform distribution μ_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 Equation (25) 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 which maximizes the after-cost average excess return. Column (8) of Table I reports the median estimate of the optimal filter to cost ratio μ_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.

III. Out-of-Sample Optimal Switching Strategy Results

We start our first-day forecast on November 28, 1986 (after the first one third of the sample).¹¹ For each of the four exchange rate return specifications, we estimate the model parameters using all observations for the first one third of the sample (up to November 27, 1986) and make the first forecast (for November 28). If the forecasted excess return (recall that the excess 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 Japanese yen, and simultaneously take a short position in the dollar. Conversely, if the forecasted excess 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 excess return from a zero-cost investment strategy. While daily data are employed in this study, we do not view non-synchronous price information to be a serious issue due to the heavy trading volume of the Japanese yen.

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

¹¹ Our 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 $\frac{1}{3}$.

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 round-trip transaction costs.¹² (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 round-trip 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 and after-cost excess return rates for the case without a filter; and the after-cost excess return rates for the cases with transaction cost filters. 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 $+0.00953$), less than one percent. This return is not statistically distinguishable from zero (t -ratio = 0.348).

Table II reports the results for the four forecasting models. Each model will be discussed separately below. For the AR(1) model, without imposing a transaction cost filter, the strategy involves 2,100 switches over 4,130 trading days (over 50 percent of the time). In the absence of transaction costs, the strategy produces an annualized excess return of 6.9 percent with a t -ratio of 2.486 which is statistically significant at the 5 percent level using a 2-sided test. However, a round-trip cost of 10 basis points completely wipes out the profits, resulting in a negative excess return of 18.8 percent. A naïve filter equal to the actual transaction cost of 10 basis points dramatically reduces the number of transactions to 34, and

¹² The transaction cost c in the theory is equivalent to the cost of one round-trip from equation (1) where c represents the cost of switching from a long position in currency i to a long position in j . The empirical case, however, is based on the difference-in-value function in equation (5) in which long and short position are combined, implying two roundtrips when a transaction occurs. An intermediate strategy – closing out one position (long or short) but maintaining the other position – is never optimal. As a result the overall return is always equal to the difference between the long and short position returns.

yields an insignificant excess return of 0.9 percent 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 5.6 percent which is significant at the 5 percent level. Similarly, the optimal filter under the normality assumption μ_N^* produces an average excess return of 5.3 percent per annum which is significant at the 10 percent level. The bootstrapped filter μ_B^* yields an insignificant excess return of 2.2 percent.

For the ARMA(1,1) model, the after-cost excess return for the case without a filter is 0.3 percent with 444 trades. The naïve filter equal to “ c ” dramatically reduces the number of trades to 12, resulting in an insignificant excess return of 0.5 percent. In contrast, the optimal filter, μ_U^* , captures many of the profitable trades and yields an excess return of 5.1 percent with a t -ratio of 1.827 which is significant at the 10 percent level in a two-tailed test (5 percent level in a one-tailed test). The filter under the bootstrapped distribution μ_B^* produces nearly the same results as μ_U^* , whereas the filter under the normality assumption, μ_N^* , yields a higher excess return of 6.6 percent which is significant at the 5 percent level.

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.

Finally, the long MA(126) process provides very smooth forecast of expected returns. While the strategy without filter yields an after-cost return of 4.9 percent which is significant at the 10 percent level, the naïve filter equal to c skips too many profitable trades, resulting in a negative return of 2 percent. 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 5.7 percent which is significant at the 5 percent level. The other two filters, μ_N^* and μ_B^* yield

somewhat smaller returns.¹³

We want to emphasize again that we are not searching for the best trading strategy. Our key point is that an optimal transaction filter can be designed to maximize the after-cost expected profitability of a particular trading strategy. Whatever the before-cost trading strategy returns, we expect the optimal filter to improve after-trading-cost profits, and to perform well compared to *ad hoc* filters, if the assumed exchange rate process is a reasonable approximation of reality.

Figures 1 through 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. 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 maximum filter value is quite close to the 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 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.

Figures 5 through 8 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.

¹³ The trading strategies for each of the forecasting models imply a reasonably even choice of each currency. For instance, with the optimal filter μ_U^* , the fraction of long Japanese yen and short U.S. dollar is: AR(1) 2555/4130, ARMA(1,1) 1830/4130, MA(21) 1987/4130, and MA(216) 1875/4130.

Table III provides the risk-adjusted trading rule returns. We correct the *ex post* trading rule returns from each of the four forecast models for market risk using the MSCI World market index and the Euro dollar 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.

IV. Conclusion

If transitory profitable trading opportunities exist, filter rules are used in practice to mitigate transaction costs. 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.

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 an ARMA(1,1) process the optimal filter is only around 30 percent 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 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 strongly suggests that caution should be applied in dismissing abnormal returns as due to transactions costs, merely because the after-cost return is negative or insignificant. For instance, Lesmond, Schill, and Zhou (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.

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, Schill, and Zhou (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 can not 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 first step in deriving a unique filter that can be related to observable variables.

Appendix

A. The Difference-In-Value Function

From equations (1) and (2):

$$V_i(x_{t-1}, \varepsilon_{t-1}) = E_{t-1} r_i(x_t, \eta_{it}) + \int_{\{\varepsilon_t: \frac{1}{1+r} V_{it} > \frac{1}{1+r} V_{jt} - c\}} \frac{1}{1+r} V_i(x_t, \varepsilon_t) f(\varepsilon_t) d\varepsilon_t + \int_{\{\varepsilon_t: \frac{1}{1+r} V_{it} \leq \frac{1}{1+r} V_{jt} - c\}} \left[\frac{1}{1+r} V_j(x_t, \varepsilon_t) - c \right] f(\varepsilon_t) d\varepsilon_t, \quad (\text{A1})$$

with $x_t = \rho x_{t-1} - \delta \varepsilon_{t-1} + \varepsilon_t$, $E_{t-1} \varepsilon_t = 0$, $\rho \neq \delta$, $f(\varepsilon_t) = f(-\varepsilon_t)$.

Taking the difference $V_{diff}(x, \varepsilon) = V_2(x, \varepsilon) - V_1(x, \varepsilon)$ and using equation (3) yields:

$$V_{diff}(x_{t-1}, \varepsilon_{t-1}) = E_{t-1} x_t + \int_{\{\varepsilon_t: c \geq \frac{1}{1+r} V_{diff} \geq -c\}} \frac{1}{1+r} V_{diff}(x_t, \varepsilon_t) f(\varepsilon_t) d\varepsilon_t + c \left[\int_{\{\varepsilon_t: \frac{1}{1+r} V_{diff} > -c\}} f(\varepsilon_t) d\varepsilon_t - \int_{\{\varepsilon_t: c > \frac{1}{1+r} V_{diff}\}} f(\varepsilon_t) d\varepsilon_t \right]. \quad (\text{A2})$$

Redefining the state variables so that $V_{diff}(x_t, \varepsilon_t) \equiv V_d(\mu_{t+1}, \varepsilon_t)$ is without loss of generality since

$\mu_{t+1} \equiv \rho x_t - \delta \varepsilon_t$ and yields:

$$V_d(\mu_t, \varepsilon_{t-1}) = \mu_t + \int_{\{\varepsilon_t: c \geq \frac{1}{1+r} V_d \geq -c\}} \frac{1}{1+r} V_d(\mu_{t+1}, \varepsilon_t) f(\varepsilon_t) d\varepsilon_t + c \left[\int_{\{\varepsilon_t: \frac{1}{1+r} V_d > -c\}} f(\varepsilon_t) d\varepsilon_t - \int_{\{\varepsilon_t: c > \frac{1}{1+r} V_d\}} f(\varepsilon_t) d\varepsilon_t \right], \quad (\text{A3})$$

with $\mu_{t+1} = \rho \mu_t + (\rho - \delta) \varepsilon_t$, $E_{t-1} \varepsilon_t = 0$, $\rho \neq \delta$, $f(\varepsilon_t) = f(-\varepsilon_t)$.

It is clear from (A3) that, once the state variable μ_t is considered, there is no additional role for the state variable ε_{t-1} . Thus we have

$$V(\mu_t) \equiv V_d(\mu_t, \varepsilon_{t-1}). \quad (\text{A4})$$

B. Proof of Proposition 1

It is easy to see that $V(\mu_t)$ is monotonically increasing in μ_t and that, in effect, for $\rho > 0$ a one unit increase in μ_t raises $V(\mu_t)$ more than one unit: there is a direct one-to-one benefit of higher μ_t given that asset 2 is held, plus the positive persistence ($\rho > 0$ in A3) means that an additional benefit also is expected to extend to the future. If we assume that the innovation ε_t 's has unbounded support, then since $\mu_{t+1} = \rho \mu_t + (\rho - \delta) \varepsilon_t$ and $\rho \neq \delta$, any value of μ_{t+1} is possible. Hence, a critical $\varepsilon_t = \varepsilon_{1t}^*$ must

exist such that $\frac{1}{1+r}V[\rho\mu_t + (\rho-\delta)\varepsilon_{1t}^*] = c$ and a critical $\varepsilon_t = \varepsilon_{2t}^*$ must exist such that $\frac{1}{1+r}V[\rho\mu_t + (\rho-\delta)\varepsilon_{2t}^*] = -c$. Moreover, these critical ε_t are unique due to the monotonicity of $V(\mu_t)$.

Thus we can write:

$$V(\mu_t) = \mu_t + \int_{\varepsilon_{2t}^*}^{\varepsilon_{1t}^*} \frac{1}{1+r} V(\mu_{t+1}) f(\varepsilon_t) d\varepsilon_t - c \left[\int_{-\infty}^{\varepsilon_{2t}^*} f(\varepsilon_t) d\varepsilon_t - \int_{\varepsilon_{1t}^*}^{\infty} f(\varepsilon_t) d\varepsilon_t \right] \quad (\text{A5})$$

with $\mu_{t+1} = \rho\mu_t + (\rho-\delta)\varepsilon_t$, $\varepsilon_{1t}^* = \{V^{-1}[c(1+r)] - \rho\mu_t\}/(\rho-\delta)$, and

$$\varepsilon_{2t}^* = \{V^{-1}[-c(1+r)] - \rho\mu_t\}/(\rho-\delta).$$

Given (A4) and (A5) and since $f(\varepsilon_t) = f(-\varepsilon_t)$ it follows that

$$V(\mu_t) = -V(-\mu_t). \quad (\text{A6})$$

Thus, given (A6) and equations (7) and (8) we have

$$\varepsilon_{1t}^* = -\varepsilon_{2t}^* - \frac{2\rho\mu_t}{\rho-\delta}. \quad (\text{A7})$$

Defining the critical mean returns as $\frac{1}{1+r}V(\mu_1^*) = c$ and $\frac{1}{1+r}V(\mu_2^*) = -c$, we have

$\mu^* \equiv \mu_1^* = -\mu_2^* = V^{-1}[c(1+r)]$. Hence, given that $V(0)=0$ and $V'(\mu_t) > 1$:

$$0 < \mu^* < c(1+r) \quad (\text{A8})$$

Concavity of $V(\mu_t)$ for $\mu_t > 0$ follows from differentiation of (A5), using Leibniz's rule:

$$V'(\mu_t) = 1 + \frac{\rho}{1+r} \int_{\varepsilon_{2t}^*}^{\varepsilon_{1t}^*} V'(\mu_{t+1}) f(\varepsilon_t) d\varepsilon_t. \quad (\text{A9})$$

Again differentiating yields:

$$V''(\mu_t) = \frac{-\rho^2}{(\rho-\delta)(1+r)} V'(\mu^*) [f(\varepsilon_1^*) - f(\varepsilon_2^*)] + \frac{\rho^2}{1+r} \int_{\varepsilon_{2t}^*}^{\varepsilon_{1t}^*} V''(\mu_{t+1}) f(\varepsilon_t) d\varepsilon_t, \quad (\text{A10})$$

where we used the fact that from (A6) $V'(\mu_t) = V'(-\mu_t)$. Given that $f(\varepsilon_t)$ is symmetric and unimodal we have for $\rho > 0$ that $f(\varepsilon_1^*) > f(\varepsilon_2^*)$ if $\mu_t > 0$ so that $V''(\mu_t) < 0$ if $\rho-\delta > 0$ and $V''(\mu_t) > 0$ if $\rho-\delta < 0$. Thus, from (A6), for $\mu_t < 0$ we have $V''(\mu_t) > 0$ if $\rho-\delta > 0$ and $V''(\mu_t) < 0$ if $\rho-\delta < 0$. \square

References

- Alexander, Sidney S., 1961, Price movements in speculative markets: trends or random walks, *Industrial Management Review* 2, 7-26.
- Balduzzi, Pierluigi, and Anthony W. Lynch, 1999, Transaction costs and predictability: some utility cost calculations, *Journal of Financial Economics* 52, 47-78.
- Balvers, Ronald J., and Douglas W. Mitchell, 1997, Autocorrelated returns and optimal intertemporal portfolio choice, *Management Science* 43, 1537-1551.
- Balvers, Ronald J., Yangru Wu, and Erik Gilliland, 2000, Mean reversion across national stock markets and parametric contrarian investment strategies, *Journal of Finance* 55, 745-772.
- Bessembinder, Hendrik, 1994, Bid-ask spreads in the interbank foreign exchange markets, *Journal of Financial Economics* 35, 317-348.
- Canova, Fabio, 1991, An empirical analysis of ex ante profits from forward speculation in foreign exchange markets, *Review of Economics and Statistics* 73, 489-96.
- Cheung, Yin-wong, and Clement Yuk-Pang Wong, 2000, A survey of market practitioners' view on exchange rate dynamics, *Journal of International Economics* 51, 401-419.
- Cooper, Michael, 1999, Filter rules based on price and volume in individual security overreaction, *Review of Financial Studies* 12, 901-935.
- Cornell, W. Bradford, and J. Kimball Dietrich, 1978, The efficiency of the market for foreign exchange under floating exchange rates, *Review of Economics and Statistics* 60, 111-120.
- Fama, Eugene F., and Marshall Blume, 1966, Filter rules and stock market trading profits, *Journal of Business* 39, 226-241.
- Frankel, Jeffrey A., and Kenneth A. Froot, 1990, Chartists, fundamentalists, and trading in the foreign exchange market, *American Economic Review* 80, 181-185.
- Grundy, Bruce, and J. Spencer Martin, 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. Douglas, and Mark J. Ready, 2001, Profitable predictability in the cross-section of stock returns, *mimeo*, University of Chicago.
- Knez, Peter J., and Mark J. Ready, 1996, Estimating the profits from trading strategies, *Review of Financial Studies* 9, 1121-1163.
- LeBaron, Blake, 1992, Do moving average trading rule results imply nonlinearities in foreign exchange markets? Technical report, University of Wisconsin-Madison.
- LeBaron, Blake, 1998, Technical trading rules and regime shifts in foreign exchange, In *Advanced Trading Rules*, eds. F. Acar and S. Satchell, 5-40, Butterworth-Heinemann, Woburn, MA.
- LeBaron, Blake, 1999, Technical trading rule profitability and foreign exchange intervention, *Journal of International Economics* 49, 125-143.

- Lee, Chun I., and Ike Mathur, 1996, Trading rule profits in European currency spot crossrates, *Journal of Banking and Finance* **20**, 949-962.
- Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* **71**, 349-380.
- Levich, Richard M., and Lee R. Thomas, 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, Hong, 2004, Optimal consumption and investment with transaction costs and multiple risky assets, *Journal of Finance* **59**, 289-338.
- Liu, Hong, and Mark Loewenstein, 2002, Optimal portfolio selection with transaction costs and finite horizons, *Review of Financial Studies* **15**, 805-835.
- Lynch, Anthony W., and Pierluigi Balduzzi, 2000, Predictability and transaction costs: the impact on rebalancing rules and behavior, *Journal of Finance* **55**, 2285-2310.
- Lynch, Anthony W., and Sinan Tan, 2002, Multiple risky assets, transaction costs and return predictability: Implications for portfolio choice, Working Paper, New York University.
- Maddala, G.S., 1983, *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York.
- Mark, Nelson C., and Yangru Wu, 1998, Rethinking deviations from uncovered interest parity: the role of covariance risk and noise, *The Economic Journal* **108**, 1686-1706.
- Marquering, Wessel, and Marno Verbeek, 2004, The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis* **39**, 407-429.
- Melvin, Michael, and Kok-Hui Tan, 1996, Foreign exchange market bid-ask spreads and the market price of social unrest, *Oxford Economic Papers* **48**, 329-341.
- Qi, Min, and Yangru Wu, 2004, Technical trading-rule profitability, data snooping, and reality check: evidence from the foreign exchange market, *mimeo*, Rutgers University.
- Sweeney, Richard J., 1986, Beating the foreign exchange market, *Journal of Finance* **41**, 163-182.
- Sweeney, Richard J., 2000, Does the Fed beat the foreign-exchange market? *Journal of Banking and Finance* **24**, 665-694.
- Taylor, Dean, 1982, Official intervention in the foreign exchange market, or, bet against the central bank, *Journal of Political Economy* **90**, 356-368.
- Taylor, Stephen J. 1992, Rewards available to currency futures speculators: compensation for risk or evidence of inefficient pricing? *Economic Record* **68** (Supplement), 105-116.

Table I 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 AR(1): $x_{t+1} = \rho x_t + \varepsilon_{t+1}$; ARMA(1,1): $x_{t+1} = \rho x_t + \varepsilon_{t+1} - \delta \varepsilon_t$; and the transformed model of MA(N): $x_{t+1} = \lambda \sum_{i=0}^{N-1} [(\frac{N-i}{N})x_{t-i}] + \varepsilon_{t+1}$, 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 1/3 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 equation (19); 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 $c = 0.1$ percent. Numbers inside parentheses are t -ratios.

	ρ	δ	λ	σ_ε	μ_U^* / c	μ_N^* / c	μ_B^* / c
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AR(1) Model	0.0548 (2.494)			0.00659	0.92	0.88	0.85
ARMA(1,1) Model	0.918 (19.232)	0.880 (15.665)		0.00658	0.32	0.34	0.32
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

Table II 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 $c = 0.1$ percent. The data cover the period from August 31, 1978 to May 3, 2003 with 6,195 daily observations. The first 1/3 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 naive 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^*				
	# of trade	excess return	t-ratio	excess return	t-ratio	# of trade	excess return	t-ratio	# of trade	excess return	t-ratio	# of trade	excess return	t-ratio	# of trade	excess return	t-ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
AR(1)	2100	0.069	2.486	-0.188	-6.702	34	0.009	0.315	42	0.056	2.011	48	0.053	1.910	66	0.022	0.794
ARMA(1,1)	444	0.057	2.054	0.003	0.093	12	0.005	0.199	81	0.051	1.827	71	0.066	2.391	81	0.052	1.870
MA(21)	448	0.040	1.453	-0.015	-0.522	12	-0.016	-0.586	125	0.036	1.314	123	0.037	1.323	125	0.040	1.451
MA(126)	131	0.065	2.338	0.049	1.756	5	-0.020	-0.725	31	0.057	2.074	29	0.046	1.650	29	0.046	1.650

Table III Risk Adjustment on Trading Performance in Foreign Exchange

This table reports after-cost average excess return (from the zero-cost investment strategy), *t*-ratio of excess return, market risk loading of excess return (β), systematic risk-adjusted excess return (α), and *t*-ratio of α for trading strategies with three optimal transaction cost filters with a round-trip transaction cost $c = 0.1$ percent. The data cover the period from August 31, 1978 to May 3, 2003 with 6,195 daily observations. The first 1/3 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). In estimating the CAPM, we use the MSCI world equity index as the market portfolio, and the Euro dollar interest rate as the risk-free rate. All returns and α measures are annualized.

Model	with optimal filter μ_U^*					with optimal filter μ_N^*					with optimal filter μ_B^*				
	return	t-ratio	β	α	t_α	Return	t-ratio	β	α	t_α	return	t-ratio	β	α	t_α
AR(1)	0.056	2.011	0.054	0.056	2.021	0.053	1.910	0.048	0.053	1.918	0.022	0.794	-0.007	0.022	0.794
ARMA(1,1)	0.051	1.827	-0.043	0.050	1.825	0.066	2.391	-0.044	0.066	2.390	0.052	1.870	-0.043	0.052	1.869
MA(21)	0.036	1.314	-0.022	0.036	1.312	0.037	1.323	-0.022	0.037	1.322	0.040	1.451	-0.025	0.040	1.450
MA(126)	0.057	2.074	-0.029	0.057	2.072	0.046	1.650	-0.035	0.046	1.649	0.046	1.650	-0.035	0.046	1.649

Fig 1. After-cost Excess Return (round) and Trading Cost (square) vs. μ/c ,
 Vertical Lines: Solid μ_V^*/c , Dash μ_N^*/c , Long Dash μ_B^*/c
 $c=0.1\%$, AR(1) Model

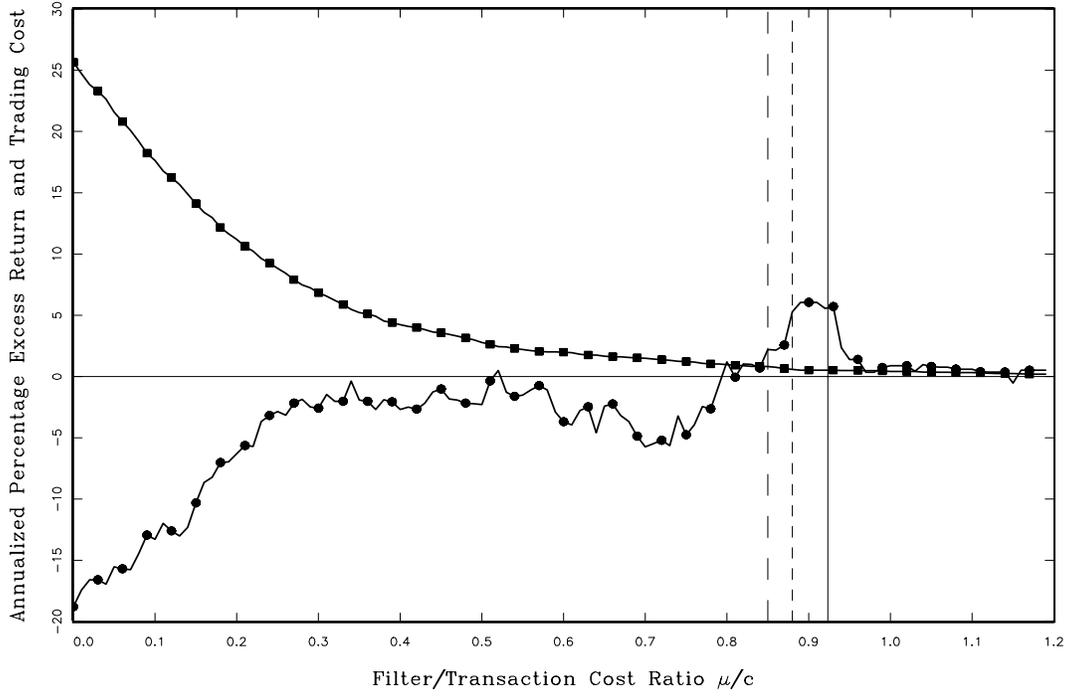


Fig 2. After-cost Excess Return (round) and Trading Cost (square) vs. μ/c ,
 Vertical Lines: Solid μ_V^*/c , Dash μ_N^*/c , Long Dash μ_B^*/c
 $c=0.1\%$, ARMA(1,1) Model

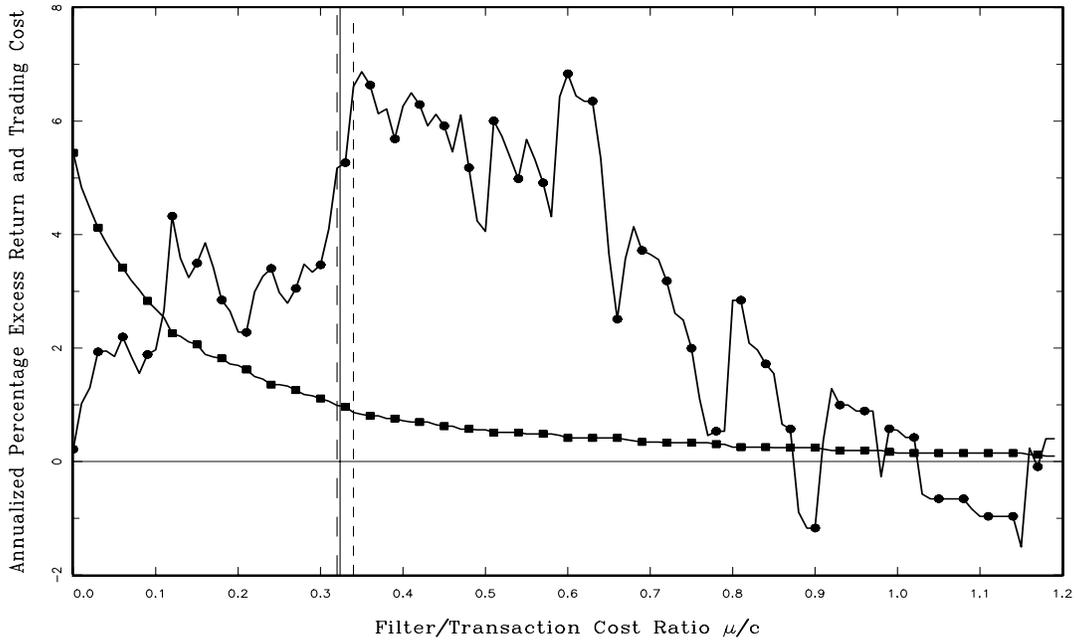


Fig 3. After-cost Excess Return (round) and Trading Cost (square) vs. μ/c ,
 Vertical Lines: Solid μ_U^*/c , Dash μ_N^*/c , Long Dash μ_B^*/c
 $c=0.1\%$, MA(21) Model

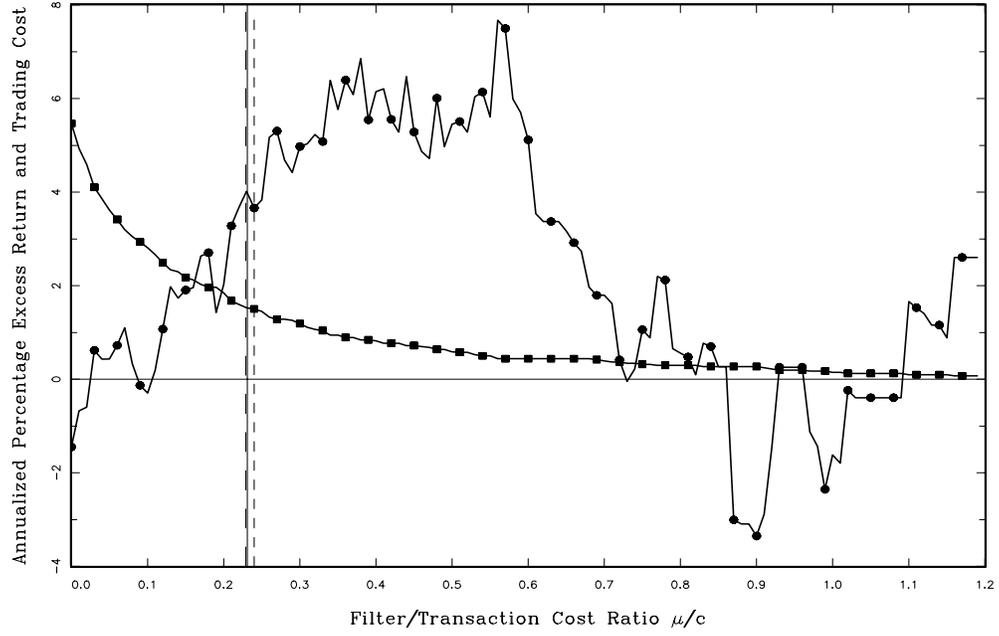


Fig 4. After-cost Excess Return (round) and Trading Cost (square) vs. μ/c ,
 Vertical Lines: Solid μ_U^*/c , Dash μ_N^*/c , Long Dash μ_B^*/c
 $c=0.1\%$, MA(126) Model

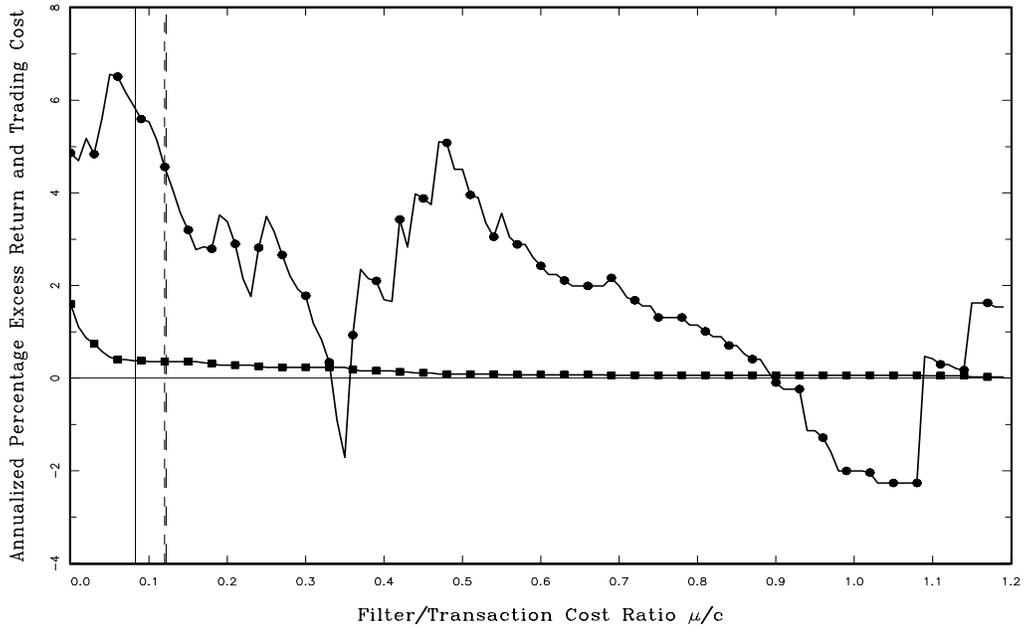


Fig 5. Trading Frequencies by Year: AR(1) Model

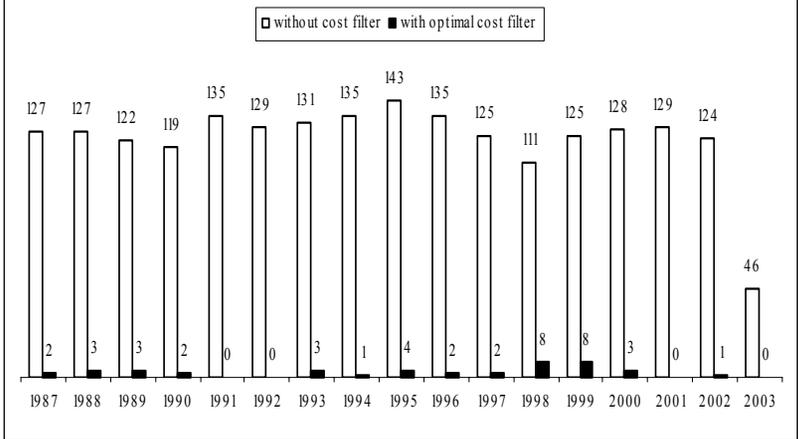


Fig 6. Trading Frequencies by Year: ARMA(1,1) Model

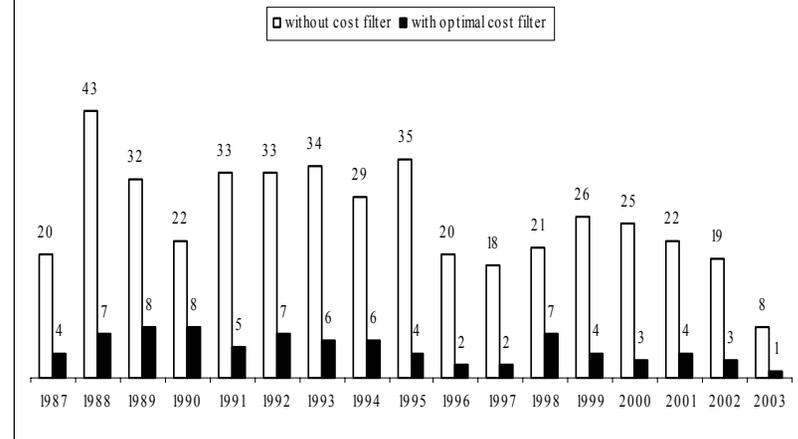


Fig 7. Trading Frequencies by Year: MA(21) Model

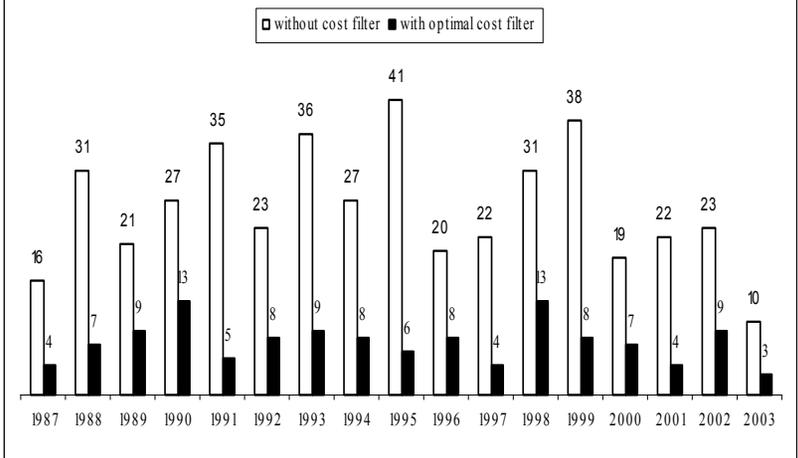


Fig 8. Trading Frequencies by Year: MA(126) Model

