14TH JACQUES POLAK ANNUAL RESEARCH CONFERENCE
NOVEMBER 7–8, 2013

Pick Your Poison: The Choices and Consequences of Policy Responses to Crises

Kristin J. Forbes
MIT- Sloan School of Management and NBER

Michael W. Klein
Tufts-Fletcher School and NBER

Paper presented at the 14th Jacques Polak Annual Research Conference
Hosted by the International Monetary Fund
Washington, DC—November 7–8, 2013

Pick Your Poison:

The Choices and Consequences of Policy Responses to Crises*

Kristin J. Forbes –MIT-Sloan School of Management and NBER

Michael W. Klein—Tufts-Fletcher School and NBER

October 16, 2013

CONFERENCE DRAFT—COMMENTS APPRECIATED

Abstract: Countries choose different strategies when responding to crises. It is challenging to assess the impact of these policy choices on key variables, however, because of endogeneity and selection bias. This paper addresses these challenges by using propensity-score matching to estimate how major reserve sales, large currency depreciations, substantial changes in policy interest rates, and increased controls on capital outflows affect real GDP growth, unemployment, and inflation during two periods marked by crises, 1997 to 2001 and 2007 to 2011. We find that sharp currency depreciations and major reserve sales significantly raise GDP growth (albeit with a lagged effect and after an initial contraction) and also increase inflation (especially depreciations). These policies have weaker benefits and greater costs in emerging and non-OECD economies (especially reserve sales). Estimates also show that increasing interest rates and new controls on capital outflows significantly lower GDP growth.

JEL Code: F41

Keywords: Crisis response, depreciation, capital controls, interest rate policy, foreign exchange reserves, propensity score matching.

*****Thanks to Ayhan Kose, Pierre-Olivier Gourinchas, and anonymous reviewers at the IMF for helpful comments. Patrick O'Halloran provided excellent research assistance. Author contact information: kjforbes@mit.edu and michael.klein@tufts.edu .

I. Introduction

In the mid- and late-1990's a series of emerging markets faced sudden stops of capital inflows and severe financial crises. The standard policy prescription included a sharp increase in interest rates and large currency depreciation.[1] Countries that followed this prescription generally experienced severe recessions and increased unemployment, at least immediately after the crisis. This experience caused many countries to take steps to attempt to avoid such severe policies during the next crisis, such as accumulating international reserves, experimenting with capital controls, and developing currency swap arrangements. When the Global Financial Crisis hit in 2008, countries adopted a range of policy responses and, in many cases, avoided substantial increases in interest rates. But these policies also created challenges. Countries must "pick their poison" and every policy choice involves tradeoffs.

This paper analyzes the causes and consequences of different policy responses to contractions in international capital flows and their corresponding crises. It focuses on two crisis windows; from 1997 through 2001 (which includes the crises in Asia, Russia, Brazil, Argentina and Turkey and the global turmoil after the collapse of LTCM), and from 2007 through 2011 (which includes the crises in Latvia, Iceland, and several euro-zone countries and the turmoil during the Global Financial Crisis (GFC)).[2] This paper concentrates on four policy responses; major reserve sales, large currency depreciations, substantial interest rate increases, and new controls on capital outflows. These are defined as major policy adjustments that do not normally occur during stable times, but are also nimble policies that can quickly be adopted in response to sudden stops in capital flows.[3] We document patterns in the use of these four policy responses during the two crisis periods and show that countries were more likely to have large depreciations during the 2007 to 2011 period (when not controlling for other variables). Then we estimate what factors determine a country's policy choices and test the efficacy of these policies in terms of their impact on three key outcome variables: GDP growth, inflation, and unemployment.

---

[1] See Fischer (2004), Chapter 3.

[2] Although the crises in the 1997-2001 period were often more regional or country-specific than during the later period (except during the fall of 1998 after LTCM's collapse), both windows are periods of volatility in global capital flows punctuated by "sudden stops". Claessens *et al.* (2013) presents a wide-ranging analysis and overview of financial crises, and the chapter by Claessens and Kose (2013) provides a detailed discussion of different types of crises. Gourinchas and Obstfeld (2012) analyze the causes and consequences of crises from 1973 – 2010.

[3] This roughly corresponds to the tradeoffs countries face in the policy trilemma, as discussed in Obstfeld, Shambaugh, and Taylor (2010), Dominguez (2013), and Klein and Shambaugh (2013). As with the policy trilemma, we do not consider fiscal policy responses since these are less nimble and generally take longer to implement. Also, fiscal policy responses are the focus of other papers prepared for this conference, such as Chari and Henry (2013).

Two major challenges complicate the analysis of the efficacy of these different policy responses to crises: selection bias and endogeneity. Countries which chose certain policies (such as large depreciations) tend to share certain characteristics and face different circumstances than other countries that do not opt for these policies. For example, estimates in this paper show that countries with larger current account surpluses, pegged exchange rates, and greater capital account openness are less likely to have large currency depreciations during the two crisis windows. Moreover, governments tend to chose policies in response to changes in variables such as economic growth and inflation, which are variables the policies are intended to influence. For example, estimates in this paper show that countries are more likely to have large currency depreciations after a slowdown in GDP growth. Although previous studies note the challenges arising from selection bias and endogeneity, there is typically a limited ability to address these problems because of the difficulties controlling for all differences across countries using a limited set of observable statistics and finding effective instruments to appropriately identify estimates.

This paper addresses these challenges by using a propensity-score matching methodology, as developed in Rosenbaum and Rubin (1985).[4] This technique is common in some areas of economic analysis (such as labor economics), but used rarely in international macroeconomics.[5] To the best of our knowledge, it has not yet been used to analyze the effect of different policy responses to crises. This methodology allows us to construct a counterfactual for each major policy response in each country. The counterfactual is created using an estimated propensity score that selects relevant observations to compare countries which use a certain policy (a treatment group) with countries that do not (a control group).

In the first stage of this propensity-score methodology, we estimate a logit model of the probability that a country undertakes each of the four policy responses as a function of global variables, domestic vulnerabilities, country characteristics, and other recent policy decisions. The model has a high degree of explanatory power and most coefficients have the expected signs. The estimates confirm that there are significant differences in the macroeconomic characteristics of countries that use these policies (showing the presence of selection bias), and that variables

[4] In Section IV.A, we discuss the methodology in more detail and the differences between the propensity-score matching methodology and the standard multivariate regression analysis.
[5] Three exceptions are: Glick, Guo, and Hutchinson (2006), Das and Bergstrom (2012), and Forbes, Fratzscher, and Straub (2013).

intended to be influenced by the policies can significantly affect the probability of using each policy (showing the presence of endogeneity).

In the next stage of the analysis, we use these estimates of the determinants of a country's policy responses to estimate propensity scores, the probability that each country adopts each major policy response in each quarter. These propensity scores are then used to match countries which adopted each policy to a control group. We focus on five matching algorithms and present a series of tests to select between matching algorithms and ensure that key assumptions to use this methodology are satisfied.[6] This stage also includes a statistical analysis that confirms that selection bias is an important issue to address in this analysis. Then we use the matched sets of countries to estimate the average treatment effect on the treated, the statistic that shows the effect of the four policy responses on key outcome variables while controlling for selection bias.

The estimates suggest that large currency depreciations and major reserve sales generate significant increases in GDP growth relative to the counterfactual, although the positive effects do not occur for at least four quarters (and even longer for depreciations) and may be preceded initially by a contraction in GDP growth. Both policies also generate an increase in inflation, although this effect is larger and only significant after depreciations. Substantial increases in interest rates and new controls on capital outflows appear to have particularly negative effects, as they generate sharp and significant decreases in GDP growth. None of these policy responses generates significant reductions in unemployment, although major reserve sales and large currency depreciations are more likely to reduce unemployment (and these effects are sometimes significant at the 10% level).

A series of robustness tests generally confirms these results. We do find, however, that emerging markets and non-OECD economies may face an even more difficult tradeoff. In this group of countries, the benefits of sharp depreciations and large reserve sales in terms of raising GDP growth are generally weaker, while large reserve sales may increase (instead of decrease) unemployment.

These results suggest that during a crisis and contraction in capital flows, there is no miracle cure of a policy response—especially in emerging and non-OECD economies. Neither a major reserves sale, sharp currency depreciation, large increase in the interest rate, or new

[6] As discussed in more detail in Section IV, these include nearest-neighbor without replacement, five-nearest neighbors with replacement, radius with caliper, kernel, and local-linear matching.

controls on capital outflows will be able to simultaneously generate stronger growth and lower unemployment while avoiding an increase in inflation.

The remainder of this paper is as follows. Section II defines the major policy responses that are the focus of this paper and documents their incidence and joint occurrence over the two crisis periods. Section III estimates a logit model predicting a country's use of each policy in each quarter. Section IV explains the propensity-score matching methodology and uses the results from the logit regression to estimate propensity scores and create control groups using different matching algorithms. Then Section V uses these results to estimate the effect of different policies on GDP growth, unemployment and inflation, including a number of robustness checks and extensions to the base-case analysis. Section VI concludes.

II. Policy Responses During Crises

Countries can, and do, deal with economic crises using a variety of policy tools. In this section we document the policy choices undertaken in our sample of 85 countries for two four-year periods which include crises affecting multiple countries and contractions in global capital flows, from 1997 – 2001 and 2007 – 2011. To simplify terminology, we will refer to the earlier period as the 1990's crises and the later window as the 2000's crises. We focus on four different policy responses: major sales in foreign exchange reserves, sharp depreciations of the nominal exchange rate, substantial increases in policy interest rates, and new or augmented controls on capital outflows.

These are four of the primary policy responses that a government might quickly undertake when faced with different exogenous shocks. For example, as described in Blanchard, Das, and Faruqee (2010), an increase in perceived risk that reduces capital flows, or a decrease in foreign output that leads to a decrease in domestic exports, often lead to currency depreciations. Depreciations can further reduce output through balance sheet effects. Policy efforts to forestall a depreciation could include selling reserves, raising interest rates, or implementing controls on capital outflows. The cost of these policies, however, might be lower output than would otherwise occur with the depreciation. As Blanchard *et al.* write, "If the policy implications seem complicated, it is because they are." (p. 276)

We focus on episodes of relatively large and infrequent changes in these four policy choices for a broad sample of 85 countries.[7] For our base case, we define large policy responses as occurring in only 5% of the country-quarter observations during the two crisis windows.[8] More specifically, using these thresholds, major reserve sales are defined as quarters during which international reserves (excluding gold) fall by at least 24 percent, compared to the previous year.[9] To ensure that this is economically meaningful, we only include cases where the reserve to GDP ratio is at least 10 percent. Large depreciations are defined as quarters in which there is at least a 23 percent depreciation over the previous year in the country's exchange rate versus the U.S. dollar. To avoid episodes in which depreciations primarily reflect a response to inflation, we only include those country-quarter observations in which the 23 percent depreciation was not preceded by a quarter with annual inflation of 20 percent or higher. A substantial increase in interest rates is defined as a quarter in which there has been an increase in the policy interest rate of at least 244 basis points over the past year.[10] To avoid a spurious change due to rising inflation, we only include country-quarters in which the annual rate of inflation is less than 20 percent.

Finally, to define changes in controls on capital outflows, it is necessary to use a different approach because there are no good measures of the intensity of different capital controls so we are unable to use the 5% sample threshold to define a "major" change. Instead, we define adjustments in capital controls as years in which there has been any increase in controls on capital outflows, based on the IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions.*[11] This capital control data is only available on an annual basis, in contrast to the other data which are available on a quarterly basis.[12] Using this measure, there were increases in controls on capital outflows in about 3 percent of the country-year observations during the two crisis periods. As a result, there is a slightly lower incidence of the use of controls on capital outflows, relative to the other three policy responses, in our sample.

[7] The countries are listed in the Appendix and include all "Advanced Economies" (as defined by the International Monetary Fund as of October 2012) and all "Emerging Markets" and "Frontier Economies" (as defined by Standard & Poor's BMI indices). Countries in this list are included if data on at least one of the policy responses is available.
[8] These thresholds translate into roughly 3% of the distribution of country-quarter observations for the years from 1995 through 2011.
[9] Details on sources and definitions for the data are in the Appendix.
[10] The policy interest rate is the interest rate related to monetary policy for each country. If the policy rate is not available, we use the short-term interest rate. Interest rate information is from Global Insight, accessed 10/1/13.
[11] The data on capital controls is from Klein (2012), and we use his approach of including changes in controls on all asset categories except foreign direct investment. This measure of capital controls (as well as most others) only captures when a new control is first added or removed, and not any subsequent modifications to the specific control.
[12] We count a control as being put in place in the third quarter of the year.

In the following analysis, we want to avoid counting a major policy change which persists over several quarters as multiple events. Therefore, we also impose the condition that after a major policy change of the type described above, we do not score another policy change of that type for the next three quarters. As a result, there can be at most one policy change in each of the four categories per year for each country. This does not preclude us from finding major changes in other types of policies at the same time (or within three quarters).

Tables 1a and 1b report the incidence of these major changes in reserves, exchange rates, interest rates, and capital controls for the two crisis periods. The numbers on the diagonal of the tables represent changes in each policy that did not occur within one quarter of a change in any other of the three policies. The numbers in the upper triangular part of the matrix represent pairs of policies that occurred either in the same quarter or within one quarter of each other; for example, element (2,3) in Table 1a shows that there were five quarters in which a substantial increase in interest rates occurred contemporaneously, or within one quarter of, a large depreciation. The numbers in the lower triangular part of the matrix represent triplets of policies that occurred contemporaneously or within one quarter; for example, there were 3 quarters in the 1997 – 2001 period when a country had a major reserve loss, a large depreciation and a substantial increase in interest rates all within one quarter of each other. The numbers in the last column of each table show the absolute number and the percentage of the total number of large policy changes.

One important point that emerges from these tables is that the great majority of the policy changes occur in isolation (represented by the diagonal elements) rather than occurring simultaneously (or within one quarter of) different policies. Only one type of policy response was chosen 80 percent of the time (113 of 143 major policy changes) in the 1997 – 2001 period and 88 percent of the time (84 of 96 changes) in the 2007 – 2011 period. Another important point is the relative incidence of the different policy responses across the two periods. Depreciations occur more frequently during the later crisis window (38% of the total policy responses from 2007 – 2011) relative to during the earlier window (only 31% of the responses from 1997 – 2001). Major reserve sales, increases in interest rates, and new controls on capital outflows, however, are used at similar frequencies during the two crisis windows. All policy responses occur more often when measured on an absolute basis during the earlier crisis window, reflecting the greater number of overall “major policy response” events during this period.

Figures 1a and 1b present the time series of the use of these four policies for the 1997 – 2001 and the 2007 – 2011 periods, respectively.[13] Figure 1a shows that during the earlier crisis window, the instances of major increases in interest rates peaked during 1997-1998. Large reserve sales peaked in 1999, and major depreciations and new controls on capital outflows were more stable. In contrast, Figure 1b shows that in the later crisis window, countries had more major currency depreciations during 2008-2009 relative to any of the other policies. The instances of large increases in interest rates peaked in 2008 before falling to zero in 2009. The use of reserve sales and new controls on capital outflows is more stable in this period and occurs less frequently (as also shown in Table 1b).

III. Explaining Policy Choices

The previous section documented the use of four different policies during crisis periods: selling reserves, currency depreciations, increasing interest rates, and adding capital controls. But what determines which of these policies, if any, a country chooses? To answer this question, this section estimates the probabilities of each of these policies being employed during the crisis periods in the late 1990's and late 2000's. These estimates not only contribute to our understanding of why countries use these policies, but also are the basis for the analysis in the next section that uses propensity scores to estimate how these policies affect key outcome variables such as economic growth.

We estimate the likelihood of a major change in each of the four policies (as defined above) in each quarter using a logit model for our panel data of 85 countries (listed in the Appendix) over the periods 1997Q1 to 2001Q4 and 2007Q1 to 2011Q4. There are a large number of potential covariates that could predict the use of these four policies. We select those which are available for a large sample of countries at (preferably) a quarterly frequency and that research has highlighted as important in predicting vulnerability to crises (e.g., Frankel and Saravelos, 2012, Rose and Spiegel, 2010, Claessens, Dell'Ariccia, Igan, and Laeven, 2010, and IMF, 2010), sudden stops in capital flows (e.g., Forbes and Warnock, 2012, Calvo, 1998, Calvo, Izquierdo, and Meijía, 2008), and the use of capital controls (e.g., Forbes, Fratzscher, and Straub, 2013, and Aizenman and Pasricha, 2013).

[13] Only countries with information for all four policy responses for each year in the given crisis period are included—so that the sample of countries is constant across time within each graph (although not across graphs). As a result, the incidence of the use of each policy differs relative to Tables 1a and 1b.

The covariates can be roughly divided into four categories, recognizing that some variables could be placed in more than one group. One set of covariates controls for changes in the global environment: an indicator of global risk and uncertainty (the VXO), the change in the U.S. interest rate, and the log of a commodity price index.[14] This category also includes a dummy variable equal to 1 for observations in the 1997 – 2001 period in order to distinguish the two crisis periods. All the covariates used for this analysis except the period dummies are lagged by one quarter (or by one year if only annual data is available). A second category includes variables capturing domestic vulnerabilities: changes in real GDP growth, changes in gross capital outflows (relative to GDP), changes in gross capital inflows (relative to GDP), the country's current account balance, changes in CPI inflation, the change in private credit (relative to GDP), and a commodity exporter dummy interacted with the commodity price.[15]

A third category represents domestic characteristics that vary more in the cross section than in the time series for any particular country. This set includes the logarithm of income per capita, an indicator of institutional quality[16], capital-account openness as measured by the Chinn-Ito KAOPEN variable, the level of reserves relative to GDP, whether the exchange rate is pegged[17], and a dummy equal to 1 if the country is a member of the euro zone at any time in the sample. The fourth and final category of covariates captures any changes in the previous quarter in the four policies on which we are focusing: changes in reserves (as a share of GDP) and the country's policy interest rate relative to the U.S. rate, percent changes in the nominal exchange rate versus the U.S. dollar, and any new capital controls (on either inflows or outflows). Details on the definitions and sources for all of these variables are available in the appendix.

Using these four categories, the logit model that is the base case for our analysis is:

$$Prob(pc_{it} = 1) =$$
$$F\left(\boldsymbol{\Phi}_{t-1}^{\text{Global}}\mathbf{B}_{\text{G}} + \boldsymbol{\Phi}_{i,t-1}^{\text{Vulnerabilities}}\mathbf{B}_{\text{V}} + \boldsymbol{\Phi}_{i,t-1}^{\text{Characteristics}}\mathbf{B}_{\text{C}} + \boldsymbol{\Phi}_{i,t-1}^{\text{RecentPolicies}}\mathbf{B}_{\text{RP}}\right), \qquad (1)$$

[14] The VXO is the Volatility Index calculated by the Chicago Board Options Exchange and is similar to the VIX (with a correlation of 99%), except the VIX is only available starting in 1990. The U.S. interest rate is the policy rate. The commodity price index is the Economist All-Commodity dollar index.

[15] Capital flow data are from the dataset created in Forbes and Warnock (2012) and from the IMF's Balance of Payments. Private credit is measured by private credit by deposit money banks and other financial institutions to GDP and reported in Beck, Demirguc-Kunt, Levine, Cihak, and Feyen (2013). The commodity interaction term is defined as the product of the log of the commodity price index (defined above) multiplied by a dummy equal to 1 if a country is a major commodity exporter.

[16] Institutions are measured using an index based on the ICRG measures of institutional quality compiled by the World Bank.

[17] From Klein and Shambaugh (2013).

where pc_{it} is an episode dummy variable that takes the value of 1 if country *i* adopts a major policy change (reserve sale, currency depreciation, interest rate increase, or new control on capital outflows) in quarter *t*; $\mathbf{\Phi}_{t-1}^{\text{Global}}$ is a vector of global variables lagged by one quarter; $\mathbf{\Phi}_{i,t-1}^{\text{Vulnerabilities}}$, $\mathbf{\Phi}_{i,t-1}^{\text{Characteristics}}$, and $\mathbf{\Phi}_{i,t-1}^{\text{RecentPolicies}}$ are vectors of variables measuring country vulnerabilities, other country characteristics, and recent changes in related policies for country *i*, lagged by one quarter. All standard errors are robustly estimated.

Table 2 shows the estimates of this logit model for each of the four crisis-response policies. The model is best able to explain large currency depreciations (with a pseudo-R^2 of 0.39), followed by major reserve sales and significant increases in interest rates (with pseudo-R^2s of 0.26 and 0.25). It is less successful in predicting new controls on capital outflows (with a pseudo-R^2 of 0.11).

The coefficients on a number of the covariates capturing domestic vulnerabilities and other country characteristics are significant (at the 5% level) and have the expected sign. For example, episodes of major reserve sales are significantly more likely when capital outflows increase, capital inflows decrease, or private credit has grown more rapidly. Currency depreciations are significantly more likely when real GDP growth is weaker, the current account surplus is smaller (or deficit is larger), and for countries that are poorer, have less open capital accounts, fewer reserves (relative to GDP) and a pegged exchange rate (in the previous quarter). Major interest rate increases and new capital controls are more likely for countries that are major commodity exporters when commodity prices rise more quickly.

The coefficients reported at the top of Table 2 show that global variables, as well as domestic vulnerabilities, affect the likelihood of choosing different policy responses. For example, sharp depreciations are more likely during periods of higher global risk. Currency depreciations and increases in interest rates are more likely when commodity prices are higher. Currency depreciations and new controls on capital outflows are less likely after the United States increases its policy interest rate.

In this set of global covariates, the coefficients on the dummy variables for the 1990's crisis window are of special interest. These estimated coefficients are positive and significant for large currency depreciations and major increases in interest rates. This suggests that countries were significantly more likely to use these two policies during the 1990's crisis window than during the more recent 2000's crises after controlling for changes in the global environment,

domestic vulnerabilities, country characteristics, and recent policy changes. Moreover, the coefficient estimates for the 1990's crisis dummy suggest that the differences across the two crisis windows are not only significant, but large in magnitude. This is a different result than shown in Figures 1a and 1b and Tables 1a and 1b, which indicated a greater incidence of currency depreciations during the later crisis window and similar incidence for interest rate increases across both windows. These earlier figures and tables, however, did not control for other global and domestic variables.

We have conducted a number of robustness tests for the results in Table 2. One set of these tests includes additional variables, such as the country's debt to GDP, fiscal balance as a share of GDP, stock market capitalization relative to GDP, changes in the real exchange rate, a dummy for high income countries (as defined by the World Bank) and/or a dummy for major financial centers. The inclusion of many of these additional variables shrinks the sample size and, especially, the number of "treated" observations in which a country makes one of the four major policy changes. Substantial reductions in sample size make it impossible to estimate the effects of these policies with any precision, but robustness tests which maintain a reasonable sample size (including of treated observations) do not alter the key results reported below. As another robustness check, we estimated the model using a complimentary logarithmic (or cloglog) estimator (instead of the standard logit model) in order to adjust for the fact that major policy adjustments occur irregularly and $F(\cdot)$ is asymmetric. This also has no significant effect on the key results.[18]

Many of the coefficients reported in Table 2 are not significant. Including variables that are irrelevant (in the sense that that they do not influence the policy choice) in the first-stage regression used for propensity-score matching increases the variance of the estimates in the second stage and can make it more difficult to find appropriate matches.[19] Therefore, we also estimate stepwise regressions which only include the explanatory variables that are significant at the 20% level in the initial regressions in Table 2. These estimates are presented in Table 3.[20] As shown in that table, virtually all the regressors predicting major reserve sales, currency

[18] This specification assumes that the $F(\cdot)$ is the cumulative distribution function (cdf) of the extreme value distribution. In other words, this estimation strategy assumes that: $F(z) = 1 - exp[-exp(z)]$. See Forbes and Warnock (2012) for details on this estimation methodology.

[19] See Heinrich, Maffioli, and Vázquez (2010). Including irrelevant variables will make it more difficult to satisfy the common-support condition, as explained below.

[20] The regressions in Table 3 often include a larger sample size because dropping explanatory variables allows additional observations that did not have data for the larger set of covariates to be included in the final regression.

depreciations and interest rate increases are significant, including many which were not significant in Table 2. This more parsimonious specification will be used for the base case for our main results, although we report results based on the full set of variables in the robustness section.

The estimates presented in both Tables 2 and 3 also confirm that selection bias and endogeneity are not just hypothetical concerns when analyzing the impact of different policy responses. Selection bias can occur when countries that adopt certain policies have different domestic characteristics and vulnerabilities than those which do not. The two tables show, for example, that countries with smaller current account surpluses and without pegged exchange rates are more likely to undergo sharp currency depreciations. Similarly, endogeneity can occur when variables that are used to assess the effectiveness of these policies can, in turn, affect the decision to use these policies. For example, as shown on the tables, countries tend to have declines in real GDP growth before undergoing large currency depreciations, complicating any measurement of how the depreciations affect real GDP growth. Propensity-score matching offers one method of controlling for these country characteristics, vulnerabilities, and recent changes by matching groups of similar countries in order to assess the impact of the policies while controlling for selection bias and endogeneity. The next section discusses this method.

IV. Propensity-Score Methodology and Matching Results

This section begins with a brief discussion of the propensity-score methodology. Then it presents results and test statistics using propensity scores to match each treatment with a control group of observations.

A. Propensity-Score Methodology

To illustrate this propensity-score methodology and issue of sampling bias, define the adoption of the treatment by the i^{th} country (such as a major depreciation) as $D_i = 1$, and the absence of this action as $D_i = 0$. The outcome variable (say GDP growth) is $Y_{1,i}$ for the i^{th} member of the treated group and $Y_{0,i}$ for the i^{th} member of the untreated (control) group. Summing over members of each group, we are able to observe $E[Y_{1,i}|D_i=1]$ and $E[Y_{0,i}|D_i=0]$, but we are interested in the average treatment effect on the treated (ATT), which is $E[Y_{1,i} - Y_{0,i}|D_i=1]$. The difference in the two observable statistics reflects both the average effect of the treatment on the treated and selection bias. Selection bias, represented by $E[Y_{0,i}|D_i=1] -$

$E[Y_{0,i}|D_i=0]$, reflects the difference in outcomes that is attributable to differences in the treated and control group (such as different country characteristics) rather than any effect of the treatment itself.[21]

The effect of sampling bias could be easily minimized if countries differed along only one or two discrete, relevant dimensions with respect to the likelihood of undertaking a policy (such as undertaking a major depreciation). If the set of countries could be easily apportioned to a small number of "cells" reflecting all differences along these dimensions, and there were enough instances of both treated and control cases in each cell, then it would be straightforward to calculate the differences between the treated and the untreated in each cell and take a weighted average of those differences in order to estimate the effect of different treatments (i.e., policy changes).

In practice, however, there are many, multidimensional differences across countries, and it is impossible to match two countries which share identical macroeconomic characteristics. Propensity score matching offers a means to address this challenge. This methodology matches countries that undertake "treatments" (i.e., the four policy responses discussed above) to a subset of countries that do not, based on a set of observable country characteristics and global variables, represented by the vector X_i for the i^{th} country.[22] This matching controls for the differences in the treated and untreated groups that affect outcomes, such that the sampling bias disappears (at least any bias that is captured in the vector X_i). In other words, the underlying assumption is: $E[Y_{0,i}|X_i,D_i=1] = E[Y_{0,i}|X_i,D_i=0]$. While this still leaves a multidimensional problem, Rosenbaum and Rubin (1985) show that it is sufficient to match treated and control observations based on a "propensity score," $P(X_i)$, which is the probability that country *i* receives the treatment. This single propensity score reduces the number of dimensions over which observations must be matched. Rubin and Thomas (1992) further show that it is possible to estimate these propensity scores based on the vector of observable characteristics.

The propensity score is the conditional probability of adopting the treatment (in our case, the policy response), given pre-treatment characteristics, X_i. Continuing to define the adoption

[21] See Angrist and Pischke (2008, Chapter 3) for an excellent and intuitive presentation of this topic.

[22] This methodology has been widely used in medical studies and labor economics, but has only recently been employed in international and macroeconomic research. See Angrist and Pischke (2008) for an excellent summary of this methodology and examples from the labor literature. Four papers that have used this methodology in the international/macroeconomic literature are: Ehrmann and Fratzscher (2006) for the effect of monetary policy shocks on firms, Forbes, Fratzscher, and Straub (2013) for the effects of capital controls and macroprudential measures, and Glick, Guo, and Hutchison (2006) and Das and Bergstrom (2012) on the link between openness and currency crises and growth, respectively.

of the treatment as $D_i = 1$ (and not adopting the treatment as $D_i = 0$), the propensity score, $p(X_i)$ is:

$$p(X_i) = \Pr[D_i=1|X_i] \; . \qquad (2)$$

In the context of our model, propensity scores are the likelihoods that a country undergoes a major reserves sale, sharp currency depreciation, substantial interest rate increase, or adds new controls on capital outflows, controlling for the set of country and global characteristics, as specified in equation (1). In practice, the propensity scores can be generated using logit regressions, such as those reported in Tables 2 and 3.

After the propensity scores have been calculated, there are several algorithms that can be used to match each treated observation with one or more untreated observations (i.e., controls). We focus on five matching algorithms: nearest-neighbor without replacement, five-nearest neighbors with replacement, radius with caliper, kernel, and local-linear matching. Nearest-neighbor matching pairs each treated observation with the one observation from the control group that has the closest propensity score. We use this method "without replacement," which means that any observation from the untreated group can be matched to only one treated observation. We also use a variant of this method focusing on the five nearest neighbors (rather than one) as a control group, and allow "replacement" (so that untreated observation can be used more than once as a match). Another algorithm which uses the same basic approach is the radius method that includes all "nearest neighbors" whose estimated propensity scores fall within a maximum radius (referred to as the caliper).[23] The kernel and local-linear matching algorithms calculate a weighted average of all observations in the control group using nonparametric estimators which use generalized weighting functions to assign a higher weight to control observations with propensity scores closer to that of the treated observation.[24] The nearest-neighbor algorithms are basically extreme forms of kernel and local-linear matching, with all the weight given to the closest propensity score(s).

Each of these matching methodologies has advantages and disadvantages. Nearest neighbor is straightforward, easy to implement, and minimizes "bad" matches with control observations that have little in common with the treated observation. This method, however,

[23] We set the caliper at 0.05.

[24] The main difference between the two methods is the weighting functions. See Heckman, Ichimura, and Todd (1997, 1998) for a detailed description of the local-linear matching method.

ignores useful information from other countries in the control group. Radius, kernel, and local-linear matching use more information and therefore tend to have lower variances, but at the risk of including bad matches. Radius matching is less sophisticated than kernel and local-linear matching, as it does not place greater weight on better matches within its "radius". Fan (1992a, b) shows that local-linear matching has several important advantages over kernel matching, such as a faster rate of convergence near boundary points and greater robustness to different data design densities. In the following analysis, we begin with each of the five different matching approaches and then use several tests to evaluate their performance and select the base case for each analysis. Including the different matching methodologies is also useful to test for the robustness of the results, especially as the significance of key results can depend on the construction of the control group. These tests used to assess the accuracy of the matching algorithms are discussed in the next section.

To conclude this section on the propensity-score methodology, it is useful to mention how this approach compares to the more familiar regression analysis. Multivariate regressions estimate the partial correlation of the treatment with the outcome variable, and, like the matching method, control for the other variables included in the equation. The main difference between the standard regression approach and the propensity-score methodology, however, is the weighting of the covariate-specific differences between the treated and the untreated observations.[25] More specifically, consider cells constructed from discrete values of X. In both approaches, it is necessary to construct weights for the difference between treated and untreated values across different cells in order to calculate the average effect for the whole sample. Using propensity-score methodology, the weights are based on the distribution of covariates among the treated, with the greatest weights put on cells representing the highest likelihood of being treated. In contrast, in regression analysis, the greatest weights are placed on cells where the conditional variance of treatment status is larger; roughly speaking, those cells with equal likelihood of its elements being treated or untreated.

These two different weighting approaches can significantly affect the estimated average effects of the treatments (i.e., policy changes) if the differences between the treated and the

[25] Angrist and Pischke (2008, Chapter 3) present an excellent discussion of the similarities and differences between regression analysis and propensity-score matching, including their relative advantages. Also, although propensity-score matching can reduce asymptotic efficiency relative to a regression framework, Angrist and Hahn (2004) show that there can be efficiency gains in a finite sample, even if there is no asymptotic efficiency gains from the use of propensity-scoring estimators. Given the small size of our sample—this suggests that this potential drawback of the propensity score methodology is less likely to be an issue.

untreated observations vary across cells. For example, when estimating whether depreciations significantly affect growth, a propensity-matching approach would put less weight on the cells representing countries that had large current account surpluses, pegged exchange rates, and open capital accounts than would be the case with a regression approach. This would occur because large depreciations are less likely in countries with large current account surpluses, pegged exchange rates, and open capital accounts (as shown in the estimates in Tables 2 and 3). In other words, the propensity score methodology would put more weight on controls that are similar according to these dimensions of the vector *X* when constructing the counterfactual.

B. Results of Propensity-Score Matching

We use the estimates from the logit model in Table 3 to calculate separate propensity scores for each of the four policies (reserve sales, currency depreciations, interest rate increases, and new controls on capital outflows) for each quarter during our crisis periods (1997-2001 and 2007-2011). We then use these estimated propensity scores to match treated observations with control observations using the five matching algorithms discussed above.[26] In order to avoid using a country that is about to make or has just made one of these policy changes as a control variable for that same policy change, we continue to include an "exclusion window" for the 3 quarters before and the 3 quarters after a major policy change. During this exclusion window, a country cannot be included as a control observation for that same type of policy (although it could serve as a control for a country that undertook a different policy change). A country can also serve as its own control, albeit not within the period of the exclusion window.

Matching algorithms can be assessed by considering whether their results satisfy certain criteria. One criterion is whether observations are "on support" or "off support". A treated observation is "on support" when its propensity score is between that of the minimum and maximum propensity scores for the untreated observations. If this is not the case, then the treated observation is "off support" and comparisons with the untreated observations are less accurate. Fortunately, the logit model in Table 3 performs well in terms of generating matches that are on-support for most of the major policy responses. Table 4 reports the number of observations in the treatment group and unmatched control group for each of the four policy responses, and the number of observations (in parentheses) that are "off-support" after using

[26] We apply these matching algorithms with the Stata module PSMATCH2, developed by Leuven and Sianesi (2003). The number of treated observations is lower than reported in Table 1 because data are not available for the all of the covariates needed to estimate propensity scores for all observations.

each matching algorithm. It shows that for each of the five matching algorithms, the treated observations are on-support for all 29 episodes of major reserve sales, all 59 of sharp interest rate increases, all 40 of new capital controls, and for 74 of the 75 episodes of large currency depreciations. The single treatment which is off-support (for each of the five matching algorithms) is Russia's depreciation in the first quarter of 2009. The literature on matching suggests that the use of off-support matches can be problematic and, therefore, we impose a common-support condition which drops all observations with a propensity score higher than the maximum or lower than the minimum propensity score of the controls in order to reduce the effect of any "bad" matches.

A second key criterion to assess if a matching methodology is valid is known as the test of "balancing" or the "independence assumption". The goal of this test is to verify that the matching was able to remove any significant differences in observable variables between the treated and control groups, i.e., that:

$$D \perp X \mid p(X), \tag{3}$$

where *D* signifies whether an observation is treated or in the control group.

Table 5 presents results of tests of the independence assumption for the policy response (treatment) of large currency depreciations. The first two columns report the mean values for the treated group (μ_T) and control group (μ_C) for the unmatched samples for each of the covariates used to estimate the propensity scores. The third column reports *t*-statistics for tests of the hypothesis that the mean of each variable in the treatment group is equal to the mean in the control group (H0: $\mu_T = \mu_C$). There are significant differences (at the 95 percent confidence level) between the treatment and unmatched control groups for twelve of these variables. Countries were more likely to have a large depreciation during periods of higher global risk, reductions in U.S. interest rates, and lower commodity prices, and if they had slower real GDP growth, bigger current account deficits, lower income per capita, worse institutions, a floating exchange rate, less financial openness, and had sold more reserves or had a larger depreciation in the last period. These significant differences across the treatment and unmatched control groups highlight that selection bias is important; countries which had a major depreciation had significantly different characteristics before the depreciation than other countries.

The remaining columns of Table 5, however, indicate that the matching algorithms are able to remove this selection bias (at least as measured by the included variables) in almost all cases. These columns show mean values for each of the covariates in the matched control groups after using different matching algorithms: five-nearest neighbors, radius, kernel, and local-linear.[27] It also reports the same t-statistics of tests for significant differences between the treatment and matched control groups for each of the variables. In contrast to the results for the unmatched control group, there are no longer significant differences for any of the variables between the treatment and the matched control groups based on five-nearest neighbor and local-linear matching. With radius and kernel matching, there is a significant difference between the treatment and matched control groups for only one variable (whether the country's exchange rate is pegged). These results suggest that when analyzing the impact of large currency depreciations, five-nearest neighbor and local-linear matching satisfy the independence assumption, but radius and kernel matching do not.

These results in Table 5 showing the ability of propensity-score matching to remove significant differences between the variables in the treatment and the matched control groups for certain matching algorithms also applies when this analysis is performed for major reserve sales, large increases in interest rates, and increased capital controls. Table 4 summarizes these results by reporting how many variables do not satisfy the independence assumption for each of the four policy responses and five matching algorithms. It shows that five-nearest neighbor matching satisfies this requirement for each of the policy responses. Local-linear and nearest-neighbor matching satisfy it for all policy responses except for increased capital controls (when the independence assumption is not met due to significant differences in the reserves/GDP ratios between the treatment and matched control groups). Radius and kernel matching do not meet the independence assumption for major depreciations (as discussed above) and for large increases in interest rates.

As a final analysis of the performance of the different matching algorithms, Table 4 reports several statistics from propensity-score matching for each of the four policy responses. The statistics include the mean propensity scores for the treated group, unmatched control group, and matched control group using each of the matching algorithms. The mean propensity scores after matching are usually closer to that of the treatment group than for the unmatched control

[27] Results for nearest-neighbor matching are nor reported due to space constraints and because this is not the preferred methodology (as discussed below), but results are similar to those for five-nearest neighbor.

group. The main exception is when matching is done using nearest neighbor—which often generates a mean propensity score further from the treatment group than with the unmatched control group. This highlights the shortcoming of using just one nearest neighbor for matching, and as a result we will not focus on results obtained using this matching algorithm.

Table 4 also reports the mean absolute bias of the treatment group relative to the unmatched control group and control groups using each of the matching algorithms. In each case, the matching reduces the mean absolute bias by a substantial amount. In order to simplify the discussion that follows, we will focus on results obtained using the five-nearest neighbor matching algorithm as it always satisfies the independence assumption. It also generally has the lowest mean absolute bias and generally performs well, relative to the other methods, with respect to the mean propensity score being close to that of the treatment group. To demonstrate the robustness of our analysis, we will also report results obtained by local-linear matching, since it uses the full set of control observations (unlike with five-nearest neighbors), it generally performs better than the radius and kernel matching, and in some of the sensitivity tests is performs better according to these tests than other algorithms.

V. Effects of Policy Choices: Average Treatment Effects on the Treated (ATT)

This section uses the matched control groups to analyze the effects of large reserve sales, sharp currency depreciations, major increases in interest rates, and new controls on capital outflows (the four treatments or policy responses) on three outcome variables: real GDP growth, inflation, and unemployment. The effect of these policies is calculated as the average treatment effect on the treated or ATT (as described in Section IV.A). This approach tests for any significant differences in outcomes between the countries which adopted the policy response relative to the outcome variables averaged over the constructed control group. This section begins by discussing the approach in more detail and reporting the results for the base case with the full sample of countries in both crisis periods. The section ends with a series of robustness tests and extensions—including specifications using different control variables, including only emerging markets and non-OECD economies, and for each of the crisis windows separately.

A. Base Case Results

In order to evaluate the effects of the four different policy responses, we test for the ATT on outcome variables at each quarter over a one-and-one-half year period, beginning with the

quarter that the major policy response (the treatment) occurs and over each of the next six quarters. This enables us to evaluate both immediate and longer-run effects. We do not consider effects beyond one-and-one-half years because the matching algorithms become less accurate over longer time periods. We calculate the change in each of the outcome variables (real GDP growth, CPI inflation, and unemployment) relative to the average over the two years before the policy change occurred.[28] The standard errors of the ATT, which are required to gauge whether there is a statistically significant difference between the treated and control groups, must be calculated through bootstrapping methods because the propensity scores are estimated.[29]

To illustrate this approach, consider the ATT for the effect of large reserve sales on real GDP growth. We calculate the average difference between the change in real GDP growth between the treated and control groups. For the treatment period, this is real GDP growth during the quarter of the large reserve sales (t=0) relative to average GDP growth over the previous eight quarters (t-1, t-2, t-3, … t-8). For the post-treatment periods, this is the change in GDP growth in the respective periods relative to average GDP growth over the same eight quarters before the major reserve sales (periods t-1, t-2, t-3, … t-8). This approach allows us to capture any differential effects of major policy responses over different time periods rather than choosing, *a priori*, the time period on which to focus. This is a *ceterus paribus* approach in that it does not incorporate any adjustment for changes in post-treatment covariates.

The average values of the outcome variables for the control groups that are used to estimate the ATT are calculated differently based on the selected matching algorithm. As discussed in Section IV.A., average values based on any of the nearest neighbor or radius algorithms use equal weights for all of the observations included in the control group, but do not include any information for observations that are not in the control group. Average values based on the kernel and local-linear algorithms use the full set of control observations, but place higher weights on observations with propensity scores closer to that of the treated observations. For our base-case analysis, we focus on ATTs calculated using the five-nearest neighbor and local-linear matching algorithms, the two that perform best according to the series of tests discussed in the last section.

The most straightforward way to characterize the results of this ATT analysis is bar graphs of the estimated effects over each of the seven quarters. Figures 2 through 4 present these

[28] We use the average over the previous eight quarters in order to minimize seasonal affects and also to smooth the baseline value that we use as a comparison to the post-treatment values.
[29] See Lechner (2002) for the appropriate methodology. We use 500 repetitions for the bootstrap.

graphs for each of the four policies on GDP growth, unemployment, and CPI inflation, respectively. The height of the bars shows the magnitude of the estimated ATT and the color of the bars indicates whether the effect is significant. Dark black indicates that the ATT for that quarter is significant at the 5% level or less, medium-blue indicates significance at the 5% to 10% level, and light blue indicates that the effect is insignificant. The black line is the fitted line for the average treatment effect.

Figure 2 indicates that each of the four policies has significant effects on GDP growth, but in different directions and at different time horizons. Focusing on results that are significant at the 5% level, major reserve sales and sharp currency depreciations have positive and significant effects on GDP growth, although these positive effects do not occur until four to six quarters after the policy is undertaken. The effects may be large, reaching a maximum increase in GDP growth of almost 4 percentage points after major reserve sales and 2 percentage points after major currency depreciations, although the initial effect in each case may be a decrease in GDP growth of close to 2 percentage points, especially for the first quarter after a sharp depreciation. This pattern after depreciations is consistent with the standard J-curve effect as well as a balance-sheet effect in which a depreciation increases liabilities denominated in foreign currency relative to assets denominated in domestic currency. Large increases in interest rates and new controls on capital outflows are associated with sharp and significant declines in GDP growth, estimated to reach a maximum effect of reducing GDP growth by 4 percentage points. This negative effect occurs more quickly after an increase in interest rates than after new capital controls. These results suggest that countries hoping to pursue policies to support GDP growth over the medium term during crises should consider reserve sales or depreciations, although the initial impact of these policies may be slower growth.

Figure 3 presents results for the average treatment effects on unemployment. In contrast to the significant effects on growth in Figure 2, none of the policy responses generates a significant (at the 5% level) impact on unemployment. The different policies appear to cause unemployment to move in different directions, however, and some of the effects are significant at the 10% level. More specifically, major reserves sales and large currency depreciations reduce unemployment relative to the counterfactual (and these effects are sometimes significant at the 10% level using different matching algorithms), while large increases in interest rates and new controls on capital outflows appear to increase unemployment over time. These different patterns are not surprising as they correspond to the direction of the different effects on growth in Figure

2. It is also not surprising that any effects of the policies on unemployment are lagged (such as from reserve sales), which is consistent with the generally lagged response of unemployment to GDP. One possible source of the imprecision of these estimates is the large differences in the institutions and employment practices across countries that determine how changes in GDP affect unemployment.

Figure 4 shows the estimated effects of each of the four policies on CPI inflation. The most striking result from this series of graphs is the large and significant effect of sharp currency depreciations on inflation. This effect increases over several quarters to peak at in increase in inflation of about 8 percentage points after one year. Major reserve sales and interest rate increases are also correlated with increased inflation, although these estimated effects are usually insignificant and smaller in magnitude.

The results presented in Figures 2, 3, and 4 show that different policy responses to crises can yield very different outcomes. Large currency depreciations and major reserve sales appear to be most beneficial in terms of supporting GDP growth relative to the counterfactual, although the benefits are lagged, may only arrive after an immediate reduction in GDP growth, and may generate a substantial increase in inflation—especially after sharp currency depreciations. Any benefits in terms of reducing unemployment are usually not significant. Sharp increases in interest rates and controls on capital outflows appear to yield the least benefit in terms of the criteria evaluated, as they correspond to significant and economically meaningful reductions in GDP growth, while yielding no benefits in terms of significantly reducing inflation or unemployment. Countries responding to crises and contractions in global capital flows therefore have no ideal policy response that yields positive outcomes for all three of these key measures of economic performance.

B. Sensitivity Tests: ATTs Based on Different Covariates in the Logit Regressions

As discussed in Section III, the selection of covariates to include in the first-stage logit regressions used to estimate the propensity scores is difficult due to the large number of variables that could be included and the lack of clear theoretical guidance on which variables are most important. In order to evaluate if the base-case results reported above are sensitive to the choice of covariates, this section discusses results when different control variables are included in the logit regressions.

Our first sensitivity test uses the full set of covariates included in the initial regression in Table 2, rather than the subset of variables reported in Table 3. Tests of the different matching algorithms indicate that local-linear matching performs the best, so we focus on these results in the discussion below.[30] Figure 5 displays the resulting ATT graphs. Major reserve sales and sharp currency depreciations continue to increase GDP growth after several quarters, while large increases in interest rates and new controls on capital outflows continue to significantly reduce GDP growth. The directions of the estimated effects on unemployment mirror those for GDP growth and continue to be insignificant. The estimated effects of the different policies on inflation are also similar—with large depreciations and interest rate increases associated with significantly higher inflation (as found in both cases with local-linear matching in Figure 4). The main difference in these results as compared to those that use the more limited set of variables is that, as expected, the larger set of covariates leads to less precise estimates with fewer significant estimated ATTs. For example, the ATTs for the impact of depreciations on GDP growth are now significant at the 10% level (instead of the 5% level) after 6 quarters, and the effect of new capital controls on GDP growth is less often significant when the first-stage logit regressions include the larger set of (often insignificant) covariates.

We have also attempted to estimate an augmented version of the base-case analysis that includes variables such as budget balances relative to GDP, government debt to GDP, and financial market size. These additional variables are not available for the full sample of observations in the base case, however, and their inclusion reduces the sample size and the number of treated observations to the point that the matching methodology no longer works well. Augmenting the analysis with other variables that are more widely available, such as a dummy variable for high-income economies, a dummy variable for financial centers, and changes in the "global" interest rate (calculated as the average of the policy rate in the United States, Japan, and euro area) does not significantly affects the key results discussed above.

C. Extension: ATTs for Emerging and Non-OECD Economies

Next we repeat the base-case analysis for a subsample of countries that excludes high-income OECD economies (based on World Bank classifications, which vary each year). High-

[30] More specifically, radius and kernel matching have more treated observations that are off-support and occasionally fail the tests of the independence assumption. Five-nearest neighbor and local-linear matching perform similarly according to these tests, but the mean propensity scores for the matched control groups based on local-linear matching are closer to the mean scores for three of the four policy responses.

income OECD economies may face a different set of tradeoffs in their choice of policy responses to crises for a number of reasons. First, they are more likely to hold a reserve currency, for which demand would more likely increase instead of decrease during crises. Second, they may have a greater ability to lower (instead of increase) interest rates during crises since they are less reliant on capital inflows (obviously excluding individual members of the euro zone). Third, they may have less ability to enact new capital controls if they have large and highly sophisticated financial sectors, or have signed agreements limiting their ability to use capital controls.

We begin this analysis by estimating the first-stage logit regressions using the large set of covariates reported in Table 2 for the sample that does not include high-income OECD countries. Then we only include variables that are significant at the 20% level (or less) for the final logit regression used to predict the use of each of the four policy responses, resulting in a set of control variables that differs from those in Table 3. For this analysis, the local-linear matching algorithm has the best performance according to the series of tests of the accuracy of the matching methodologies, so we focus on results based on this algorithm.[31]

The ATTs from this exercise are presented in Figure 6. Many of the results correspond to those for the larger sample which included high-income OECD economies. Major increases in interest rates and new capital controls initially significantly reduce GDP growth relative to the counterfactual. Large currency depreciations and major increases in interest rates continue to be correlated with significant increases in inflation. But there are also differences between the results presented in Figure 6 and those presented in Figures 2 and 3. Major reserve sales and currency depreciations no longer have significant effects on GDP growth at the 5% level (although the lagged effect on GDP growth is still positive and significant at the 10% level for reserve sales). More striking, major reserve sales are now correlated with a significant increase—instead of decrease—in unemployment relative to the counterfactual. Moreover, the estimated magnitude of this effect is large—peaking at an increase in unemployment of 2.5 percentage points. These results suggest that sharp depreciations, and especially major reserve sales, may be less beneficial for emerging and non-OECD economies in terms of supporting growth and employment.

[31] More specifically, each of the five matching algorithms has a similar number of countries that are off-support for each policy response and matching algorithm. (In each case, more observations tend to be off support—including 3 for major reserve sales, 2 for sharp depreciations, and 1 for increases in interest rates). Each policy response and matching algorithm also satisfies the independence assumption. The mean propensity scores for the control group resulting from local-linear matching, however, are closest to those for the treatment groups for 3 of the 4 policy responses (and 2nd closest for the other policy response).

D. Extension: ATTs for Each Crisis Period

Our final robustness exercises estimate effects separately for each crisis window, the "1990's crises" from 1997 to 2001 and the "2000's crises" from 2007 to 2011. While both crisis periods included increases in volatility and sudden stops in capital flows, they also differed in ways that could affect the choices of policy responses and the tradeoffs of using different policies; for example, the more global nature of the slowdown during the GFC could reduce inflationary pressures and thereby reduce the costs of a policy such as a sharp depreciation.[32]

We begin this extension by estimating separate logit regressions for each crisis window using the large set of covariates listed in Table 2, and then use a smaller set of covariates to calculate the propensity scores for the matching. One difference between this exercise and the base case is that the matching algorithms have a substantially larger number of country-quarter observations that are off-support. For example, four observations are off-support for each of the matching algorithms predicting increased interest rates during the 2000's crises. Between 10 observations (for five-nearest neighbor and local-linear matching) and 17 observations (for radius and kernel matching) are off-support when predicting major currency depreciations in the 2000s window. In contrast, no more than 1 observation was ever off support in the base case which combined the two crisis periods. Most of the matching algorithms also fail the tests of the independence assumption for at least one of the policy responses during each of the separate crisis windows. The only matching algorithm that satisfies the independence assumption for each policy response in each crisis window is five-nearest neighbors, so we focus on results obtained using this methodology.[33]

Figures 7 and 8 graph the ATTs corresponding to the graphs in Figures 2, 3, and 4—except for just the period from 1997-2001 (Figure 7) and from 2007-2011 (Figure 8). These results should be interpreted cautiously due to the inferior performance of the propensity-score matching in these more limited samples. In general, however, the estimated direction of the effects of the different policy responses agrees with those reported in the base case and sensitivity tests. Major reserve sales and sharp currency depreciations are more likely to generate

[32] Claessens and Kose (forthcoming 2013) discusses similarities and differences across crises.

[33] Even five-nearest neighbor matching, however, yields a large number of observations that are off-support. For example, 1 observation is off-support for major depreciations and increases in interest rates during the 1990's crises. During the 2000s' crises, 3 observations are off-support for major reserve sales, 10 for sharp depreciations, and 4 for major increases in interest rates.

higher GDP growth with a lag, and major increases in interest rates and new capital controls tend to lower GDP growth. But many of these effects are now insignificant, especially during the 2000's crisis window. It is impossible to discern if any differences between Figures 7 and 8 reflect different effects of policies during the two crisis periods, as the insufficient number of observations with sufficient variation in the smaller samples makes the propensity-score methodology less accurate.

VI. Conclusions

For years, economists have debated how best to respond to periods of sudden stops in global capital flows. This debate is likely to continue for many more years. This paper contributes to this debate by proposing a new approach for analyzing the consequences of different policy responses by using a propensity-score matching methodology. This methodology helps address the challenge that countries which adopt certain policy responses are generally different than countries which do not choose these policies, and are generally responding to changes in variables which they are intending to influence. Propensity-score matching takes these econometric issues seriously when evaluating the effects of policy responses to crises.

The results indicate that large currency depreciations and major reserve sales can provide an important benefit in terms of supporting GDP growth relative to the counterfactual, although this effect is lagged, may only arrive after an initial reduction in GDP growth, and may generate a substantial increase in inflation—especially after large depreciations. Moreover, the growth benefits from sharp depreciations and major reserve sales appear to be weaker in emerging markets and other non-OECD economies. None of these policies significantly improves unemployment in the full sample of countries, and large reserve sales significantly worsen unemployment in emerging and other non-OECD economies relative to the counterfactual. Sharp increases in interest rates and controls on capital outflows appear to yield even less benefit, as they correspond to significant and economically meaningful reductions in GDP growth for the full sample, while yielding no benefits in terms of reducing inflation or improving unemployment. Countries responding to crises and contractions in global capital flows therefore have no ideal policy response that yields positive outcomes for growth, unemployment, and inflation. Policy makers must "pick their poison" and face challenging trade-offs during crisis periods.

References

Aizenman, Joshua and Gurnain Pasricha. (2013). "Why Do Emerging Markets Liberalize Capital Outflow Controls? Fiscal Versus Net Capital Flow Concerns." NBER Working Paper 18879.

Angrist, Joshua and Jinyong Hahn. (2004). "When to Control for Covariates? Panel Asymptotics for Estimates of Treatment Effects." *Review of Economics and Statistics*. 86: 58-72.

Angrist, Joshua and Jorn-Steffen Pischke. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Beck , Thorsten, Asli Demirguc-Kunt , Ross Eric Levine , Martin Cihak and Erik H.B. Feyen. (2013). *Financial Development and Structure Dataset (updated April 2013). Available at: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:20696167~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html*

Blanchard, Olivier, Mitali Das and Hamid Faruquee. (2010). "The Initial Impact of the Crisis on Emerging Market Countries," *Brookings Papers on Economic Activity*, (Spring): 263 – 307.

Calvo, Guillermo. (1998). "Capital Flows and Capital-Market Crises: The Simple Economics of Sudden Stops." *Journal of Applied Economics* (Nov): 35-54.

Calvo, Guillermo, Alejandro Izquierdo, and Luis-Fernando Mejía. (2008). "Systemic Sudden Stops: The Relevance of Balance-Sheet Effects and Financial Integration." NBER Working Paper 14026.

Chari, Anusha and Peter Blair Henry. (2013). "This Time IS Different: East-Asian Lessons for a Post-Crisis World." Mimeo.

Chinn, Menzie and Hiro Ito. (2008). "A New Measure of Financial Openness." *Journal of Comparative Policy Analysis* 10(3, September): 309-322. Dataset updated as of 04/24/2103 available at: http://web.pdx.edu/~ito/Chinn-Ito_website.htm

Claessens, Stijn, Giovanni Dell'Ariccia, Deniz Igan and Luc Laeven. (2010). "Global Linkages and Global Policies." *Economic Policy* (April): 267 - 293.

Claessens, Stijn and Ayhan Kose. (2013). "Financial Crises: Explanations, Types and Implications." In Claessens, Stijn, Ayhan Kose, Luc Laeven and Fabian Valencia, eds., *Financial Crises: Causes, Consequences and Policy Responses*, forthcoming December.

Claessens, Stijn, Ayhan Kose, Luc Laeven and Fabian Valencia, eds. (2013). *Financial Crises: Causes, Consequences and Policy Responses*, forthcoming December, IMF.

Das, Kuntal, and Katy Bergstrom. (2012). "Capital Account Liberalization, Selection Bias, and Growth." Mimeo.

Dominguez, Kathryn. (2013). "Exchange Rate Implications of Reserve Changes." Unpublished mimeo.

Ehrmann, Michael and Marcel Fratzscher. (2006). "Global Financial Transmission of Monetary Policy Shocks." European Central Bank Working Paper No. 616 (April).

Fischer, Stanley. (2004). "The IMF and the Asian Crisis." Chapter 3 in *IMF Essays from a Time of Crisis*. Cambridge, MA: MIT Press.

Forbes, Kristin, Marcel Fratzscher, and Roland Straub (2013). "Capital Controls and Prudential Measures: What are they Good For?". Mimeo.

Forbes, Kristin and Francis Warnock. (2012). "Capital Flow Waves: Surges, Stops, Flight and Retrenchment." *Journal of International Economics* 88(2, Nov.): 235-251.

Frankel, Jeffrey and George Saravelos. (2012). "Can Leading Indicators Assess Country Vulnerability? Evidence from the 2008-09 Global Financial Crisis." *Journal of International Economics* 87(2, July): 216-231.

Glick, Reuven, Xueyan Guo, and Michael Hutchison. (2006). "Currency Crises, Capital-Account Liberalization, and Selection Bias." *Review of Economics and Statistics* 88(4, Nov.): 698-714.

Gourinchas, Pierre-Olivier and Maurice Obstfeld. (2012). "Stories of the Twentieth Century for the Twenty-First." *American Economic Journal: Macroeconomics* 4(1, January): 226 – 265.

Heckman, J.J., Ichimura, H. and Todd, P.E. (1997). "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* (64): 605-654.

Heckman, J.J., Ichimura, H. and Todd, P.E. (1998.) "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65: 261-294.

Heinrich, C., A. Maffioli and G. Vazquez. (2010), "A Primer for Applying Propensity-Score Matching." Technical notes No. IDB-TN-161, Inter-American Development Bank.

Klein, Michael (2012). "Capital Controls: Gates versus Walls," *Brookings Papers on Economic Activity*, (2, Fall): 317–355.

Klein, Michael and Jay Shambaugh (2013). "Rounding the Corners of the Policy Trilemma: Sources of Monetary Policy Autonomy," *N.B.E.R. Working Paper No.* 19461, September.

Lechner M. (2002). “Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods.” *Journal of the Royal Statistical Society*, Series A 165: 59-82.

Leuven, E. and B. Sianesi. (2003). “PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.” http://ideas.repec.org/c/boc/bocode/s432001.html.

Obstfeld, M., J. Shambaugh and A. Taylor. (2010). “Financial Stability, the Trilemma, and International Reserves,” *American Economic Association Journal – Macroeconomic*, 2(2, April): 57 – 94.

Rose, Andrew. and Mark Spiegel. (2010). “The Causes and Consequences of the 2008 Crisis: International Linkages and American Exposure.” *Pacific Economic Review* 15(3, August): 340 - 363.

Rosenbaum, Paul R. and Donald B. Rubin. (1985). “The Bias Due to Incomplete Matching,” *Biometrics* 41: 106 – 16.

Rubin, Donald and Neal Thomas. (1992). “Characterizing the Effect of Matching Using Linear Propensity Score Methods with Normal Distributions.” *Biometrika* 79: 797-809.

Table 1a: Policy Changes, 1997 – 2001

	Large Reserve Sales	Sharp Depreciations	Major Increases in Interest Rates	New Capital Controls	*Total*
Reserve Sales	26	1	2	1	*33 (23%)*
Depreciations	3	35	5	0	*45 (31%)*
Interest Rates		1	28	0	*39 (27%)*
Capital Controls				24	*26 (18%)*
					143

Table 1b: Policy Changes, 2007 – 2011

	Large Reserve Sales	Sharp Depreciations	Major Increases in Interest Rates	New Capital Controls	*Total*
Reserve Sales	15	4	0	0	*19 (20%)*
Depreciations		31	0	1	*36 (38%)*
Interest Rates			23	1	*24 (25%)*
Capital Controls				15	*17 (18%)*
					96

Notes: Reports instances of major policy changes, with thresholds defined as occurring in only 5% of the country-quarter observations in the full sample of crisis years (1997-2001 and 2007-2011) and subject to the 3 quarter exclusion window. Large reserve sales are at least a 24% decrease in reserves between one quarter and that quarter in the previous year (and only if the reserve-to-GDP ratio is at least 10 percent). Sharp depreciations are at least a 23% deprecation in the nominal exchange rate versus the US$ between one quarter and that quarter in the previous year (and not preceded by a quarter with annual inflation of 20% or more); Major increases in interest rates are at least a 244 basis points increase between one quarter and that quarter in the previous year (and not preceded by a quarter with annual inflation of 20% or more). New capital controls are new controls on capital outflows (as described in the text). Diagonal elements represent policy changes that occurred when none of the other policy changes occurred in that quarter, in the preceding quarter, or in the following quarter. Upper triangular elements represent instances of pairs of policies, lower triangular elements represent instances of triplets of policies, with pairs and triplets representing polices that occurred in the same quarter, or in adjoining quarters.

Table 2: Probability of Adopting Policies During Crises (Full Set of Covariates)

		Reserve Sales	Currency Depreciations	Interest Rate Increases	New Capital Controls
Global Measures	Global risk	-0.016	0.070***	0.029	-0.095*
		(0.031)	(0.020)	(0.028)	(0.055)
	US interest rate (ch)	-0.141	-0.455***	-0.024	-0.460**
		(0.173)	(0.173)	(0.156)	(0.233)
	Commodity price Index (log)	0.831	3.911**	5.472***	1.199
		(1.813)	(1.696)	(1.442)	(1.386)
	1990's crisis dummy	2.174	2.892**	5.589***	1.698
		(1.511)	(1.424)	(1.453)	(1.204)
Domestic Vulnerabilities	Real GDP growth (ch)	0.013	-0.071***	-0.030*	-0.017
		(0.024)	(0.021)	(0.017)	(0.027)
	Capital outflows, % of GDP (ch)	5.000***	-0.797	0.383	-0.607
		(1.235)	(1.189)	(3.960)	(0.834)
	Capital inflows, % of GDP (ch)	-4.839***	0.089	0.788	-0.099
		(1.345)	(1.317)	(2.143)	(0.754)
	Current account % of GDP	0.531	-2.008**	-1.141	-0.299
		(1.046)	(0.937)	(1.124)	(0.737)
	Inflation (ch)	0.028	0.036	0.001	-0.034
		(0.055)	(0.033)	(0.021)	(0.027)
	Private credit, % of GDP (ch)	0.052***	-0.019	-0.009	0.007
		(0.017)	(0.018)	(0.024)	(0.018)
	Commodity index * Comm. exporter	0.086	0.560	1.189***	0.905**
		(0.613)	(0.370)	(0.456)	(0.429)
Other Country Characteristics	Income per capita (log)	-0.353	-0.613***	-0.273	-0.059
		(0.268)	(0.216)	(0.211)	(0.208)
	Institutions index	-0.185	2.024	-2.150	-0.818
		(1.605)	(1.563)	(1.917)	(1.076)
	Reserves as % of GDP	-1.941*	-2.646***	-0.948	-0.483
		(1.041)	(0.861)	(0.776)	(0.507)
	Peg dummy	-0.246	-3.598**	-0.383	0.485
		(0.671)	(1.452)	(0.855)	(0.639)
	Openness	-0.168	-0.320**	-0.332*	-0.088
		(0.210)	(0.136)	(0.197)	(0.187)
	Euro member dummy	1.113	-1.377	-1.443	-0.330
		(0.718)	(1.512)	(1.636)	(0.849)
Recent Policy Changes	Reserves as % of GDP (ch)	-12.874***	-5.709***	-3.345	-0.402
		(1.974)	(2.013)	(2.149)	(1.350)
	ER vs. US $ (pch)	0.024*	0.019	-0.007	-0.003
		(0.013)	(0.012)	(0.016)	(0.017)
	Interest rate vs. U.S. rate (ch)	-0.083	-0.023	0.117*	-0.075*
		(0.086)	(0.064)	(0.062)	(0.042)
	Capital controls (ch)	0.466	0.331	0.036	1.041*
		(1.233)	(0.640)	(0.698)	(0.570)
Observations		***1,412***	***1,297***	***1,369***	***1,432***
Pseudo R-squared		***0.264***	***0.388***	***0.254***	***0.114***

Notes: Logit regressions predicting probability of policy listed at top and defined in Section II. "Ch" denotes change and "Pch" is percentage change. Constant not reported. * is significant at the 10% level, ** at the 5% level and *** at the 1% level. Includes robust standard errors.

Table 3: Probability of Adopting Policies During Crises (Limited Set of Covariates)

		Reserve Sales	Currency Depreciations	Interest Rate Increases	New Capital Controls
Global Measures	Global risk		0.065***		-0.092**
			(0.015)		(0.038)
	US interest rate (ch)		-0.439***		-0.389**
			(0.113)		(0.163)
	Commodity price Index (log)		2.650**	3.074***	
			(1.070)	(0.812)	
	1990's crisis dummy	1.483***	2.481***	2.870***	1.024**
		(0.496)	(0.830)	(0.668)	(0.444)
Domestic Vulnerabilities	Real GDP growth (ch)		-0.045***	-0.028*	
			(0.015)	(0.017)	
	Capital outflows, % of GDP (ch)	3.900***			
		(0.860)			
	Capital inflows, % of GDP (ch)	-3.825***			
		(0.832)			
	Current account % of GDP		-1.800***		
			(0.580)		
	Inflation (ch)				
	Private credit, % of GDP (ch)	0.041***			
		(0.012)			
	Commodity index * Comm. exporter		0.592**	0.720**	0.232
			(0.299)	(0.289)	(0.342)
Other Country Characteristics	Income per capita (log)	-0.176	-0.063	-0.603***	
		(0.191)	(0.154)	(0.169)	
	Institutions index		-0.109		-1.029
			(1.137)		(0.771)
	Reserves as % of GDP	-1.266*	-0.111	-0.431*	0.178
		(0.733)	(0.338)	(0.243)	(0.174)
	Peg dummy		-2.752***		
			(0.578)		
	Openness		-0.212**	0.058	
			(0.104)	(0.137)	
	Euro member dummy				
Recent Policy Changes	Reserves as % of GDP (ch)	-9.965***	-4.105***	-3.314***	
		(1.375)	(0.964)	(0.885)	
	ER vs. US $ (pch)	0.022**	0.024***		
		(0.011)	(0.007)		
	Interest rate vs. U.S. rate (ch)			0.119***	-0.069**
				(0.044)	(0.032)
	Capital controls (ch)				0.637
					(0.500)
Observations		***1,623***	***2,523***	***2,644***	***2,388***
Pseudo R-squared		***0.204***	***0.297***	***0.126***	***0.060***

Notes: Covariates included in each regression if they are significant at the 20% level (or less) in the corresponding regressions reported in Table 2. "Ch" denotes change and "Pch" is percentage change. Constant not reported. * is significant at the 10% level, ** at the 5% level and *** at the 1% level. Includes robust standard errors.

Table 4
Summary Statistics for Different Matching Algorithms

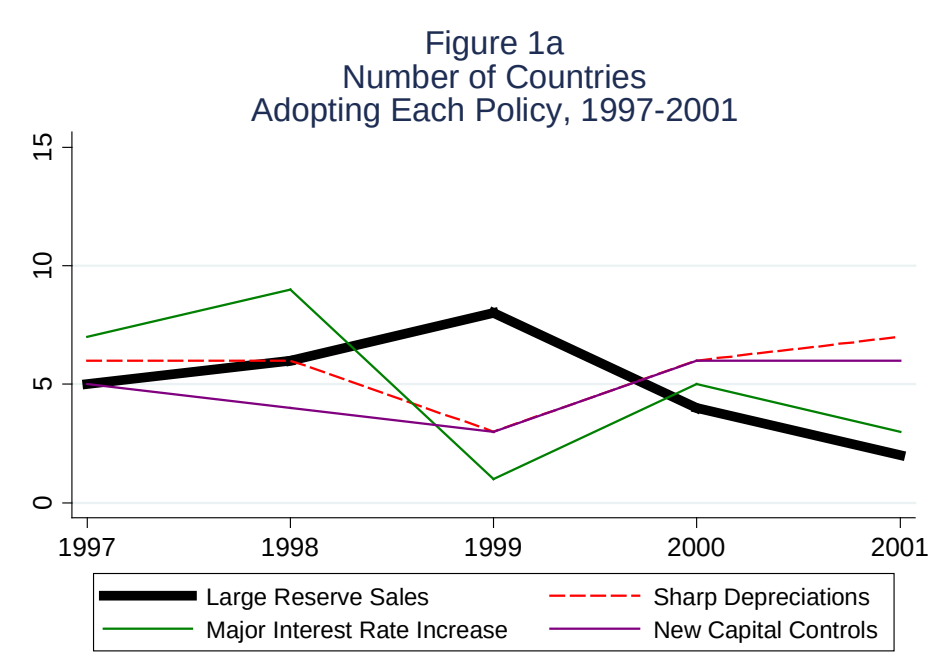
	Treatment Group	Unmatched Control Group	Control Group Based on Matching Algorithm:				
			Nearest Neighbor (no replacement)	5 Nearest Neighbors	Radius	Kernel	Local-Linear
Major Reserve Sales							
Observations (off support)[1]	29	1594	(0)	(0)	(0)	(0)	(0)
Variables failing independence test[2]			0	0	0	0	0
Mean Propensity Score[3]	467.7	423.5	530.3	467.0	483.1	477.1	478.7
Mean Absolute Bias[3]		36.9	18.6	9.2	15.3	15.3	27.9
Sharp Currency Depreciations							
Observations (off support)[1]	75	2448	(1)	(1)	(1)	(1)	(1)
Variables failing independence test[2]		12	0	0	1 (peg ER)	1 (peg ER)	0
Mean Propensity Score[3]	508.8	458.7	573.2	560.5	540.7	541.1	557.6
Mean Absolute Bias[3]		51.7	8.4	5.4	10.8	10.2	10.5
Large Interest Rate Increases							
Observations (off support)[1]	59	2585	(0)	(0)	(0)	(0)	(0)
Variables failing independence test[2]		6	0	0	1 (income per cap)	1 (income per cap)	0
Mean Propensity Score[3]	517.6	469.8	542.1	551.1	503.1	505.1	548.8
Mean Absolute Bias[3]		35.3	10.4	4.1	19.8	18.8	11.3
New Controls on Capital Outflows							
Observations (off support)[1]	40	2388	(0)	(0)	(0)	(0)	(0)
Variables failing independence test[2]		1	1 (res/gdp)	0	0	0	1 (res/gdp)
Mean Propensity Score[3]	490.9	442.7	472.6	467.5	448.8	451.0	475.7
Mean Absolute Bias[3]		20.6	14.9	6.9	16.2	15.5	10.8

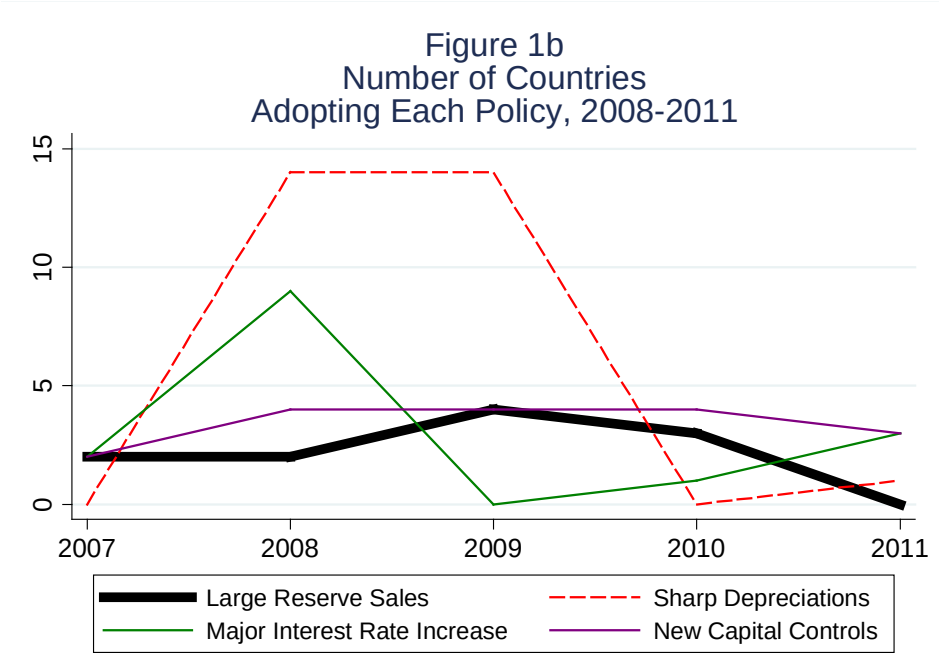
Notes: Results from matching based on regression results reported in Table 3. (1) Observations in the treatment and control groups reported in the first two columns. Observations that are off-support based on each matching method are reported in the right-hand columns in parentheses. (2) The number of variables that do not satisfy the independence test at the 5% significance level—as shown for the case of major depreciations in Table 5. (3) Reports the mean propensity score and mean absolute bias for the treatment group and control group for each matching algorithm listed at the top.

Table 5
Sharp Currency Depreciations: Means for Treatment and Control Groups using Different Matching Algorithms

	Mean: Treatment Group (μ_T)	Mean: Unmatched Control (μ_C)	t-Statistics (H0: $\mu_T = \mu_C$)	5 Nearest Neigh.		Radius		Kernel		Local-linear	
				Mean: Matched Control	t-stat	Mean: Matched Control	t-stat	Mean Matched Control	t-stat	Mean Matched Control	t-stat
Global Risk	31.703	24.786	7.06***	31.503	0.050	30.375	0.770	30.415	0.750	31.726	-0.090
ΔU.S. interest rate	-1.446	-0.455	-6.35***	-1.319	-0.320	-1.248	-0.560	-1.258	-0.530	-1.308	-0.360
Commodity prices	4.654	4.762	-2.34**	4.632	0.350	4.673	-0.360	4.670	-0.310	4.670	-0.300
1990's crisis dummy	0.533	0.465	1.16	0.581	-0.490	0.522	0.220	0.526	0.170	0.541	0.000
ΔReal GDP growth	-5.229	0.073	-5.67***	-5.267	0.210	-3.916	-0.710	-3.986	-0.660	-6.295	0.930
Current account/GDP	-0.097	0.004	-3.55***	-0.093	-0.210	-0.084	-0.410	-0.085	-0.400	-0.090	-0.280
Commodity interact.	0.413	0.293	2.24**	0.386	0.230	0.377	0.350	0.378	0.330	0.338	0.850
Income per capita	7.646	8.234	-3.85***	7.568	0.400	7.709	-0.310	7.704	-0.280	7.536	0.580
Institutions index	-0.449	-0.379	-3.20***	-0.449	0.000	-0.441	-0.250	-0.442	-0.220	-0.420	-0.840
Reserves/GDP	0.459	0.503	-0.54	0.484	-0.360	0.489	-0.390	0.485	-0.350	0.531	-0.850
Peg dummy	0.053	0.477	-7.32***	0.030	0.730	0.159	-2.08**	0.153	-1.98**	0.027	0.830
Openness	0.369	1.164	-4.52***	0.466	-0.370	0.577	-0.810	0.564	-0.760	0.529	-0.630
ΔReserves/GDP	-0.044	0.026	-4.61***	-0.043	0.210	-0.034	-0.270	-0.034	-0.240	-0.053	0.730
%ΔER/US$	15.282	2.545	8.50***	12.972	0.740	10.957	1.450	11.107	1.390	10.816	1.410
Observations	***75***	***2,488***		***74***		***74***		***74***		***74***	

Notes: Reports difference in means between treatment and control groups, with control group created based on regression results reported in Table 3 and matching performed using algorithms listed at top of table. See Appendix for detailed variable definitions. * indicates significant at the 10% level, ** at the 5% level and *** at the 1% level.



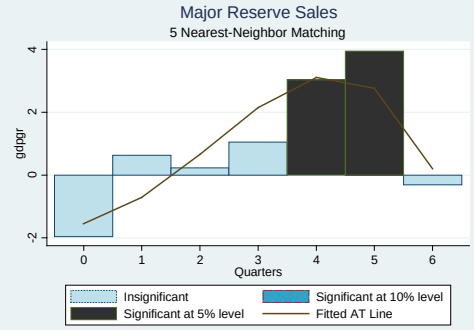


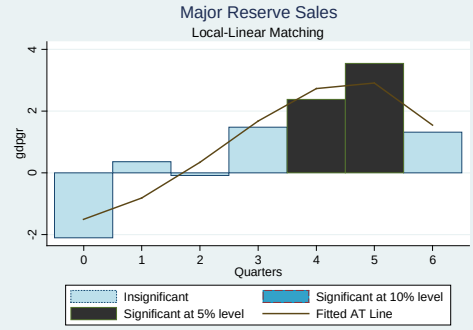
Notes: Reports annual time series of major policy changes. Large reserve sales are at least a 24% decrease in reserves between one quarter and that quarter in the previous year (and only if the reserve-to-GDP ratio is at least 10 percent). Sharp depreciations are at least a 23% deprecation in the nominal exchange rate versus the US$ between one quarter and that quarter in the previous year (and not preceded by a quarter with annual inflation of 20% or more); Major increases in interest rates are at least a 244 basis points increase between one quarter and that quarter in the previous year (and not preceded by a quarter with annual inflation of 20% or more). New capital controls are new controls on capital outflows (as described in the text). Only countries with information to calculate changes in all of the four policy responses for each year in the given period are included—so that the sample of countries is constant across time within each graph (although not across graphs).

Figure 2: Average Treatment Effects on Real GDP Growth

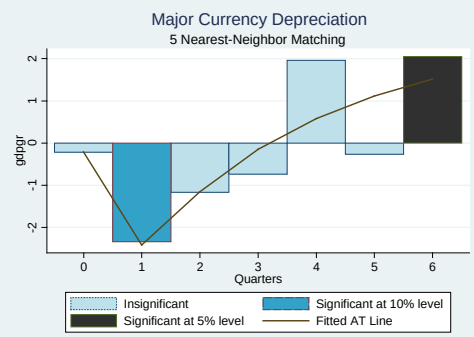
Five-Nearest Neighbors Matching | **Local-Linear Matching**

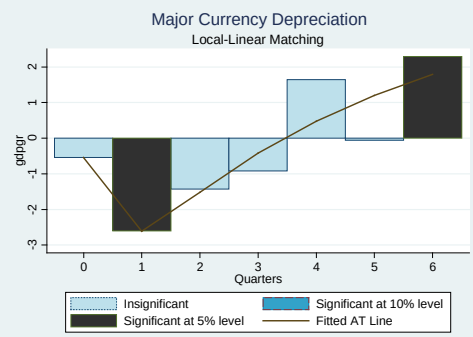
Major Reserve Sales



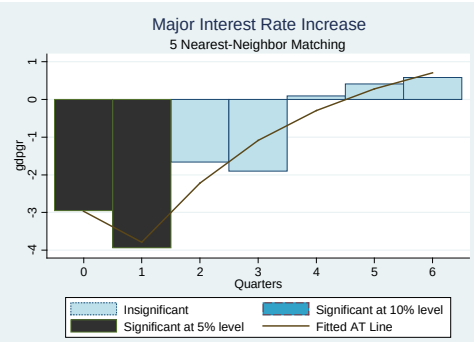


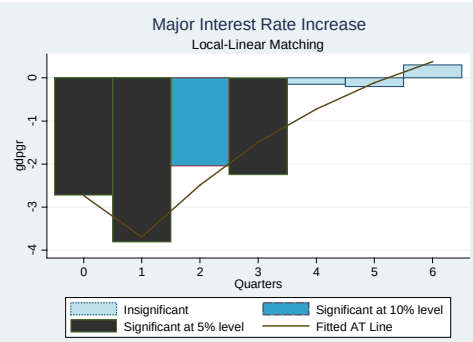
Large Currency Depreciation



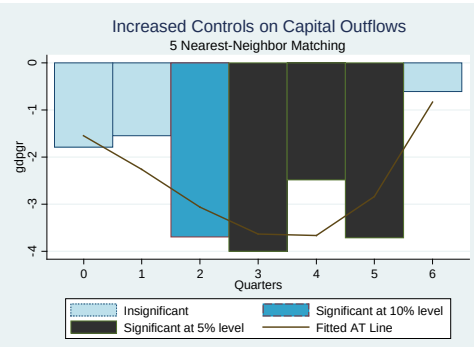


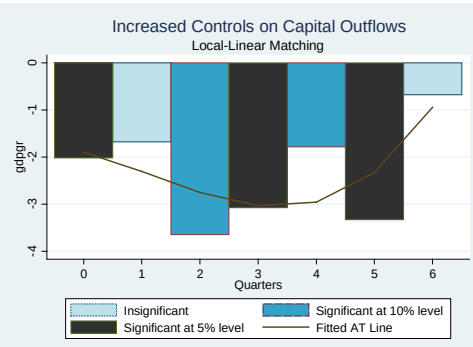
Major Interest Rate Increase





Increased Capital Controls



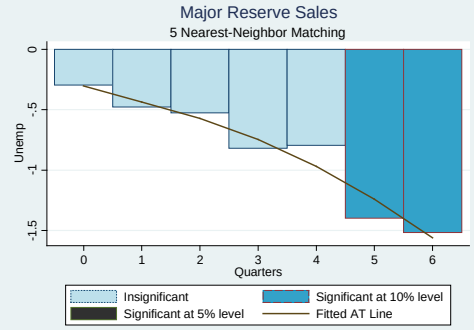


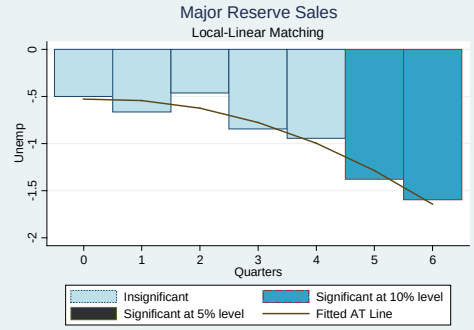
Notes: Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 3: Average Treatment Effects on Unemployment

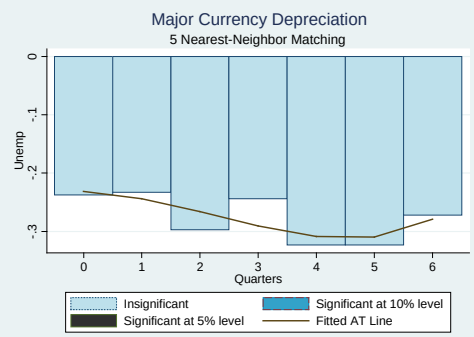
Five-Nearest Neighbors Matching | **Local-Linear Matching**

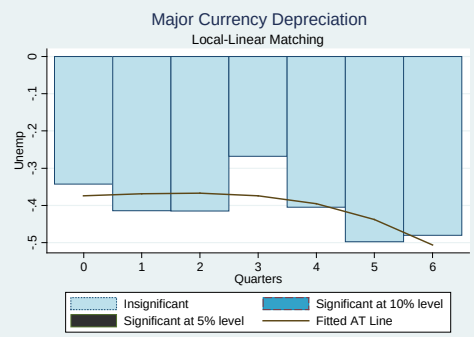
Major Reserve Sales



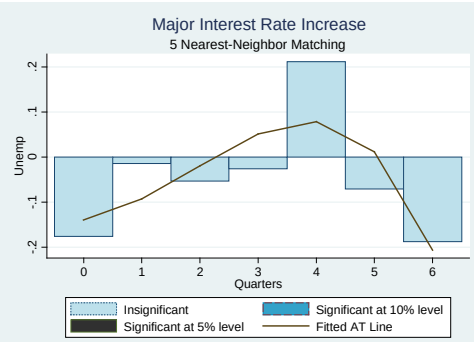


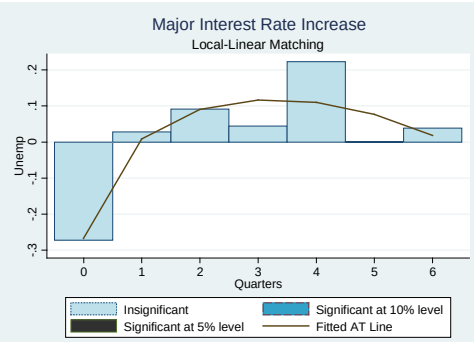
Large Currency Depreciation



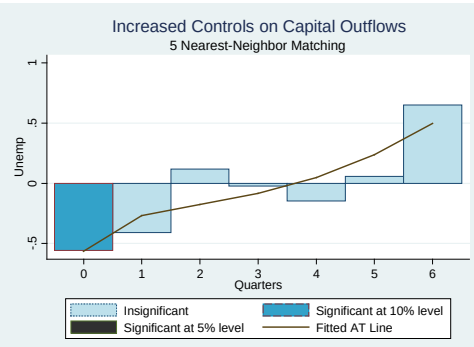


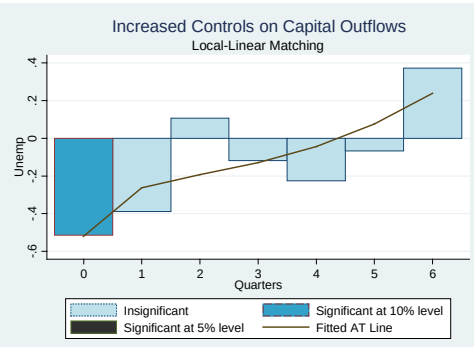
Major Interest Rate Increase





Increased Capital Controls



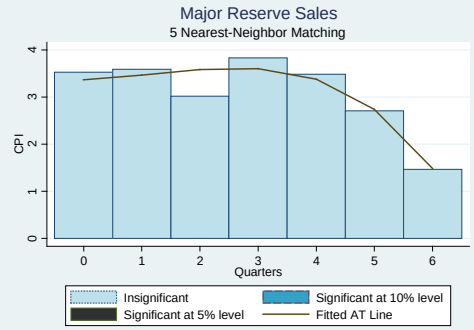


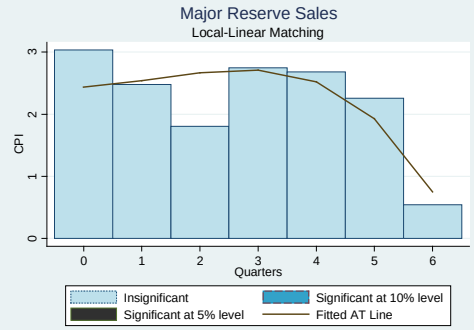
Notes: Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 4: Average Treatment Effects on Inflation (CPI)

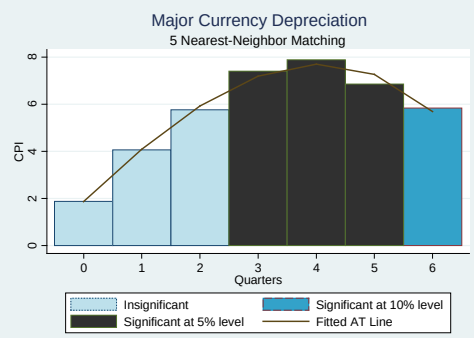
Five-Nearest Neighbors Matching | **Local-Linear Matching**

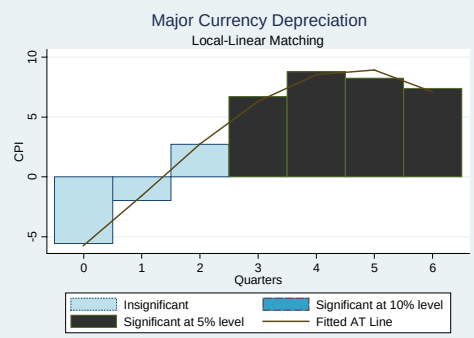
Major Reserve Sales



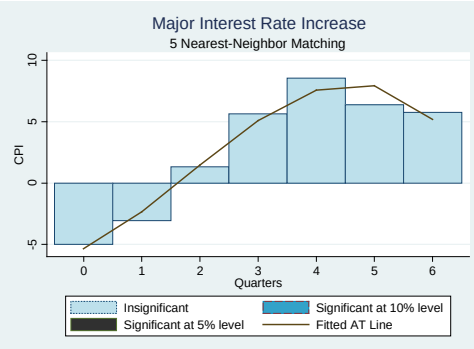


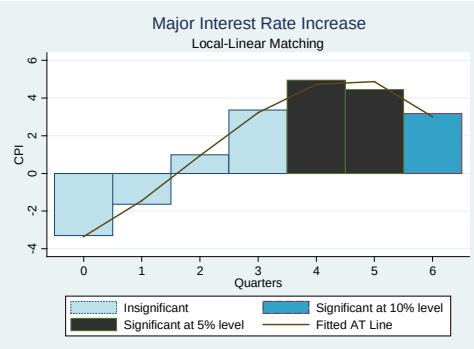
Large Currency Depreciation



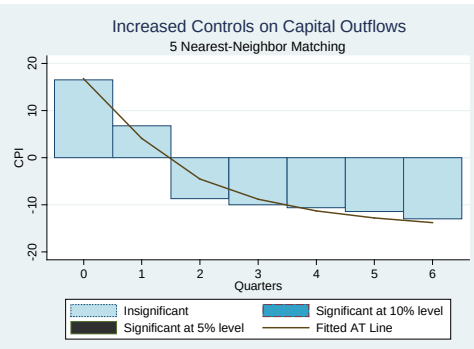


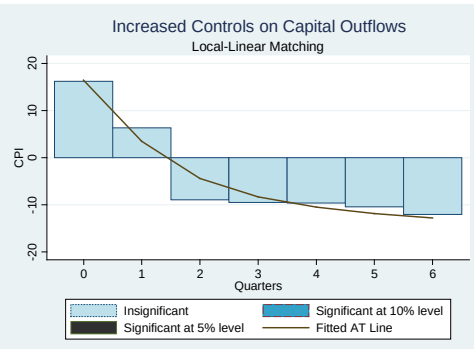
Major Interest Rate Increase





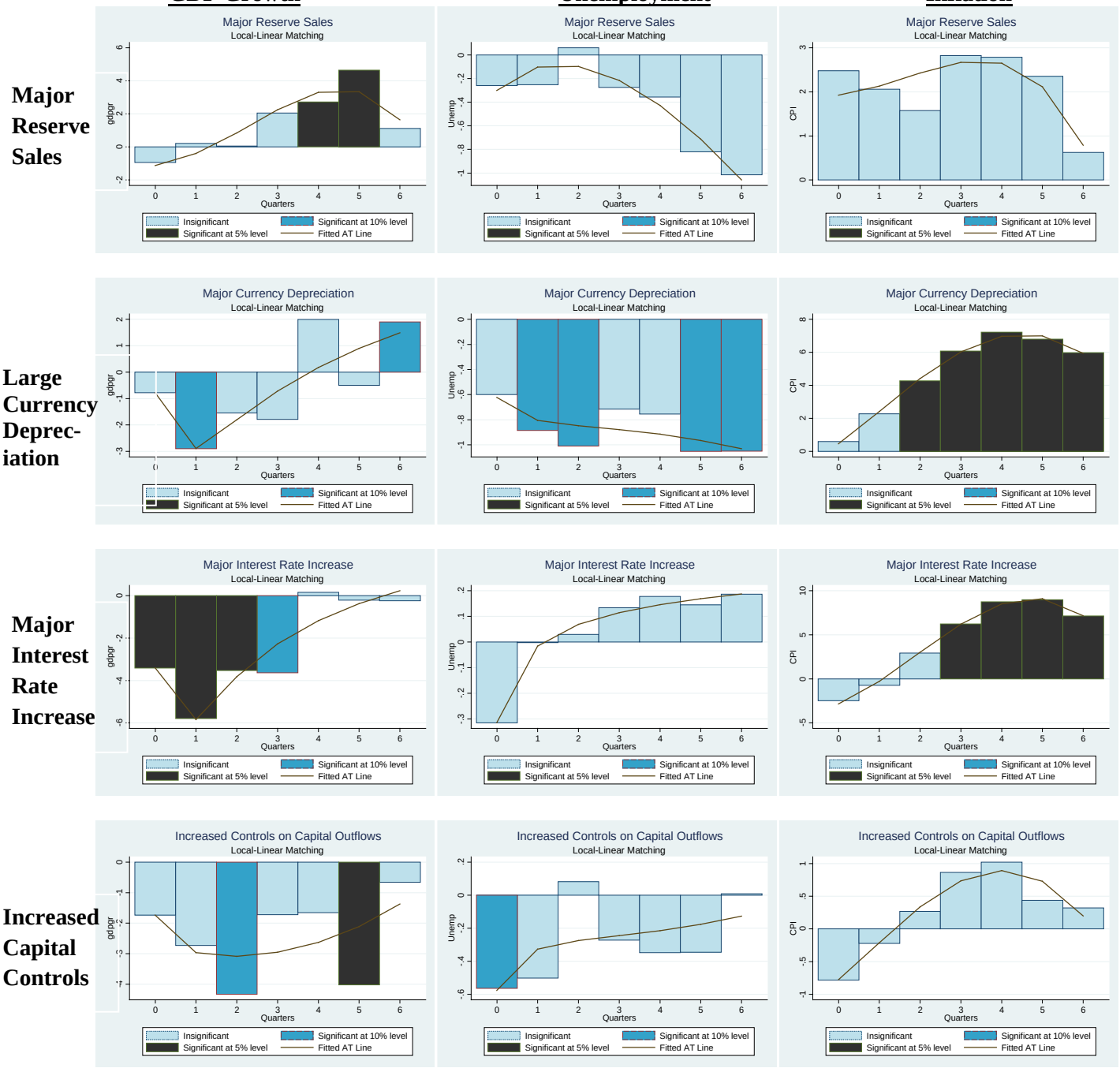
Increased Capital Controls





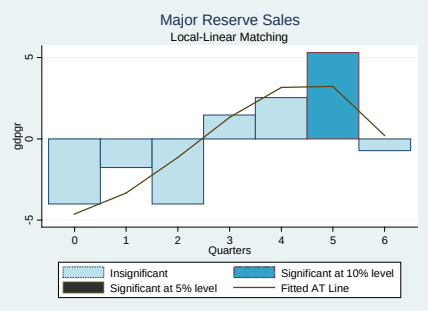
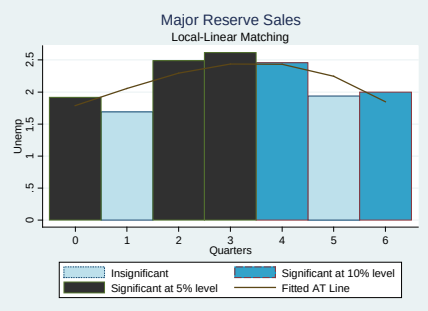
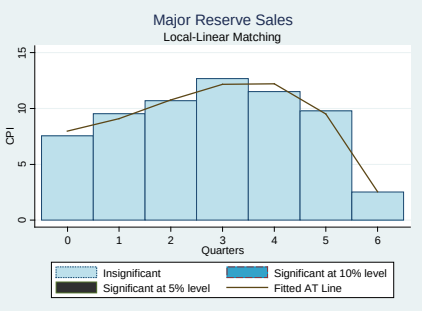
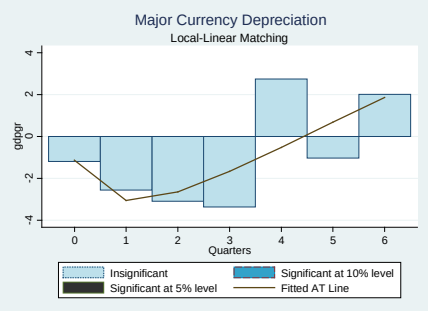
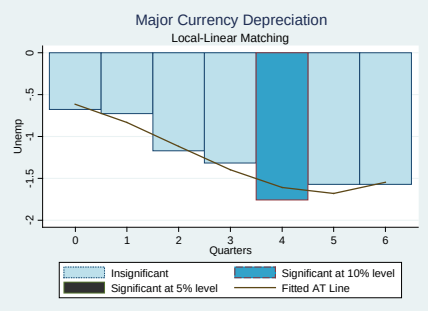
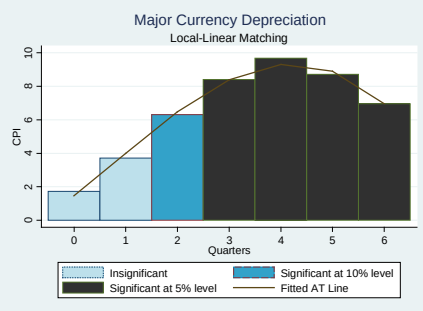
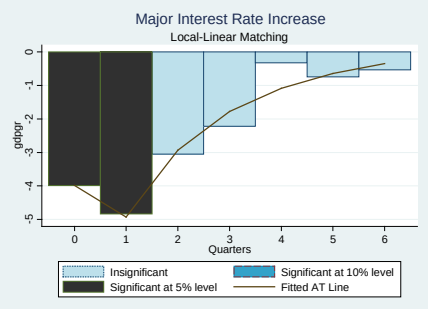
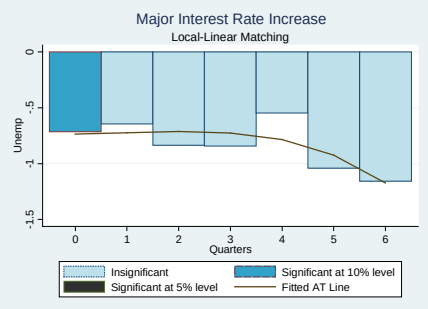
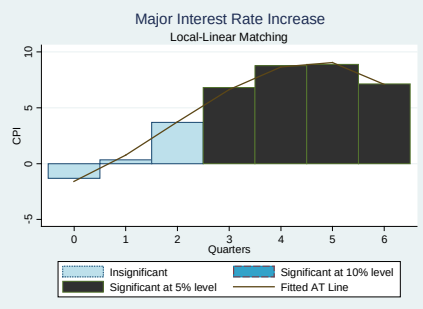
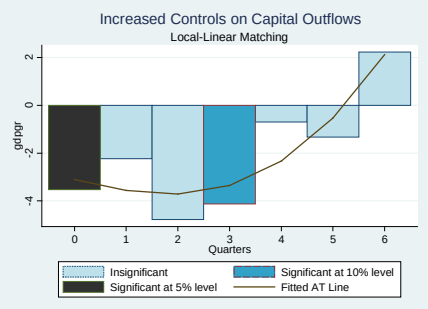
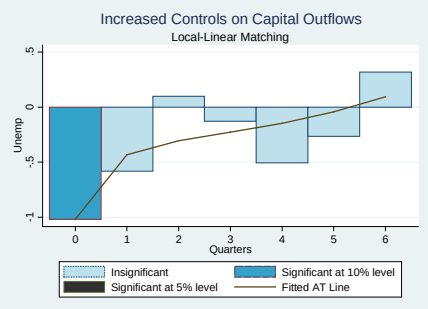
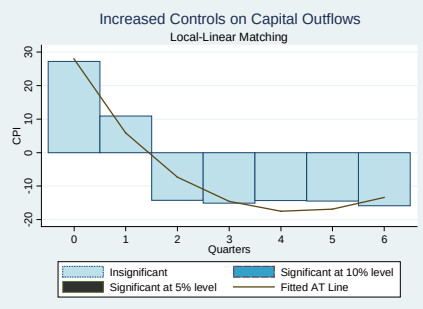
Notes: Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 5: Sensitivity Tests
Average Treatment Effects Using Full Set of Covariates



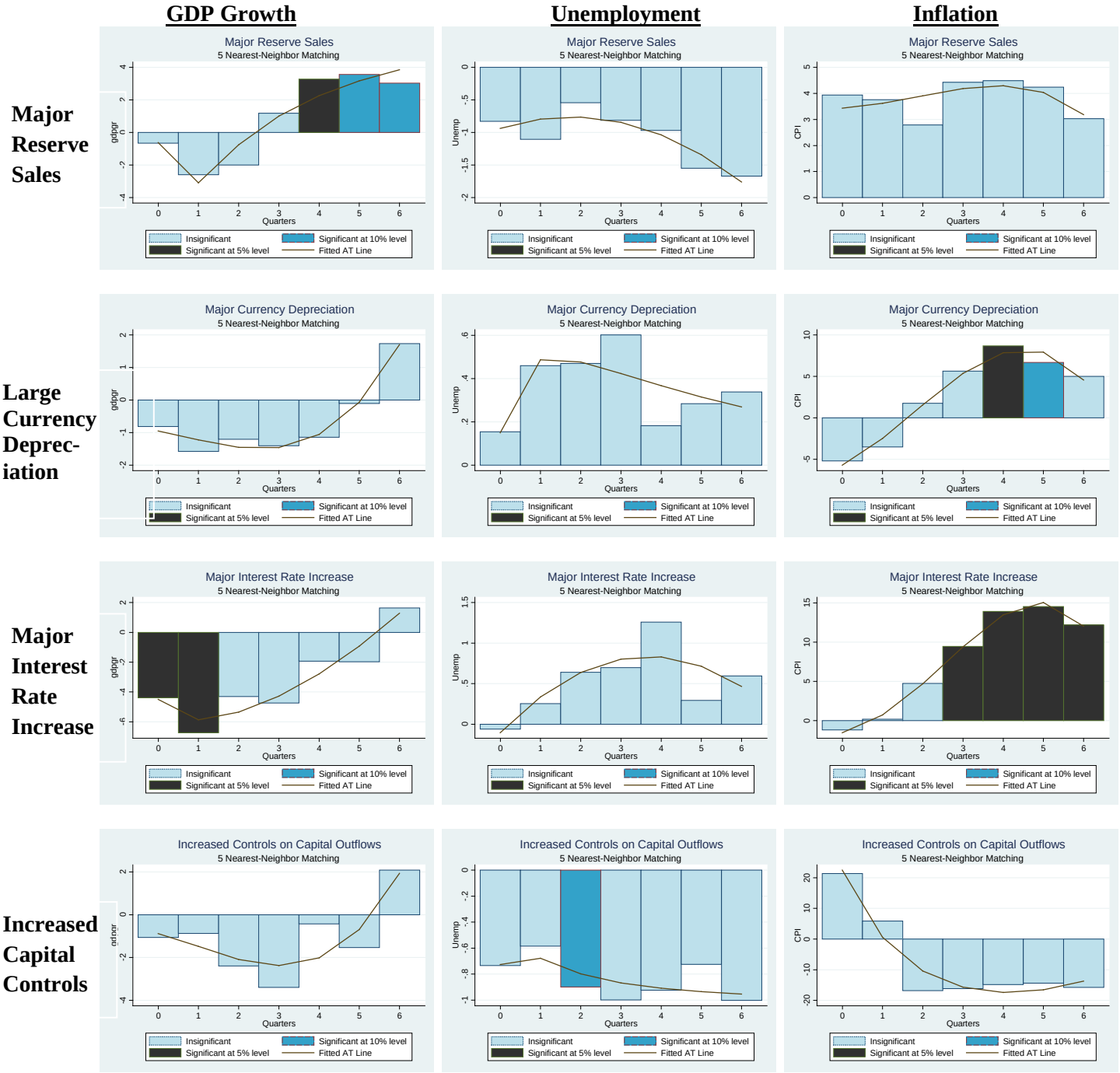
Notes: All results based on local-linear matching. Results based on first-stage logit regressions using the full set of covariates as reported in Table 2. Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 6: Sensitivity Tests
Average Treatment Effects for Emerging, Developing and non-OECD Economies

	GDP Growth	Unemployment	Inflation
Major Reserve Sales			
Large Currency Deprec-iation			
Major Interest Rate Increase			
Increased Capital Controls			

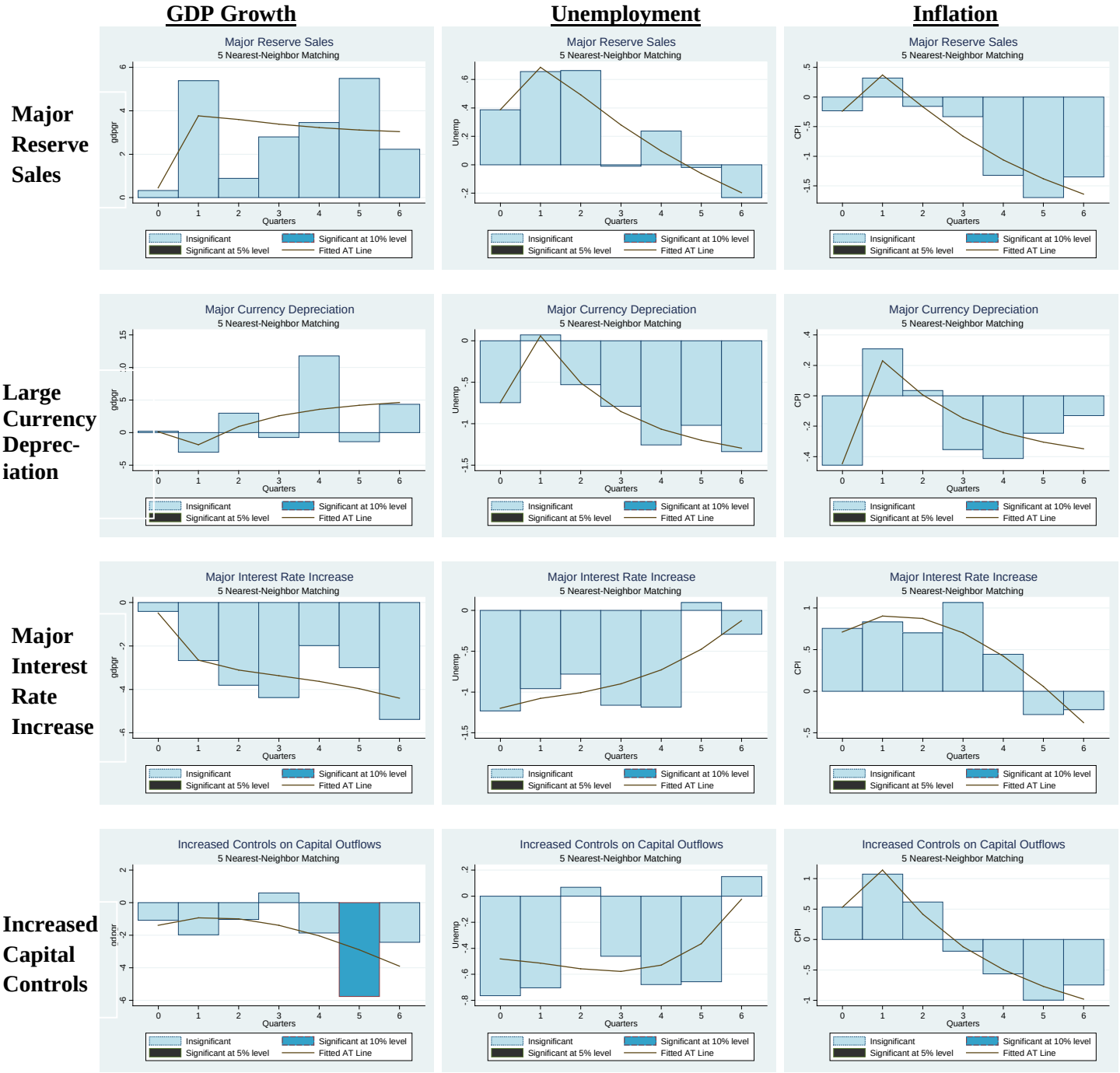
Notes: All results based on local-linear matching. Sample excludes all high-income, OECD economies as defined by the World Bank on an annual basis. Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 7: Sensitivity Tests
Average Treatment Effects for the 1997-2001 Crisis Window



Notes: All results based on five-nearest neighbor matching. Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 8: Sensitivity Tests
Average Treatment Effects for the 2007-2011 Crisis Window



Notes: All results based on five-nearest neighbor matching. Average treatment effect measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Data Appendix:

Countries in the Sample

Argentina	Ecuador	Jamaica	Namibia	Slovenia
Australia	Egypt	Japan	Netherlands	South Africa
Austria	Estonia	Jordan	New Zealand	Spain
Bahrain	Finland	Kazakhstan	Nigeria	Sri Lanka
Bangladesh	France	Kenya	Norway	Sweden
Belgium	Germany	Korea	Oman	Switzerland
Botswana	Ghana	Kuwait	Pakistan	Taiwan
Brazil	Greece	Latvia	Panama	Thailand
Bulgaria	Hong Kong	Lebanon	Peru	Trinidad & Tobago
Canada	Hungary	Lithuania	Philippines	Tunisia
Chile	Iceland	Luxembourg	Poland	Turkey
China	India	Malaysia	Portugal	Ukraine
Colombia	Indonesia	Mauritania	Qatar	United Arab Emirates
Cote d'Ivoire	Iran	Mauritius	Romania	United Kingdom
Croatia	Ireland	Mexico	Russia	United States
Czech Republic	Israel	Moldova	Singapore	Vietnam
Denmark	Italy	Morocco	Slovak Republic	Zambia

Variable Definitions and Sources

Variable	Definition	Data Source and Frequency
Reserves	International Reserves, includes IMF loans, SDRs, some SWFs, and drawn swap lines. Excludes gold for calculations of policy responses. Expressed as a share of GDP for some specifications. Millions of US$. (Q)	IMF, IFS CD-ROM, June 2013.
Nominal exchange rate	Units of local currency per dollar, end-of-period. (Q)	IMF, IFS, accessed 07/09/13.
Interest rates	The interest rate most closely related to monetary policy for each country; is the policy interest rate if available; if not available, is the short-term interest rate. (Q)	Global Insight, accessed 10/1/13.
New controls on capital outflows	A dummy equal to 1 if there is an increase in the average level of controls on capital outflows, including controls on any types of capital flows (equities debt securities, collateralized investments, and commercial credits) except FDI. (A)	Changes are based on data from Klein (2012), which is based on information in the IMF, AREARS.

Global risk	Measured by the VXO or Volatility Index calculated by the Chicago Board Options Exchange. This index measures implied volatility using prices for a range of options on the S&P 100 index.	Global Financial Data, accessed 07/11/13.
U.S. interest rate	The policy interest rate. (Q)	Global Insight, accessed 10/1/13.
Commodity price index	The Economist All-Commodity Dollar index. Measured at the end-of period and expressed in logs. (Q)	Global Financial Data, accessed 07/11/13.
Real GDP growth	Real GDP growth, measured q-o-q. (Q)	Global Insight, accessed 10/2/13.
Capital outflows, Capital inflows	Total Capital outflows or inflows in million of US$; expressed as share of GDP. Capital outflows are reported with a positive sign (unlike BOP accounting). Data for 1995-2005 uses the old balance-of-payments definitions, so there is a series break between the two crisis periods in order to avoid inconsistencies in the changes calculated over time. (Q)	IMF, BOP as of 09/13 for data from 2005-2012; Forbes and Warnock (2012) for data from 1995-2003.
Current account balance	Current account balance in millions of US$, expressed as share of GDP. (A)	IMF, WEO database, accessed 07/17/13.
Inflation	The percent change in the consumer price index relative to the previous year. Data for Hong Kong and Chile is annual. (Q)	IMF, IFS, accessed 07/09/13.
Private credit	Private credit by deposit money banks and other financial institutions as a percent of GDP. The annual data is smoothed across quarters. If data on private credit is not available, bank credit is substituted.	Beck , Demirguc-Kunt , Levine , Cihak and Feyen (*updated April 2013). Available at:* *http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:20696167~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html*
Commodity index interacted with Commodity exporter	The interaction of the log of the commodity price index (defined above) with a dummy equal to 1 if a country is a major commodity exporter. Major commodity exporters are countries for which ((food exports + fuel exports)/merchandise exports) >30%. (A)	Calculated based on export data from World Bank's, WDI, accessed 10/8/13.

Real GDP per capita	Real GDP per capita. Expressed on a quarterly basis, but if quarterly data is not available than fitted annual data is used. (Q)	Quarterly data from IMF, IFS, accessed 07/09/13; annual data from IMF WEO database, spring 2013.
Institutions index	The log of an index of institutional strength, with higher values representing stronger institutions. Index is calculated as the average of the 6 ICRG institutional variables, with each weighted by the maximum value of the variable. The variables are: legal strength, law and order, investment profile, government stability, corruption and bureaucracy quality. (Q)	Based on ICRG data compiled by the World Bank; List of variables and definitions: http://www.prsgroup.com/VariableHelp.aspx
Peg dummy	A dummy variable equal to 1 if country has an exchange rate pegged at +/- 2%. (A)	Klein and Shambaugh (2013), updating Shambaugh (2004).
Openness	The KAOPEN measure of capital account openness, which is calculated as the principal component of four binary variables from the IMF's AREARs. The four variables are: (1) capital account openness; (2) current account openness; (3) the stringency of requirements for the repatriation and/or surrender of export proceeds; and (4) the existence of multiple exchange rates for capital account transactions. Higher values indicate greater openness. (A).	Chinn and Ito (2013), with data updated as of July 2013 on their website.
Euro member dummy	A dummy equal to 1 if a country is a member of the euro area at any point in the sample.	
Change in capital controls	The sum of any increase in controls on capital inflow or outflows, except for FDI, over the previous year.	Changes are based on data from Klein (2012), which is based on information in the IMF, AREARS.
Unemployment	Unemployment rate. (Q)	IMF, IFS, accessed 07/09/13.