



ELSEVIER

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf



Evaluating foreign exchange market intervention: Self-selection, counterfactuals and average treatment effects

Rasmus Fatum^a, Michael M. Hutchison^{b,*}

^a School of Business, University of Alberta, Edmonton, Alberta T6 G 2R6, Canada

^b Department of Economics, University of California, Santa Cruz, CA 95064, USA

A B S T R A C T

JEL Classification:

E58

F31

G15

Keywords:

Foreign exchange intervention

Bank of Japan

Self-selection

Matching methods

Estimating the effect of official foreign exchange market intervention is complicated by the fact that intervention at any point entails a “self-selection” choice made by the authorities and that no counterfactual is observed. To address these issues, we estimate the “counterfactual” exchange rate movement in the absence of intervention by introducing the method of propensity-score matching to estimate the “average treatment effect” (ATE) of intervention. To derive the propensity scores we estimate central bank intervention reaction functions. We estimate the ATE for daily official intervention in Japan over the January 1999–March 2004 period. This sample encompasses a remarkable variation in intervention frequencies as well as unprecedented frequent intervention towards the latter part of the period. We find that only sporadic and relatively infrequent intervention is effective.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Intervention is not a random occurrence but a process where officials “self select” in deciding when to intervene. Since an exchange rate movement at any given point in time coincides with either intervention or no intervention, we cannot observe both what is the exchange rate movement coinciding with intervention and what would have been the “counterfactual”, i.e. what would have been the exchange rate movement if intervention had not occurred when in fact it did. In other words, the counterfactual is not directly observable and, as such, this constitutes a missing data problem.

* Corresponding author. Tel.: +1 831 459 2600; fax: +1 831 459 5077.

E-mail address: hutch@ucsc.edu (M.M. Hutchison).

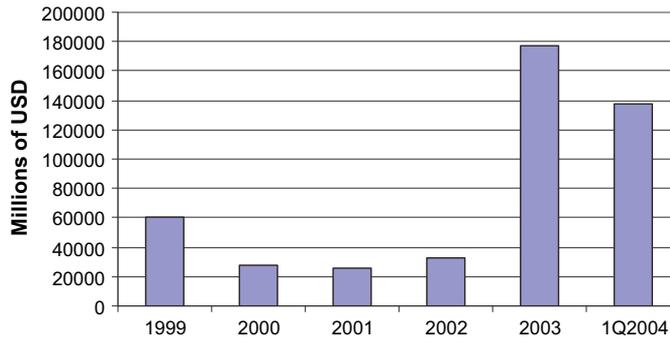


Fig. 1. Official Japanese Intervention 1999–2004. Notes: a) Yearly aggregates of daily intervention in the JPY/USD exchange rate market. The daily intervention data obtained from the Japanese Ministry of Finance data bank. b) There has been no Japanese intervention since March 2004.

These inherent issues of self-selection and missing data complicate the assessment of the effects of intervention. Following the modern literature on treatment effects, we address the issue of self-selection and the missing counterfactual by estimating the “average treatment effect” (ATE) of intervention in the JPY/USD exchange rate over the 1999–2004 period using a propensity-score matching methodology.

The approach taken here to evaluate the effectiveness of intervention, while addressing the aforementioned methodological issues, is to postulate a counterfactual and, in turn, match pairs of observations (or an average of control observations) of exchange rate movements – each pair consisting of an exchange rate movement coinciding with intervention and one that coincides with no intervention – on similar observable characteristics. Although the JPY appreciated strongly against the USD over our sample period, intervention may still have been effective in reducing the magnitude of this appreciation. This highlights the necessity of estimating a relevant counterfactual. We consider intervention as a “treatment” and, using matched counterfactuals, investigate the exchange rate movements with and without intervention in otherwise identical circumstances (as far as can be determined by observable market characteristics that lead up to the decision by the central bank to intervene). By using similar economic circumstances that lead to intervention (similar probabilities of intervention) for “matching up” observations that differ only in whether intervention occurs or not, we are able to address the missing observations and the sample selection bias issues.

Our sample of official daily Japanese intervention in the JPY/USD exchange rate market over the January 1999–March 2004 period constitute a fascinating and unprecedented period in the history of foreign exchange market intervention and fits our methodological framework perfectly.¹ First, the magnitude of intervention was extremely large. Japanese foreign exchange market intervention jumped in 2003, shown in Fig. 1, with the official selling of JPY 20.2 trillion (USD 177 billion) in exchange for USD. Massive intervention operations in support of the USD continued in the first quarter of 2004, during which time the authorities sold another JPY 14.8 trillion (USD 139 billion). Although Japan has been the most active amongst the larger industrial economies in its foreign exchange market operations during the past decade and more, the recent magnitude dwarfs the previous experience. Second, there are distinct periods of intervention frequency during this sample period. Fatum and Hutchison (2005) and several others observe that a sharp departure from past Japanese intervention policy began in early 2003 when the frequency of interventions jumped dramatically. Official intervention continued in the first quarter of 2004 and, in fact, this quarter stands out with an intervention frequency of 73% of business days. Moreover, Fatum and Hutchison (2005) demonstrate that

¹ See Humpage (2003) and Sarno and Taylor (2001) for recent surveys of the intervention literature. See Neely (2005) for a critical assessment and a very useful overview of recent studies of intervention.

intervention operations in Japan during this time were automatically sterilized and had no independent effect on monetary growth (i.e. over and above what would otherwise have been the case in the absence of intervention operations). Consistent with the studies by Ito (2003, 2005) and Kearns and Rigobon (2005), we identify (three) sub-samples of separate intervention regimes according to, in our case, highly noticeable changes in the Japanese intervention frequency. Formal tests of reaction function parameter instability across the sub-samples confirm the existence of three separate intervention regimes.²

The basic methodology consists of two parts. In the first part, models of the decision to intervene (decisions by the Ministry of Finance, carried out by the Bank of Japan as its agent through the Foreign Exchange Fund Special Account) are estimated separately across the full and across the three sub-samples. From the model estimates, the probability of intervention (a propensity score) for each day in the given sample is derived. The sample is then split into a group of days when intervention occurs and a group of days when no intervention occurs. Regardless of whether or not intervention occurs on a given day, there is a uniquely defined intervention probability associated with each day in both groups as well as a realized (day-to-day) change in the JPY/USD exchange rate. In the second part, a matching algorithm – the so-called “nearest neighbor” algorithm where each intervention observation is matched with the no-intervention observation that has the “nearest” propensity score – is implemented and the ATE of intervention on exchange rates is examined using difference-in-means tests.

Focusing on all intervention days and the general issue of effectiveness, the results of the ATE-matching analysis show that the effect of official intervention in Japan varies dramatically across the three sub-samples under study: significant effect (in the “right” direction) during the period of infrequent interventions, no significant effect during the period of relatively frequent interventions, and either an insignificant or perverse (“counterproductive”) effect during the period of very frequent interventions. Furthermore, we find a systematic pattern of non-uniform intervention effects across specific types of intervention days, indicating structural parameter instability within different intervention regimes. These findings are consistent with the view that infrequent intervention operations may surprise markets and prove an effective policy strategy, while frequent intervention operations – even very large scale – are incorporated into market expectations with little or even counterproductive effects.

While our ATE estimations address the fundamental issue of sample selection, our treatment methodology does not solve the endogeneity problem inherent in all intervention studies at the daily frequency. In particular, endogeneity is not less of a concern in a binary treatment framework such as ours than it is in a traditional intervention study aimed at estimating the quantity response of exchange rates to intervention. Therefore, we also carry out an instrumental variable estimation of the ATE in order to ensure that our results are not severely affected by simultaneity bias. The result of the instrumental variable estimation, discussed in detail in the robustness section, suggests that simultaneity is not severely biasing our results.

The rest of the paper is organized as follows. Section 2 describes the official Japanese intervention data. Section 3 further discusses the matching methodology and its application to the study of intervention. This section also describes the reaction function estimations necessary for extracting the propensity scores used in the matching. Section 4 presents the main results. Section 5 considers several robustness tests, including radius matching and a procedure to deal with serial dependence. Section 6 discusses and concludes.

2. Data and summary statistics

The official Japanese intervention data is provided by the Ministry of Finance and consists of official, daily operations in the JPY/USD foreign exchange market. During the period under study, 1 January

² Despite the evident departure from past intervention policies, there was no official announcement of a policy change in January 2003 or in January 2004. Furthermore, there was no official announcement made when the active BoJ intervention policy ended abruptly on 16 March 2004 and no intervention took place in the remainder of 2004.

Table 1
Official Japanese intervention, 1 January 1999–31 March 2004.

Purchases of USD (million USD)	Number of days	Cumulated amount
Full sample: 1 January 1999–31 March 2004		
>1000	113	443,796
>500	21	16,613
>250	5	1694
>0	20	2148
Total	159	464,251
Sample 1: 1 January 1999–31 December 2002		
>1000	28	147,629
>500	2	1799
>250	0	1694
>0	0	2148
Total	30	149,428
Sample 2: 1 January 2003–31 December 2003		
>1000	52	165,101
>500	11	8864
>250	4	1465
>0	15	1671
Total	82	177,101
Sample 3: 1 January 2004–31 March 2004		
>1000	33	131,066
>500	8	5950
>250	1	229
>0	5	477
Total	47	137,722

Notes.

(a) Daily Bank of Japan intervention data obtained from the Japanese Ministry of Finance data bank.

(b) Daily intervention operations of USD 1000 million or greater: >1000; daily intervention operations of USD 500 million or greater, but less than USD 1000 million: >500; daily intervention operations of USD 250 million or greater, but less than USD 500 million: >250; daily intervention operations of less than USD 250 million: >0.

1999–31 March 2004, all official interventions in the JPY/USD market are sales of JPY against purchases of USD.³

Table 1 shows that during the full sample period, Japan intervenes in the JPY/USD exchange rate market on a total of 159 days. On most intervention days the magnitude of intervention is substantial, with purchases of over USD 1000 million and larger dominating (46 intervention days of less than USD 1000 million are reported and only 20 intervention days consist of less than USD 250 million). Table 1 also shows that only 30 of the intervention days occur during the first four years of our sample (between January 1999 and December 2002), 82 intervention days occur during 2003, while a remarkable 47 intervention days occur during the last three months of our sample (between January 2004 and March 2004). The described variation in intervention frequencies across the three sub-samples suggests that the January 1999–March 2004 time period encompasses not one but three different intervention regimes.

We follow Ito (2003) and others in using New York close quotes of the daily JPY/USD exchange rate. The exchange rate data are obtained from Global Financial Data (GFD).

Table 2 presents some basic summary statistics of the exchange rate and intervention data. The number of intervention days associated with three categories of exchange rate changes (small, medium and large) is displayed for the full and for each sub-sample period. The number of intervention days falling into a given category, as a percentage of total intervention days, is shown in parentheses below the absolute number of intervention days. The number of days as a percentage of total days in the

³ The US did not intervene in the JPY/USD exchange rate market during this period.

Table 2

Summary statistics on exchange rate changes and intervention.

	Full sample: January 1999–March 2004	Sample 1: January 1999–December 2002	Sample 2: January 2003–December 2003	Sample 3: January 2004–March 2004
Change in JPY/USD (pct.)	-6.8574	6.1600	-10.6271	-2.6628
Average absolute daily pct. change in JPY/USD	0.4832	0.5124	0.3888	0.3926
Standard deviation of daily pct. change in JPY/USD	0.4317	0.4539	0.3244	0.3772
Intervention on days of SMALL JPY/USD changes	74 (47%) [37%]	6 (20%) [35%]	45 (55%) [46%]	23 (49%) [48%]
Intervention on days of MEDIUM JPY/USD changes	52 (33%) [39%]	10 (33%) [40%]	10 (30%) [37%]	17 (36%) [36%]
Intervention on days of LARGE JPY/USD changes	33 (21%) [23%]	14 (47%) [26%]	12 (15%) [17%]	7 (15%) [16%]
REPORTED intervention	100	27	49	24
Total intervention	159	30	82	47

Notes.

- (a) SMALL JPY/USD changes are defined as daily changes (in pct.) that are smaller than the sum of the average of the daily JPY/USD exchange rate (in pct.) and $(-1/2)$ the standard deviation of the daily JPY/USD exchange rate (in pct.).
- (b) MEDIUM JPY/USD changes are defined as daily changes (in pct.) that fall in between the range of the sum of the average of the daily JPY/USD exchange rate (in pct.) $+/- 1/2$ the standard deviation of the daily JPY/USD exchange rate (in pct.).
- (c) LARGE JPY/USD changes are defined as daily changes (in pct.) that are larger than the average of the daily JPY/USD exchange rate (in pct.) plus $(1/2)$ the standard deviation of the daily JPY/USD exchange rate (in pct.).
- (d) REPORTED intervention includes days of Reuters news reports of intervention coinciding with actual intervention.
- (e) Parentheses () report intervention days as a percent of total intervention days, and brackets [] report number of days as a percent of total days in the sample, that were associated with days of SMALL, MEDIUM or LARGE exchange rate changes.

sample falling into the particular exchange rate change category is given in brackets below the number of intervention days. According to the summary statistics it appears that a shift in exchange rate policy took place. Intervention was mainly concentrated during episodes of (relatively infrequent) large exchange rate changes in the first sub-sample sample, and shifted to a fairly uniform distribution (percentage of intervention days corresponding to percentage of days in the sample) of intervention across all days in the second and third sub-samples.⁴

3. Method of matching and average treatment effects (ATE)

The matching method addresses the issue of non-random sample selection.⁵ The effect of the official Japanese intervention on the JPY/USD exchange rate is assessed by matching observations with similar characteristics in terms of intervention propensities (i.e. the likelihood of intervention on a given day), using that one group of observations consists of days when the government intervened (the “treatment” group) while the other group consists of days when the government did not intervene (the “control” group). In turn, the matching of observations allows us to capture the effect of intervention (the “treatment” effect) by measuring the difference in the average JPY/USD exchange rate change between the two groups.

In order to characterize the similarity among observations with and without official intervention, we consider a set of observable variables that can explain the decision to intervene. These variables, described in detail in the next sub-section, include the standard explanatory variables used when

⁴ Table 2 also displays the number of intervention days coinciding with Reuters news reports of intervention. Some studies question the accuracy of such news reports (see Fischer, 2006 and Osterberg and Humes, 1993) and we do not incorporate this information into our analysis.

⁵ See Glick et al. (2006) for a recent application of the matching method to an analysis of the effects of capital account liberalization on the risk of currency crises. See Persson (2001) for a useful explanation of the matching methodology. A textbook exposition is provided by Wooldridge (2002), and a survey is offered by Imbens (2004).

estimating official intervention reaction functions and a number of “news” variables. Rosenbaum and Rubin (1983, 1985) show in a general context that it is sufficient to match according to the one-dimensional probability of an observation being subject to the “treatment”. Therefore, by using the observable intervention determinants for estimating the probability of an intervention occurrence for each day in our sample, the matching is carried out according to the estimated probability of official intervention (the propensity scores).

The nearest neighbor algorithm matches each intervention observation to the no-intervention observation that has the nearest propensity score. After the no-intervention observation is used, it is “returned” to the pool of no-intervention observations. The effect of intervention, i.e. the “treatment” effect, is computed as a simple average of the differences in outcomes across the paired matches.

3.1. Propensity scores: foreign exchange intervention reaction functions

In order to estimate the Japanese official intervention reaction function and, in turn, extract the propensity scores, we follow Ito and Yabu (2007). Ito and Yabu (2007) build on the friction model developed by Almekinders and Eijffinger (1996) and use daily data for estimating an ordered probit threshold model of Japanese intervention operations over the 1991–2002 time period. They develop their reaction function model from first principles by assuming that the intervening authority has a loss function (in exchange rate deviations from a target) that it seeks to minimize by intervening in the foreign exchange rate market. Furthermore, they assume that the exchange rate is a random walk, and that there are “political costs” associated with intervention. These political costs are independent of the size of intervention and may help explain why intervention tends to be correlated such that intervention on day t is likely to be followed by intervention on day $t + 1$.

As explanatory variables, Ito and Yabu (2007) use three measures of the past exchange rates (the first lag of the JPY/USD rate, the 21 business day moving average of the JPY/USD rate, and the one-year moving average of the JPY/USD rate) as well as the first lag of a $(-1, 0, 1)$ intervention indicator variable that takes on non-zero values on intervention days only.⁶ In addition, they employ (potentially asymmetric) threshold values of intervention in order to capture when the costs of intervening are exceeded by the benefits.

We extend this approach by including Japanese news announcements that could influence the decision to intervene.⁷ We are interested in the “surprise” component of news, defined as the difference between official announcements regarding GDP, CPI, the unemployment rate and the trade balance, and results of surveys of expectations of these announcements conducted by Bloomberg during the days preceding the announcements. The official value of these news variables is announced once a month, or at a lower frequency. Our news variables capture the associated surprise element on announcement dates, thus these variables are non-zero only on announcement dates and only when the announcement differs from market expectations.⁸

Our reaction function estimates are free of simultaneity bias due to the inclusion of lagged exchange rate changes and the exclusion of contemporaneous exchange rate changes as explanatory variables. The cost of avoiding endogeneity in the reaction functions, however, is that we do not account for the possibility of within-day exchange rate movements triggering some interventions. In other words, if the true structural reaction function also depends on contemporaneous information, the estimated

⁶ While Almekinders and Eijffinger (1996) use the intervention amount as their dependent variable, consistent with Baillie and Osterberg (1997), Ito and Yabu (2007) use the indicator variable of intervention as their dependent variable in a binary choice modeling framework. They argue that the decision of the monetary authority to intervene or not is more important than the magnitude of intervention.

⁷ Bonser-Neal and Tanner (1996) and, more recently, Galati et al. (2005) and Fatum and Scholnick (2006) have found news variables to impact day-to-day exchange rate changes. Our reaction function specification allows such news to also influence the decision to intervene. We thank a referee for pointing out that the news variables proxy for contemporaneous movements in the exchange rate.

⁸ Neely (2008) notes that central banks say in survey responses that they do not intervene directly in response to macro news. We therefore also carry out the analysis without news variables in the reaction functions. The results are completely robust to the exclusion of the news variables. These results are not shown for brevity but available from the authors upon request.

parameters will suffer from omitted variable bias instead of endogeneity. As a result, our reaction function estimations give only an imprecise measure of the intervention probabilities.

As mentioned earlier, all the official intervention operations in our sample period are sales of JPY against USD purchases. This allows us to use a standard binary choice model, reflecting that there are only two (intervention or no intervention) as opposed to three (intervention purchases of USD, no intervention, and intervention sales of USD) possible actions taken by the authorities.⁹

We use a logit model framework and estimate the following regression model:

$$IINT_t = \beta_0 + \beta_1 JPYUSD_{t-1} + \beta_2 TARGET_{t-1} + \beta_3 MADAY_{t-1} + \beta_4 MAYEAR_{t-1} + \beta_5 NEWS + \beta_6 IINT_{t-1} + \epsilon_t \quad (1)$$

where $IINT$ is the (0, 1) indicator variable that takes on the value 1 on days when there is intervention and 0 otherwise, $JPYUSD$ is the first-difference of the log of the JPY/USD exchange rate, $TARGET$ is the first-difference of the log of the JPY/USD deviation from an exchange rate target of 125 JPY/USD, $MADAY$ is the 21-day moving average of the log of the JPY/USD exchange rate and $MAYEAR$ is the one-year moving average of the log of the JPY/USD exchange rate.¹⁰ $NEWS$ is the vector of variables capturing the unexpected component of Japanese macroeconomic news on days when an official macroeconomic announcement differs from market expectations. $NEWS$ surprises cover GDP (GDP-UNEXP), CPI (CPI-UNEXP), the unemployment rate (UNEM-UNEXP) and the trade balance (TRDE-UNEXP). $NEWS$ variables are daily observations (on the announcement date) that are divided into positive and negative surprises (giving us a total of eight $NEWS$ announcement variables), in order to take into account the fact that *only* dollar support intervention operations were observed during the five-year sample period.

In order to take into account the possibility of heteroskedasticity in the error term, ϵ_t , all our logit model estimations are carried out using [White's \(1980\)](#) heteroskedasticity-consistent (robust) standard errors. The constant term, β_0 , is included to allow for the possibility of a threshold value consistent with the political costs of intervention.¹¹ The reaction function model is estimated over the full sample period and separately across the three sub-samples. The sub-samples are demarked by the striking changes in intervention frequencies.¹² As noted earlier, despite no official announcement of a policy change in January 2003 when the frequency of intervention jumped to 32% of business days, or in January 2004 when the frequency of intervention jumped even further to 73% of business days, it seems evident that the de-facto intervention policy of Japan has changed twice during the period under study.

We first estimate the model including all the explanatory variables in Eq. (1). A common characteristic of the estimations across all samples is the insignificance of $TARGET$ and six of the $NEWS$ variables.¹³ We then exclude the insignificant explanatory variables and re-estimate the reduced models. These models are reported in [Table 3](#). The significant variables for the full sample (column 1) are the lagged exchange rate change $JPYUSD$, lagged $MAYEAR$, lagged intervention, positive GDP announcement surprises and positive CPI announcement surprises. The significant variables differ across the sub-sample estimations.

The displayed model diagnostics suggest that the four regressions fare reasonably well in explaining intervention operations. The McFadden R^2 ranges from 0.27 to 0.41 in the estimated equations, and they all pass the Likelihood Ratio test against the constant-only alternative. None of the models are

⁹ In other words, analyzing the entire April 1991–March 2004 period of publicly available Japanese intervention data (encompassing both purchases and sales of USD) using the binary treatment framework is not possible.

¹⁰ The variable $TARGET$ is included (and significant) in the reaction function estimations displayed in [Ito \(2003\)](#), but not included in [Ito and Yabu \(2007\)](#). Inclusion of $TARGET$ is possibly problematic due to a high degree of collinearity with $JPYUSD$. As it turns out, $TARGET$ is insignificant in all our estimations and, therefore, subsequently excluded from the analysis.

¹¹ As explained in [Ito and Yabu \(2007\)](#), this “conventional” reaction function specification is a linearization of the general friction model of central bank intervention.

¹² This is similar to [Kearns and Rigobon \(2005\)](#), [Ito \(2003\)](#), and [Ito and Yabu \(2007\)](#) who identify June 1995 as a turning point in the Japanese intervention policy of the previous decade due to the noticeable change in the frequency of intervention.

¹³ These results are not reported for brevity but are available from the authors upon request.

Table 3

Logit model estimates of Japanese intervention function dependent variable: IINT.

	Full sample: January 1999–March 2004	Sample 1: January 1999–December 2002	Sample 2: January 2003–December 2003	Sample 3: January 2004–March 2004
Constant	−3.60*** (0.18)	−4.69*** (0.31)	−2.05*** (0.24)	−1.10* (0.58)
JPYUSD(−1)	−95.30*** (16.99)	−116.75*** (27.59)	−57.24* (31.69)	n.a.
MADAY(−1)	n.a.	−32.92*** (5.05)	−18.60** (9.10)	n.a.
MAYEAR(−1)	−4.60*** (1.11)	n.a.	n.a.	n.a.
IINT(−1)	4.05*** (0.26)	2.84*** (0.71)	2.60*** (0.34)	3.23*** (0.75)
GDP−UNEXP > 0	1.50** (0.63)	n.a.	n.a.	40.18*** (1.11)
CPI−UNEXP > 0	1.70** (0.80)	2.02*** (0.77)	n.a.	n.a.
Total Obs	1364	1040	260	64
Obs with IINT = 1	159	30	82	47
McFadden R ²	0.41	0.27	0.28	0.33
LR statistic	406.14	73.45	90.00	24.24
P(LR)	0.00***	0.00***	0.00***	0.00***
H−L statistic	11.00	3.54	9.05	8.31
Total gain	77.04	79.02	56.82	59.38
Percent gain	43.18	15.38	43.18	40.18

Notes.

(a) * denotes significance at 90%; ** denotes significance at 95%; *** denotes significance at 99%.

(b) Heteroskedasticity and autocorrelation consistent standard errors in parentheses to the right of the point estimates.

(c) Logit models are defined in Eq. (1) in the text.

(d) The dependent variable IINT is a (0, 1) indicator variable that takes on the value 1 on intervention days and 0 otherwise.

(e) The independent variables are defined as follows: JPYUSD is the first-difference of the log of the JPY/USD exchange rate; TARGET is the first-difference of the log of the JPY/USD deviation from a target rate of 125 JPY/USD; MADAY is the 21-day moving average of the log of the JPY/USD exchange rate; MAYEAR is the one-year moving average of the log of the JPY/USD exchange rate; (−1) denotes the first lag of a variable. The control variables measure the surprise element of Japanese macroeconomic announcements concerning GDP (GDP−UNEXP), CPI (CPI−UNEXP), Unemployment (UNEM−UNEXP), and Balance of Trade (TB−UNEXP). Positive (>) and negative (<) surprises are controlled for separately.

(f) n.a. indicates that an independent variable is omitted from the second model due to lack of significance.

(g) The Likelihood Ratio (LR) statistic tests the overall significance of the estimated model against a constant-only alternative. P(LR) shows the *p*-value of the LR test statistic.(h) Hosmer–Lemeshow test statistic for goodness-of-fit based on a $\chi^2(8)$ -distribution. The 95 (90) percent critical value for rejecting the null of a fitting model is 15.51 (13.36).

(i) The estimated logit models' prediction evaluation is based on expected value calculations: Total Gain captures the percentage point gain/loss of correct predictions when compared to the naïve constant probability model; Percent Gain shows the percent of incorrect predictions according to the naïve model corrected by the estimated model.

rejected by the Hosmer–Lemeshow test and, lastly, all four models improve the predictive ability against the naïve constant probability model. These results compare favorably with other models explaining official intervention policy in Japan.¹⁴

The model coefficient estimates reported in Table 3 indicate structural breaks across the sub-samples, in line with the observed shifts in the Japanese intervention intensity. We use a standard Wald test (*F*-test) to assess whether the coefficient estimates of the reduced models are significantly different across the three sub-samples. Table 4 displays the results. Overall, the Wald test shows that the estimated reaction function parameters are generally unstable, indicating the existence of three separate intervention regimes.

Our estimations suggest that the intervening authority reacts systematically to previous day exchange rate changes in deciding to intervene, i.e. a depreciation of the JPY against the USD leads to a lower probability of USD-support intervention the following day. This holds true in every sub-sample except sub-sample 3 (first quarter of 2004). Intervention on a given day is an excellent predictor of intervention the following day in every sub-sample, indicating that intervention operations are frequently clustered. The 21-day moving average of the exchange rate (MADAY) is also a good predictor of intervention in the first and the second sub-sample (1999–2002 and 2003), indicating that trend

¹⁴ Although not directly comparable with our results, it is noteworthy that R-Bar squares for Ito (2003) intervention functions, estimated by GMM, was 0.345 over April 1991–June 1995, but only 0.025 over June 1995–March 2001 (the period closest to our sample), and 0.026 for his full sample (April 1991–March 2001).

Table 4

Wald tests of reaction function parameter stability.

	Sub-sample 1	Sub-sample 2	F-statistic
Constant	−4.85	−2.05	71.04***
JPYUSD(−1)	−116.85	−57.24	4.65**
MADAY	−35.66	−18.60	8.04***
IINT(−1)	2.84	2.60	0.11
GDP−UNEXP > 0	n.a.	n.a.	n.a.
CPI−UNEXP > 0	2.02	n.a.	n.a.
	Sub-sample 1	Sub-sample 3	F-Statistic
Constant	−4.85	−1.10	131.33***
JPYUSD(−1)	−116.85	n.a.	n.a.
MADAY	−35.66	n.a.	n.a.
IINT(−1)	2.84	3.23	0.30
GDP−UNEXP > 0	n.a.	40.18	n.a.
CPI−UNEXP > 0	2.02	n.a.	n.a.
	Sub-sample 2	Sub-sample 3	F-statistic
Constant	−2.05	−1.10	15.85***
JPYUSD(−1)	−57.24	n.a.	n.a.
MADAY	−18.60	n.a.	n.a.
IINT(−1)	2.60	3.23	5.76**
GDP−UNEXP > 0	n.a.	40.18	n.a.
CPI−UNEXP > 0	n.a.	n.a.	n.a.

Notes.

(a) * denotes significance at 90%; ** denotes significance at 95%; *** denotes significance at 99%.

(b) The null hypothesis of the Wald test is that a coefficient estimate is the same across two sub-samples (parameter stability); the alternative hypothesis is that they are different (parameter instability).

(c) The dependent variables are defined in the notes to Table 2. The displayed coefficient estimates are associated with the models described in Table 2 (columns 2, 3 and 4). Sub-sample 1 covers January 1, 1999–December 31, 2002; sub-sample 2 covers January 1, 2003–December 31, 2003; and sub-sample 3 covers January 1, 2004–March 31, 2004.

depreciation of the JPY reduces the likelihood of intervention. Over the entire sample (1999–2004) the one-year moving average exchange rate change (MAYEAR) is a significant predictor of intervention, i.e. long-term trend depreciation of the JPY leads to a lower likelihood of USD-support intervention (even when controlling for contemporaneous exchange rate changes). Unexpected increases in Japanese GDP growth (growth announcements less expected growth based on survey data) increases the likelihood of intervention to support the USD in the full sample and in the third sub-sample (columns 1 and 4). Unexpectedly high CPI announcements also lead to USD-support intervention in the full sample and the first sub-sample.

4. Results

As a benchmark for comparison, we test for significant effects of intervention without invoking the matching technique. In order to do so, we use a standard *t*-test for assessing whether the average exchange rate change across intervention days is significantly different from zero. We do so across the full sample as well as separately across the three easily identified sub-samples.

Table 5 shows that intervention is, on average, effective over the full January 1999–March 2004 period. Table 5 also shows that the effects of intervention vary across the three samples and that the full-sample effectiveness result is driven by the intervention that occurred during the January 1999–December 2002 sub-period. Once again, the idea of uniform effects of intervention across different regimes, therefore, seems questionable and we focus our matching analysis on the sub-samples separately rather than on the full sample.

4.1. Nearest-neighbor matching

Using the cumulative logistic distribution function it is straightforward to extract from the estimated intervention reaction functions the conditional probability of observing an intervention on any

Table 5

Without matching: effectiveness of bank of Japan intervention.

	Full sample: January 1999–March 2004	Sample 1: January 1999–December 2002	Sample 2: January 2003–December 2003	Sample 3: January 2004–March 2004
ALL	0.00107** (0.00524)	0.0065*** (0.0021)	−0.0003 (0.0005)	0.0001 (0.0008)
Intervention days	159	30	82	47

Notes.

(a) The table shows the mean of the JPY/USD exchange rate changes on intervention days; significance of whether the mean is different from zero is assessed using a standard *t*-test.

(b) * denotes significance at 90%; ** denotes significance at 95%; *** denotes significance at 99%.

(c) Standard errors in parentheses below the point estimates.

(d) ALL includes all intervention days; IINT is a (0, 1) indicator variable that takes on the value 1 on days when the Bank of Japan intervenes and 0 otherwise.

given day in our sample.¹⁵ These probabilities constitute the propensity scores necessary for the paired matching of observations. Based on these propensity scores, we employ the nearest-neighbor matching algorithm to evaluate the effect of intervention (computed as a simple weighted average of the differences in the outcomes across the paired matches). We first focus on all intervention days (ALL) and address the general issue of effectiveness, paying particular attention to the results from analyzing separately the three sub-samples. We then compare these effectiveness results to the previously discussed preliminary results obtained from testing for significant effects of intervention on the average exchange rate changes across the intervention days without using matching.

4.2. Nearest-neighbor matching results: all intervention days (ALL)

The result of the matching analysis of all intervention days (ALL) across the full sample is displayed in Table 6. The propensity scores underlying the matching procedure are derived from the model reported in column 1 of Table 3. Without taking into account the possibility of different intervention effects associated with different time periods, this result indicates that intervention has a small impact on the JPY/USD rate once we correct for the selection bias. The 0.0014 estimate suggests that, on average, days of USD-support intervention over 1999–2004 led to a same-day 0.14% appreciation of the USD against the JPY. This finding is significant at the 90% level of confidence. This estimated effect of intervention is similar to the 0.11% appreciation estimated without matching, reported in Table 5.

Table 6 also shows the estimated effects of intervention for the three sub-samples. The underlying propensity scores are derived from the models displayed in columns 2, 3 and 4, respectively, of Table 3. The row labeled ALL again displays the results of the effectiveness analysis when all of the intervention days for the sub-sample in question are considered.

Focusing first on the January 1, 1999–December 31, 2002 sub-sample, Table 6 shows that Japanese intervention (USD purchase against JPY sale) is associated with an exchange rate movement of the correct sign (USD appreciation vis-à-vis the JPY). The point estimate suggests that an intervention day is, on average, associated with a 0.61% same-day increase in the JPY/USD exchange rate. This finding is significant at the 95% level.

Turning to the January 1, 2003–December 31, 2003 sub-sample, the displayed point estimate of 0.02% is of the correct sign, but insignificant. This suggests that during 2003, intervention had, on average, no effect on the JPY/USD exchange rate.

Finally, the effectiveness result based on the January 1, 2004–March 31, 2004 sub-sample suggests that intervention is associated with a JPY/USD movement of the wrong sign. Specifically, an intervention (USD purchase) is, on average, associated with a 0.13% decrease in the JPY/USD exchange rate (USD depreciation). This finding is significant at the 90% level and illustrates that intervention may not only be ineffective (as is the case of the 2003 sub-sample), but even counterproductive.

¹⁵ See Humpage (1999) for an earlier application of logit models to the study of intervention.

Table 6

Nearest-neighbor matching: effectiveness of intervention.

	Full sample: January 1999– March 2004	Sample 1: January 1999– December 2002	Sample 2: January 2003– December 2003	Sample 3: January 2004– March 2004
ALL	0.0014* (0.0008)	0.0061** (0.0021)	0.0002 (0.0008)	–0.0013* (0.0008)
SA	n.a.	0.0080** (0.0032)	0.0009 (0.0028)	n.a.
CL	n.a.	0.0027 (0.0021)	0.0001 (0.0008)	–0.0013* (0.0008)
CLFD	n.a.	0.0043 (0.0035)	–0.0000 (0.0019)	–0.0031** (0.0015)
CLMD	n.a.	n.a.	–0.0002 (0.0010)	–0.0010 (0.0009)
CLLD	n.a.	–0.0003 (0.0025)	0.0010 (0.0015)	–0.0022 (0.0036)
FD	n.a.	0.0072*** (0.0026)	0.0003 (0.0029)	–0.0031** (0.0015)
Obs with IINT = 1	159	30	82	47

Notes.

(a) Matching based on the nearest neighbor algorithm.

(b) * denotes significance at 90%; ** denotes significance at 95%; *** denotes significance at 99%.

(c) Standard errors in parentheses to the right of the point estimates.

(d) ALL includes all intervention days; SA includes only “stand-alone” intervention days, i.e. intervention days not immediately preceded or succeeded by another intervention day; CL includes all intervention days belonging to clusters, i.e. intervention days immediately preceded or/and succeeded by an intervention day (ALL minus SA); CLFD includes CL intervention days that are not immediately preceded by an intervention day; CLMD includes only CL intervention days that are both immediately preceded and succeeded by other intervention days; CLLD includes only CL intervention days that are not immediately succeeded by another intervention day; FD includes only SA and CLFD intervention days; IINT is a (0, 1) indicator variable that takes on the value 1 on days when the Bank of Japan intervenes and 0 otherwise.

(e) For the full sample, only the analysis of ALL is carried out. For Samples 1 through 3, n.a. indicates when a sample does not include the given type of intervention days.

In order to assess the importance of carrying out the matching procedure and the importance of addressing the sample selection issue in the context of an intervention study, we compare the discussed ALL sub-sample results based on matching (displayed in Table 6) to the previously mentioned ALL sub-sample results based on no matching (displayed in Table 5).

For the first sub-sample, both approaches yield very similar results. Although the estimated effect of intervention based on no matching is slightly higher (and associated with significance at the 99% level) a standard *t*-test rejects that the two point estimates are significantly different from each other. This indicates that Japanese intervention works as intended during this period. For the second sub-sample, the sign of the point estimate changes when the matching methodology is employed. However, regardless of whether we use matching or not, we find no significant effects of intervention during the 2003 time period. Assessing the effects of intervention without the use of matching in the third sub-sample, however, yields a correctly signed though insignificant point estimate, while employing the matching procedure yields an unexpected directional sign and significant point estimate.¹⁶

5. Robustness

In order to test the robustness of the results, we carry out the analysis using a different matching algorithm, address the issue of independence of observations, assess whether simultaneity presents

¹⁶ The rows labeled SA and CL in Table 6 display the results of the matching procedure when the effects of intervention across SA (stand-alone or single-day interventions) and CL (clusters of two or more intervention days in succession) are analyzed separately. We also distinguish between the first intervention day in a cluster (CLFD), the intervention day(s) surrounded by other intervention days (CLMD), and the last intervention day in a cluster (CLLD). Overall, we find that the previously discussed significant results found across the 1999–2002 period can be ascribed to the impact of single-day intervention operations. From the perspective of the non-governmental exchange rate market participant, on a given intervention day that does not succeed a previous intervention day there is observational equivalence between a single-day intervention and the first intervention day in a cluster. Therefore, we also pool the single-day interventions with the first intervention day of clusters, denote these as general “first days” (FD), and redo the matching exercise in order to assess the effect across these two types of intervention days. Consistent with the previous findings, FD is significant and of the correct sign in the first sub-sample, while it is insignificant in the second sub-sample. There are no SA intervention days in the third sample, thus the effect of FD mirrors the effect of the first intervention day of clusters (CLFD).

a severe bias in our ATE estimates, and control for the possibility that macroeconomic news affect the exchange rate. All results pertaining to this section are available from the authors upon request.

First, we implement the radius algorithm. The radius algorithm matches each intervention observation to the average of all the no-intervention observations with propensity scores falling within a pre-set radius from the propensity score of the intervention observation.¹⁷ The effect of intervention is again computed as an average of the difference in outcomes, weighted by the number of no-intervention observations used in the construction of each match. The radius matching averages over the control sample so the results are not only dependent upon individual “matches” but, rather, rely on a control group of matches with similar propensity scores. Unlike nearest-neighbor matching, which takes the closest match regardless of the propensity score distance, radius matching will drop intervention observations from the sample that do not find support for the specified propensity score radius within the set of control observations. Radius matching, therefore, effectively trims the intervention sample for outliers among the intervention (treated) observations. The results across the first two sub-samples are completely robust to this change in matching procedure, and the results are also very similar regardless of the radius distance. The third sub-sample of intense intervention operations, however, does not lend itself to radius matching since so few days are available in the control sample, i.e. most days of the sample are associated with intervention.¹⁸

Second, we address the fact that our observations are not independent draws and that this leads to a possible correlation between the treatment and the control sample observations.¹⁹ We adjust for this concern by excluding from the control sample (consisting of days of no intervention) all observations immediately following an intervention. By this adjustment we ensure that the control sample, and potential “matches” that are calculated in the ATE, do not include observations that are potentially correlated with the previous day’s intervention operation. For completeness, we also carry out the matching analysis on sub-samples where we drop both the first and the second no-intervention days following an intervention day. Our results based on these trimmed samples are qualitatively identical to our baseline findings.²⁰

Third, we assess whether simultaneity presents a severe bias in our ATE estimates based on propensity-score matching. In order to do so, we follow Wooldridge (2002, p. 623: Procedure 18.1) and estimate the ATE of intervention on the exchange rate for the full sample using an instrumental variable procedure.²¹ We then compare our regression-based instrumental variable ATE estimate to the propensity-score matching ATE estimate. The estimated ATE effect of intervention using the instrumental variable procedure is 0.0013 and significant at the 90% level of confidence (these results are not shown for brevity but available from the authors upon request). The ATE effect from the propensity-score

¹⁷ The selection of the radius distance is somewhat arbitrary, however, and the choice is usually determined judgmentally by the range and distribution of propensity scores over the [0,1] spectrum. In order to be comprehensive in our robustness tests we choose three alternative values for the radius: 0.01, 0.05 and 0.10.

¹⁸ Nearest-neighbor matching is also problematic in the third sub-sample since the nearest-neighbor control observation is the same for many of the intervention observations. In general, any sample with a large number of intervention observations and few control observations will not be well suited to matching methods.

¹⁹ The matching methodology assumes that each observation is independent and drawn from an identical distribution from the underlying population. This assumption implies that the effect of intervention on the exchange rate on the intervention day does not affect the exchange rate on any other day. In our context, however, it is possible that this assumption is violated if, for example, an intervention operation on day t affects the exchange rate movement on day $t + 1$. If this is the case, then the $t + 1$ observation is “contaminated” by the treatment (i.e. the intervention) of observation t .

²⁰ This robustness test was not implemented for the third sub-sample. Dropping the no-intervention days following intervention days during the third sub-sample of very frequent intervention did not leave enough control observations to make this test meaningful.

²¹ The procedure is related to the dummy endogenous variable model of Heckman (1978) and involves two steps. The first step is to estimate a binary response model for the selection equation (in our case the likelihood of foreign exchange market intervention estimated by a logit model) and obtain the fitted probabilities (the propensity scores). The second step is to estimate the outcome equation (in our case the change in the exchange rate) by instrumental variables where the explanatory variables are intervention (substituted with the propensity scores as the instrument) and the other variables in the selection equation except for one that is excluded for identification. In our context, we include lagged intervention in the selection equation (since the first lag of intervention is a good predictor of contemporaneous intervention) and exclude it from the outcome equation (since lagged intervention is a poor predictor of the contemporaneous exchange rate change). We only consider the full sample since, in order to use this procedure, we need enough variation in the independent variables to obtain sufficiently accurate coefficient estimates.

matching (nearest neighbor) over the full sample is 0.0014 and also significant at the 90% level of confidence (see Table 6). This suggests that simultaneity is not severely biasing our results

Fourth, we control for the possibility that macroeconomic news surprises affect the exchange rate by excluding from both the intervention and the no intervention samples days coinciding with news surprises. Unsurprisingly, our results are not affected by this exclusion of data points. Few news surprises coincide with intervention days.

6. Conclusion

The intervening authority makes a conscious decision to enter the foreign exchange market when intervention occurs. The self-selection of the timing of an official intervention operation, and the fact that we don't observe what would have occurred in its absence, is a methodological challenge in estimating the effect on the exchange rate. Estimating an appropriate "counterfactual" under these circumstances in order to properly evaluate the effects of intervention on exchange rate movements is a central methodological problem. We address the issue of self-selection and the missing counterfactual by estimating the "average treatment effect" (ATE) of intervention on the exchange rate. We use a propensity-score matching methodology to do so.

In our analysis of daily official intervention in the JPY/USD exchange rate market over the January 1999–March 2004 period, the exchange rate movement is the "outcome" variable and intervention is the "treatment." Our propensity-score matching compares pairs of observations of exchange rate movements – each pair consisting of an exchange rate movement coinciding with intervention and one that coincides with no intervention – that are similar in observable characteristics (and associated with similar probabilities of intervention). To derive the propensity scores we estimate a central bank intervention reaction function. The ATE is the average difference in terms of exchange rate movements across these matched pairs (or, for radius matching, an average of several control observations that are matched with the intervention observation).

We find significant effects of intervention in the right direction in the January 1999–December 2002 sub-sample. This general finding is consistent with several other studies analyzing Japanese intervention over a similar period.²² By contrast, we find a complete lack of significant effects of official intervention in the 2003 sub-sample. Looking at the first quarter of 2004, and the extended sub-sample that includes the first and second quarters of 2004, the effects of official intervention are once again significant, but this time in the wrong direction. Intervention appears to be counterproductive. All our results are robust to various methodological changes.

It is interesting that intervention is effective during the first sub-sample of infrequent interventions (3% of business days), ineffective during the second sub-sample of more frequent interventions (32% of business days) and possibly counterproductive during the third sub-sample where the interventions occur at an extremely high frequency (73% of business days).²³ Although not a testable hypothesis, given that the three sub-samples of our analysis essentially constitute three "observations", it seems plausible that the dramatic increase in the Japanese intervention frequencies constitute an important element towards understanding why intervention in one direction, in one exchange rate, carried out by one central bank over a total time span of little more than 5 years, turns from effective to ineffective and, perhaps, counterproductive.

This "frequency explanation" is consistent with our finding that the significant results of the 1999–2002 period stem from single-day intervention operations, and points to effectiveness being conditional on the surprise element of intervention.²⁴

²² Fatum and Hutchison (2006) find significant support for effectiveness over the April 1991–December 2000 interval, Kearns and Rigobon (2005) find effectiveness over the April 1991–June 2002 time period, and Ito (2003) finds effectiveness over the June 1995–March 2001 period. By contrast, Galati et al. (2005) find little or no effect of Japanese intervention in their study of perceived intervention.

²³ This pattern is also found in Ito (2003). He shows that Japanese intervention during the 1991–1995 period, characterized by relatively frequent interventions (16% of business days), is ineffective or even counterproductive while intervention during the 1995–2001 period, characterized by infrequent interventions (3% of business days), is effective.

²⁴ This result is also in line with work by Chaboud and Humpage (2005) and Fatum (2002).

We have some methodological reservations about the ATE-matching results from the 2004 sub-sample when intervention operations were intense and very frequent. Therefore, we interpret the results with caution and conclude as follows: Our results strongly support effectiveness of official Japanese intervention during an extended period of relatively infrequent operations (1999–2002), while no evidence to support effectiveness is found during periods of frequent and large-scale intervention operations (2003–2004).

Acknowledgements

Fatum gratefully acknowledges financial support from the Social Sciences of Humanities Research Council of Canada (SSHRC) and from a University of Alberta School of Business Fellowship. Part of this research was carried out while Fatum was Visiting Scholar at the Bank of Japan's Institute for Monetary and Economic Studies (IMES). We thank seminar and conference participants at the University of Copenhagen, the Solvay Business School, the Hong Kong Institute for Monetary Research, the NBER Japan Project, the Institute for Economic Policy Research (IEPE/CdG, Rio de Janeiro), and Kathryn Dominguez, Ippei Fujiwara, Mark Huson, Fabio Kanczuk, Franz Palm, Barry Scholnick, Nao Sudou and Carl Walsh for very helpful comments. We are also very appreciative of helpful referee comments. We thank Xueyan Guo and Michael Simmons for valuable research assistance.

References

- Almekinders, G.J., Eijffinger, S.C.W., 1996. A friction model of daily bundesbank and federal reserve intervention. *Journal of Banking and Finance* 20, 1365–1380.
- Baillie, R.T., Osterberg, W.P., 1997. Why do central banks intervene? *Journal of International Money and Finance* 16, 909–919.
- Bonser-Neal, C., Tanner, G., 1996. Central bank intervention and the volatility of foreign exchange rates: evidence from the options market. *Journal of International Money and Finance* 15, 853–878.
- Chaboud, A.P., Humpage, O.F., 2005. An Assessment of the Impact of Japanese Foreign Exchange Intervention: 1991–2004. Board of Governors of the Federal Reserve International Finance. Discussion Paper no. 824.
- Fatum, R., 2002. Post-plaza intervention in the DEM/USD exchange rate. *Canadian Journal of Economics* 35, 556–567.
- Fatum, R., Hutchison, M.M., 2005. Foreign exchange intervention and monetary policy in Japan, 2003–04. *International Economics and Economic Policy* 2, 241–260.
- Fatum, R., Hutchison, M.M., 2006. Effectiveness of official daily foreign exchange market intervention operations in Japan. *Journal of International Money and Finance* 25, 199–219.
- Fatum, R., Scholnick, B., 2006. Do exchange rates respond to day-to-day changes in monetary policy expectations when no monetary policy changes occur? *Journal of Money, Credit and Banking* 38, 1641–1657.
- Fischer, A.M., 2006. On the inadequacy of newswire reports for empirical research on foreign exchange interventions. *Journal of International Money and Finance* 25, 1226–1240.
- Galati, G., Melick, W., Micu, M., 2005. Foreign exchange market intervention and expectations: an empirical study of the dollar/yen exchange rate. *Journal of International Money and Finance* 24, 982–1011.
- Glick, R., Guo, X., Hutchison, M.M., 2006. Currency crises, capital account liberalization, and selection bias. *Review of Economics and Statistics* 88, 698–714.
- Heckman, J.J., 1978. Dummy endogenous variables in a simultaneous equations system. *Econometrica* 46, 931–959.
- Humpage, O.F., 1999. US intervention: assessing the probability of success. *Journal of Money, Credit and Banking* 31, 731–747.
- Humpage, O.F., 2003. Government Intervention in the Foreign Exchange Market. Federal Reserve Bank of Cleveland. Working Paper no. 03–15.
- Imbens, G.W., 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics* 86, 4–29.
- Ito, T., 2003. Is foreign exchange intervention effective? The Japanese experience in the 1990s. In: Mitzen, P. (Ed.), *Monetary History, Exchange Rates and Financial Markets, Essays in Honour of Charles Goodhart*, vol. 2. Edward Elgar, UK, and NBER Working Paper no. 8914.
- Ito, T., 2005. Interventions and Japanese economic recovery. *International Economics and Economic Policy* 2, 219–239.
- Ito, T., Yabu, T., 2007. What promotes Japan to intervene in the forex market? A new approach to a reaction function. *Journal of International Money and Finance* 26, 193–212.
- Kearns, J., Rigobon, R., 2005. Identifying the efficacy of central bank interventions: evidence from Australia and Japan. *Journal of International Economics* 66, 31–48.
- Neely, C.J., 2005. An analysis of recent studies of the effect of foreign exchange intervention. *Federal Reserve Bank of St. Louis Review* 87, 685–717.
- Neely, C.J., 2008. Central bank authorities' beliefs about foreign exchange intervention. *Journal of International Money and Finance* 27, 1–25.
- Osterberg, W.P., Humes, R.W., 1993. The inaccuracy of newspaper reports of US foreign exchange intervention. *Federal Reserve Bank of Cleveland Economic Review* 29, 25–33.
- Persson, T., 2001. Currency unions and trade: how large is the treatment effect? *Economic Policy* 33, 433–462.
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55.

- Rosenbaum, P.R., Rubin, D.B., 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician* 39, 33–38.
- Sarno, L., Taylor, M.P., 2001. Official intervention in the foreign exchange markets: is it effective and, if so, how does it work? *Journal of Economic Literature* 39, 839–868.
- White, H., 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica* 48, 817–838.
- Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.