NBER WORKING PAPER SERIES

PICK YOUR POISON:
THE CHOICES AND CONSEQUENCES OF POLICY RESPONSES TO CRISES

Kristin J. Forbes
Michael W. Klein

Working Paper 20987
http://www.nber.org/papers/w20987

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2015

Thanks to Rajeev Dehejia, Pierre-Olivier Gourinchas, Ayhan Kose, Jeffrey Sachs, anonymous reviewers at the IMF, and conference participants at the IMF for helpful comments. Patrick O'Halloran provided excellent research assistance. We gratefully acknowledge financial support from the IMF for this paper which was prepared for the IMF's Annual Research Conference in November 2013. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Pick Your Poison: The Choices and Consequences of Policy Responses to Crises
Kristin J. Forbes and Michael W. Klein
NBER Working Paper No. 20987
February 2015
JEL No. F41

ABSTRACT

Countries choose different strategies when responding to crises. An important challenge in assessing the impact of these policies is selection bias with respect to relatively time-invariant country characteristics, as well as time-varying values of outcome variables and other policy choices. This paper addresses this challenge 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 none of these policies yield significant improvements in growth, unemployment, and inflation. Instead, a large increase in interest rates and new capital controls are estimated to cause a significant decline in GDP growth. Sharp currency depreciations may raise GDP growth over time, but only with a lagged effect and after an initial contraction.

Kristin J. Forbes
MIT Sloan School of Management
100 Main Street, E62-416
Cambridge, MA 02142
and NBER
kjforbes@mit.edu

Michael W. Klein
Fletcher School
Tufts University
Medford, MA 02155
and NBER
michael.klein@tufts.edu

I. Introduction

In the mid- and late-1990s, a series of emerging markets faced sudden stops of capital inflows. The standard policy prescription included a sharp increase in interest rates in an effort to attract capital.[1] Countries that followed this prescription generally experienced severe recessions and increased unemployment, and while the medicine was not necessarily worse than the alternative, there were adverse side effects. 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 late 2007, many countries responded in ways other than through substantial increases in interest rates. But these other policies also had their own painful side effects. Thus, these crisis episodes suggest that countries must "pick their poison," since virtually every policy response to a crisis involves costs and 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, as well as the turmoil during the Global Financial Crisis [GFC]).[2] We analyze four policy responses: major reserve sales, large currency depreciations, substantial interest rate increases, and new controls on capital outflows. These major policy shifts do not typically occur during stable times and are nimble enough to be quickly adopted in response to sudden stops in capital flows.[3] We document patterns in the use of these four policies during the two crisis periods. We also estimate the determinants of which countries use these policies and at what times. These estimates are then

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

[2] Although the crises in the period from 1997 to 2001 were often more regional or country-specific in nature than during the later period (except perhaps during the fall of 2008 after the collapse of LTCM), both crisis windows are periods of volatility in global capital flows punctuated by "sudden stops." Claessens *et al.* (2014) presents a wide-ranging analysis and overview of financial crises, and the chapter by Claessens and Kose (2014) provides a detailed discussion of the different types of crises and patterns over time. Gourinchas and Obstfeld (2012) present an analysis of the causes and consequences of crises from 1973 to 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. We also do not analyze fiscal policy responses as these are discussed in detail in Chari and Henry(2015), as part of the same conference for which this paper was prepared.

used, in a manner described below, in our analysis of the impact of these policies on three key outcome variables: GDP growth, inflation, and unemployment.

Estimates of the impact of these policies cannot simply compare outcomes in the wake of a particular policy change to outcomes in the absence of that policy due to "selection bias." Selection bias can occur because of three potentially important differences between countries that do and do not undertake these policies: longstanding underlying economic differences, differences in outcome variables before and at the time when the policy is enacted, and differences in the use of other policies. These differences could potentially affect outcome variables and, therefore, confound estimates of the effects of policies on outcomes. For example, estimates in this paper show that countries are more likely to have large depreciations during the two crisis periods if they were less open, if recent GDP growth was slower, and if they did not recently sell reserves, all of which could potentially affect GDP growth, unemployment, or inflation independent of whether a large depreciation occurred. Although previous studies note these issues when considering the impact of crisis responses, there is typically a limited ability to address them. This is because of the difficulties controlling for all differences across countries using a limited set of observable statistics and because of the difficulty in finding effective instrumental variables.

In this paper, we address selection bias between countries that undertake these policy responses at particular times and those that do not 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 rarely used 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, based on the estimated likelihood that a policy is undertaken, that selects relevant observations to compare countries which use a certain policy at a certain time (a treatment group) with countries that do not (a control group).

In the first stage of this propensity-score methodology, we estimate logit models of the probability that a country undertakes each of the four policy responses as a function of global

[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 Hutchison (2006); Das and Bergstrom (2012); and Forbes, Fratzscher, and Straub (2015).

variables, domestic vulnerabilities, country characteristics, and other recent policy decisions. The models have 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 between countries that use these policies and those that do not. The estimates also show that, in some cases, the use of one policy may influence the likelihood of the subsequent use of another one.

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 select control groups using five different matching algorithms, matching controls for time-varying attributes, such as the lagged values of other policies and outcome variables, as well as for relatively stable country characteristics. A key result is provided in the set of *independence tests* that show that the matched set of countries do not exhibit the selection bias on observables that would be present if the entire sample served as a control group in a simple comparison of outcomes. The matched sets of countries are used to estimate the average treatment effect on the treated (ATT), the statistic that shows the effect of the four policy responses on key outcome variables while controlling for selection bias in the quarter that the policy is undertaken as well as in the subsequent six quarters.[6]

The estimates suggest that none of these four policies yield significant improvements in growth, unemployment and inflation; instead, several yield negative consequences. 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 with no benefit in terms of reducing inflation or unemployment. Large currency depreciations appear to be most likely to generate significant increases in GDP growth relative to the counterfactual, although the positive effects do not occur for at least six quarters, are often preceded by a contraction in GDP growth, and are often only marginally significant. A series of robustness tests generally confirms these results. Extensions suggest that policies may have larger effects in non-OECD economies, but these effects are often less precisely estimated. Extensions also suggest that during the Global Financial Crisis, depreciations may have yielded a reduction in unemployment (relative

[6] Propensity score matching does not directly control for endogeneity. This is a potential concern for the ATT statistic of the effect of the policy on the outcome variable in the quarter in which the policy is enacted. But the ATT statistics in the six subsequent quarters consider the effect of the policy on subsequent outcomes and do not have the same potential problem with endogeneity, especially given the exclusion restriction described below.

to the counterfactual), and the effects of interest rate increases differed than during the 1990s—but all of these extensions are only suggestive as they are based on small samples that limit the effectiveness of the estimation technique.

These results suggest that during a crisis and contraction in capital flows, there is no miracle cure of a policy response. Neither a major reserves sale, sharp currency depreciation, large increase in the interest rate, nor new 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. Section V then 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 to 2001 and 2007 to 2011. To simplify terminology, we will refer to the earlier period as the 1990s crises and the later window as the 2000s 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 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 leads 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. More specifically, using these thresholds, major reserve sales are defined as quarters during which international reserves (relative to GDP) fall by at least 13.4 percentage points, compared to the previous year.[8] Large depreciations are defined as quarters in which there is at least a 22.6 percent depreciation over the previous year in the country's exchange rate versus the U.S. dollar. 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 200 basis points over the past year.[9] To avoid episodes when large depreciations or interest rate hikes primarily reflect a response to inflation, we exclude country-quarters in which the annual rate of inflation is more than 10 percent. We exclude countries that are currently members of the euro zone from the sample of countries that can have a large change in reserves, exchange rates, or interest rates due to the constraints on these countries as members of this currency union.

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, 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 new controls on capital outflows were added,

[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 in the final analysis if data on all four of the policy responses is available. Details on sources and definitions for the data are in the Appendix.

[8] There are a variety of ways to measure reserve loss in order to obtain an indicator of a "major" change; for example, in an earlier version of this paper we considered a threshold change in reserves combined with a requirement that reserves-to-GDP meet or exceed a minimum level (there were 52 instances of large reserve losses using that method, versus 62 with our current method). We have chosen our current approach because GDP is an appropriate variable for directly scaling reserves. But, as with any approach, there are potential concerns; for example, the percentage change in reserves-to-GDP could be large and negative (and count as a major reserve loss) if GDP growth is especially large and reserves are not repleted (this is the case in about 8 of the 62 instances of major reserve loss). Also, we would not count as a major reserve loss an instance where there is a large drop in reserves in a time when GDP growth has a large negative value (although, given that we have more instances of major reserve loss with this method than with our former approach, this might not be an important concern).

[9] 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.

based on the IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions.*[10] This capital control data is only available on an annual basis, in contrast to the other data which are available on a quarterly basis.[11] Using this measure, controls on capital outflows were added 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.

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 we do not score a policy change until at least four quarters after the occurrence of a policy change of the same type. 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 of the other 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 eight episodes in which a substantial increase in interest rates occurred contemporaneously, or within one quarter of, a large depreciation. The numbers in the boxes that span more than two cells indicate triplets or quadruplets of policies that occurred contemporaneously or within one quarter; for example, there was one episode in the 1997-2001 period when a country had a large depreciation, a substantial increase in interest rates, and the addition of capital controls all within one quarter of each other (Indonesia in the third quarter of 1997) and one other episode in which a country undertook all four policies (Brazil in the first quarter of 1999). The last column of each table shows the total number of each of the large policy changes.

One important point that emerges from these tables is that the majority of the policy changes occur in isolation (represented by the diagonal elements) rather than occurring

[10] The data on capital controls is from an expanded version of the data set in 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.

[11] We count a control as being put in place in the third quarter of the year.

simultaneously, or within one quarter of, different policies. Countries chose only one of these policies 54 percent of the time (53 of 98 major policy changes) in the 1997-2001 period and 64 percent of the time (64 of 100 changes) in the 2007-2011 period. Another important point is the relative incidence of the different policy responses across the two periods. Major increases in interest rates occurred more frequently during the earlier crisis period window than during the later period (36% versus 17% of the total policy responses, respectively). In contrast, major reserve sales were used more frequently in the latter period (44% of policy responses) than in the earlier period (18% of policy responses). Major exchange rate depreciation and new capital controls occurred in roughly the same percentages of total policy responses across the two periods.

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.[12] Figure 1a shows that, during the earlier crisis window, interest rates were most often increased sharply in 1997, 1998, and 2001, and used more than any of the other policies during these three years. The incidence of major depreciations increased fairly steadily from 1997 through 2001. Figure 1b shows that large reserves sales was by far the most common policy response during 2008, although the use of all three policies were at or close to their peak incidence in the same year. Major policy responses were used less often in 2009 and 2010, but then 2011 shows an increased use of reserve sales, interest rate increases, and sharp depreciations.

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 new capital controls. But what determines which of these policies, if any, a country chooses, and when? To answer these questions, this section estimates the probabilities of each of these policies being employed during the crisis periods in the late 1990s and late 2000s. These estimates not only contribute to our understanding of why countries use these policies, but also are the basis for the

[12] Only countries with information for all four policy responses for each year in the given crisis period are included in these figures, 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. In these graphs, the episodes are converted from a quarterly frequency to an annual frequency.

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—but as discussed in the sensitivity tests in Section V.D, the key results in this paper are robust to the choice of different variables. Therefore, for our base case we begin by selecting a wide set of variables 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; and Claessens, Dell'Ariccia, Igan, and Laeven, 2010), sudden stops in capital flows (e.g., Forbes and Warnock, 2012; Calvo, 1998; and Calvo, Izquierdo, and Mejía, 2008), and the use of capital controls (e.g., Forbes, Fratzscher, and Straub, 2015, and Aizenman and Pasricha, 2013). We also draw on the framework of explanatory variables used in other work using propensity-score matching at higher frequencies than annual data (e.g. Forbes, Fratzscher and Straub, 2015). The Appendix includes details on all of these variables—including definitions, sources, and coverage.

The full set of 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.[13] This category also includes a dummy variable equal to 1 for observations in the 1997QI-2001QIV period in order to distinguish the earlier crisis period from the latter period. 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 as a percent of GDP, changes in CPI inflation, the change in private credit (relative to GDP), and a commodity exporter dummy interacted with the commodity price.[14] A third category represents domestic characteristics that vary more in the cross section

[13] 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.

[14] 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

than in the time series for any particular country. This set includes the logarithm of income per capita, an indicator of institutional quality[15], capital-account openness as measured by the Chinn-Ito KAOPEN variable, the level of reserves relative to GDP, whether the exchange rate is pegged[16], 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 over the previous year in the four policies on which we are focusing: changes in reserves (as a share of GDP), changes in the country's policy interest rate relative to the U.S. rate, the percent change in the nominal exchange rate versus the U.S. dollar, and the addition of any new controls on capital outflows. All these covariates enter the regression lagged by one quarter, or by one year if only annual data is available.[17]

We first estimate a base-case logit model for each of the four policies, using a specification that can be represented as

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

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

Given this wide set of variables in the initial regression, it is natural to find that many of the coefficients are not significant. In the first-stage regression used for propensity-score matching, however, including variables that are irrelevant (in the sense that they do not influence the policy choice) will increase the variance of the estimates in the second stage and can make it

---

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.

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

[16] From Klein and Shambaugh (2013).

[17] Details on the definitions and sources for all of these variables are available in the appendix.

more difficult to find appropriate matches.[18] Therefore, for our base case we estimate regressions which only include the explanatory variables that are significant at least at the 20% level in the initial estimates. The resulting estimates are presented in Table 2.

The estimates in Table 2 show that the logit model best explains large currency depreciations (with a pseudo-R^2 of 0.39) and major reserve sales (with a pseudo-R^2 of 0.41). The logit model is least successful in predicting new controls on capital outflows (with a pseudo-R^2 of 0.08).[19] The coefficients on a number of the covariates representing global variables are significant, indicating that global variables can influence the choice and timing of the four major policy responses. For example, an increase in global risk is associated with a higher likelihood of a large depreciation in the subsequent quarter, an increase in commodity prices with a higher likelihood of a major interest rate increase, and, in a less intuitive result, an increase in U.S. interest rates are associated with a lower likelihood of large reserve sales. The coefficient on the dummy variable for the 1990s crisis window shows that major interest rate increases are more likely in the earlier crisis period than in the later one, a finding consistent with the unconditional statistics presented in Tables 1a and 1b. Domestic vulnerabilities can also significantly affect the choice and timing of the policy responses. For example, a country is more likely to have a large currency depreciation in the quarter after slower growth and is more likely to increase interest rates sharply after an increase in capital outflows.

The estimates in Table 2 also show that a number of country characteristics, such as reserve holdings, institutional quality, and the exchange rate regime, are significantly associated with most policies. For example, countries with higher reserves (as a share of GDP) are more likely to have large reserve sales and less likely to have large currency depreciations. Countries that are more open are less likely to have large reserves sales, major currency depreciations, or large interest rate increases. Finally, previous policy changes affect the likelihood of major policy changes in the current period. For example, a decrease in reserves over the previous year is associated with a lower likelihood of a subsequent major reserve sale, current depreciation, or interest rate increase. In contrast, a depreciation against the dollar over the previous year is

---

[18] See Heinrich, Maffioli, and Vázquez (2010). Including irrelevant variables will make it more difficult to satisfy the common-support condition, as explained below. The sensitivity analysis also reports results when the full set of covariates is included in the first-stage regression.

[19] Even though the R^2 statistic is low in the probit explaining capital controls, statistics presented in Table 3 below indicate that no variables fail the independence. If, in fact, the addition of capital controls is not correlated with other variables, then sample selection issues are less pronounced.

associated with a greater likelihood of a subsequent major depreciation and reduced likelihood of large reserve sale.

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, 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. Robustness tests which maintain a reasonable sample size (including the number of treated observations) can yield different estimates for the first-stage regressions based on which variables are included in the regression, but do not alter the key results reported below on the impact of the four policy responses.

The estimates presented in Table 2 suggest that selection bias is not just a hypothetical concern 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 that do not. The estimates show, for example, that countries with lower levels of reserves and with floating exchange rates in the previous period are more likely to undergo sharp currency depreciations. Similarly, differences across time also affect the likelihood of major policy changes; for example, sharp currency depreciations are more likely at times when global risk is high and interest rate increases are more likely after commodity prices increase. These estimates indicate the presence of selection bias (an assessment which is confirmed by the more rigorous independence tests presented below), and therefore the importance of addressing this issue through techniques such as propensity score matching.

IV. Propensity-Score Methodology and Matching Results

A. *An Overview of Propensity-Score Methodology*

Countries which undertake a particular policy at a particular time may differ from countries which do not undertake that policy in ways that affect outcomes beyond the effects of the policy itself. This raises the possibility that an analysis comparing the outcomes of countries depending upon whether or not they implement a policy may capture selection bias as well as the impact of that policy. One means of addressing this issue is through the use of propensity score matching. This technique represents an effort to isolate the effects of policies from other, confounding effects correlated with the likelihood of their implementation. This section discusses propensity score matching, which has only recently been used in macroeconomics, monetary economics, and international economics, although it has been a staple in other fields such as labor economics for several decades.[20]

To illustrate the issue of selection bias, we define the adoption of the treatment or policy (such as a major increase in interest rates) by the i^{th} country as $D_i = 1$, and the absence of this action as $D_i = 0$. The outcome variable (such as the change in GDP) 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 $E[Y_{0,j}|D_j=0]$. The observed difference between the outcomes for these two groups is

$$E[Y_{1,i}|D_i=1] - E[Y_{0,i}|D_i=0] = E[Y_{1,i} - Y_{0,i}|D_i=1] + \{E[Y_{0,i}|D_i=1] - E[Y_{0,i}|D_i=0]\}.^{21} \qquad (2)$$

We are interested in $E[Y_{1,i} - Y_{0,i}|D_i=1]$, which is the effect of the policy relative to the outcome the members of the treated group would have had in the absence of that policy. As noted in the introduction, this is called the average effect of the treatment on the treated, or ATT. The observed difference in outcomes between the two groups, however, consists of both the ATT and $\{E[Y_{0,i}|D_i=1] - E[Y_{0,i}|D_i=0]\}$, which represents a selection bias. Selection bias occurs if the treatment is not randomly assigned and if there are differences in outcomes solely because of pre-treatment differences between the treated and control groups.

The effect of sampling bias could be easily minimized if there were a large set of countries that differed along only one or two discrete and observable dimensions with respect to the likelihood of undertaking a policy.[22] In this case, countries could be readily apportioned to a

[20] Propensity-score matching is discussed in Dehejia and Wahba (2002), Angrist and Pischke (2008, chapter 3) and Heinrich, Maffioli, and Vazquez (2010). Recent research in macroeconomics that has used this technique includes analyses of monetary policy (Angrist and Kuersteiner (2011), Angrist, Jordá, and Kuersteiner (2013) and Ehrmann and Fratzscher (2006)), the effect of openness on growth (Das and Bergstrom (2012)), financial liberalization (Das and Bergstrom (2012), and Levchenko, Rancière, and Thoenig (2009)), foreign ownership of firms (Chari, Chen, and Dominguez (2011)), the response of economies to crises (Glick, Guo, and Hutchison (2006)), fiscal policy (Jordà and Taylor (2013)) and the effects of capital controls and macroprudential measures (Forbes, Fratzscher and Straub (2015)).

[21] Angrist and Pischke (2008) show the full derivation of these equations. The observable outcomes are $Y_i = Y_{0,i} + (Y_{1,i} - Y_{0,i})D_i$. The expected values conditional on D_i are $E[Y_i|D_i=1] = E[Y_{1,i}|D_i=1]$, $E[Y_i|D_i=0] = E[Y_{0,i}|D_i=0]$, and $E[(Y_{1,i} - Y_{0,i})D_i|D_i=0] = 0$. Thus, $E[Y_{1,i}|D_i=1] - E[Y_{0,j}|D_j=0] = E[Y_{1,i} - Y_{0,i}\ |D_i=1] + E[Y_{0,i}|D_i=1] - E[Y_{0,i}|D_i=0]$.

[22] In this discussion, to illustrate propensity score matching, we focus on differences across countries and, for example, the vector X_i for the i^{th} country. Our empirical analysis focuses on differences across countries and across

small number of "cells" reflecting all differences along these dimensions. It would be straightforward to calculate the differences between the treated and the untreated in each cell (providing there are enough observations in each cell), and take a weighted average of those differences in order to estimate the effect of different treatments. In practice, however, there are many, multidimensional differences across countries and across time, and it is impossible to simply match treated and control observations with identical macroeconomic characteristics.

It may be possible, however, to match treated countries to control countries based on a set of observable country characteristics, represented by the vector X_i for the i^{th} country. If this matching takes into account the differences in the treated and untreated groups that affect outcomes, then the sampling bias (or at least any bias that is captured in the vector X_i) disappears, and $E[Y_{0,i}|X_i,D_i=1] - E[Y_{0,i}|X_i,D_i=0] = 0$. This could still leave a multidimensional problem. This problem can be resolved, however, because 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 (Rosenbaum and Rubin, 1985). This single propensity score reduces the number of dimensions over which observations must be matched.[23]

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

$$p(X_i) = \Pr[D_i=1|X_i] \; . \tag{3}$$

In the context of our model, propensity scores are the likelihoods that a country undertakes a major increase in the policy interest rate, currency depreciation, or reserve sale, or introduces new controls on capital outflows. The propensity scores, based on the function $p(X)$, can be generated using logit regressions like those reported in Section III.[24]

---

time, so that the appropriate vector is $X_{i,j}$, representing variables' values for the country i in year j. Thus, rather than pair a particular country with one or more other countries identified through propensity score matching, we will pair a particular country-year observation with one or more other country-year observations.

[23] Rubin and Thomas (1992) show that it is possible to estimate these propensity scores based on the vector of observable characteristics.

[24] In our analysis, we consider the change in a variable (e.g. GDP growth) before and after the treatment. For example, if $G_{w,I,z}$ = GDP growth for the i^{th} observation which is either treated ($w = 1$) or untreated ($w=0$), either after ($z = After$) or before ($z = Before$) the policy, we consider $E[G_{1,i,After} - G_{1,i,Before}|D_i=1] - E[G_{0,i,After} - G_{0,i,Before}\ |D_i=0]$. Following the algebra in Equation (2), and using matching to generate a probability function $p(X)$ in order to select the untreated observations to compare with the treated observations, we can show that

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) with similar propensity scores. We consider five matching algorithms: nearest-neighbor without replacement, five-nearest neighbors with replacement, radius with caliper, kernel, and local-linear matching. Each of these has advantages and disadvantages, and some techniques may not satisfy key requirements based on the specific data set and problem analyzed.[25] Using different matching methodologies is important to be able to assess which technique performs best as well as to assess the robustness of the results. This is especially important because even if several techniques all satisfy the necessary tests, the significance of key results can still depend on the matching method used and corresponding construction of the control group.

There are several criteria and tests that can be used to evaluate the performance of matching algorithms. To begin, it is useful to compare the differences in the mean propensity scores between the treated observations and both the untreated observations and the control group observations. By construction, the difference in mean propensity scores should be larger between the treated and untreated groups than for the treated and control groups, and comparing mean propensity scores helps gauge the relative performance of various matching methods and extent of selection bias before the matching. Another check is whether the propensity scores of treated observations are between that of the minimum and maximum propensity scores for the

---

$$\text{ATT} = \text{E}[G_{1,i,After}|D_i\text{=1}] - \text{E}[G_{0,i,After}|p(X),D_i\text{=1}] - \{\text{E}[G_{1,i,Before}|D_i\text{=1}] - \text{E}[G_{0,i,Before}\ |p(X),D_i\text{=1}]\}.$$

The term in curly brackets does not equal 0, even if GDP growth is one of the elements used in the *p(X)* function, because *p(X)* is matching country-year observations based on the likelihood of treatment, not the commonality of the outcome variable before the treatment (and, of course, the selection is for the same country-year observations for both the Before and After periods). Thus, in our analysis, the ATT consists of differences between GDP growth between the treated and control observations for both the periods before and after the treatment.

[25] All these methods use the propensity score of each treated observation and propensity scores of untreated (control) observations. Nearest-neighbor selects the single control observation with the closest propensity score, and "without replacement" refers to the fact that any control observation can be matched to only one treated observation. Five-nearest neighbors uses more observations from the control group and allows replacement. The radius method includes all nearest neighbors within a maximum radius (referred to as the caliper). The kernel and local-linear matching algorithms each calculate a weighted average of all observations in the control group and assign higher weights (which differ between these two methods) to control observations closer to the treated observation. Nearest neighbor is straightforward, easy to implement, and minimizes bad matches with control observations that have little in common with the treated observation. It is also straightforward to check which country is matched as the nearest neighbor in a control group. But this 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 since it does not place greater weight on closer matches. Local-linear matching has several advantages over kernel matching, such as a faster rate of convergence near boundary points and greater robustness to different data design densities.

untreated observations, which is called “on support”. If the propensity score of a treated observation is not in the range of any of the untreated observations, it is called “off support”. We will report the number of off-support observations for each of the four policies and drop any such treated observations because the untreated observations are not useful comparisons for treated off-support observations. This procedure is also known as imposing a “common support condition” and can help reduce the effect of bad matches.

An important criterion to assess if a matching methodology is valid is to verify that the matching removed any significant differences in observable variables between the treated and control groups. This test of “balancing” or the “independence assumption” requires that:

$$\mathrm{D} \perp X \mid p(X). \tag{4}$$

This is implemented by calculating the *t*–statistic of the difference in means for each variable between the treated and control group (as identified by one of the matching methods). A successful matching would remove any differences in observables between the treated and the control group.

It is instructive to mention how this propensity score matching compares to the more familiar regression analysis.[26] Both matching and regression methodologies estimate the partial correlation of the treatment with the outcome variable conditional on the values of covariates. One difference between these two methods is the weighting of the covariate-specific differences between the treated and the untreated (i.e. control) groups. Some type of weighting is needed to calculate the average effect for the whole sample. In 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, i.e., the observations that are most similar to the treated but were untreated. In contrast, in regression analysis, the greatest weights for the comparison group 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 different weighting strategies can lead to large differences between regression and propensity-score matching results.

[26] Angrist and Pischke (2008, Chapter 3) and Imbens (2014) present excellent discussions of the similarities and differences between regression analysis and propensity-score matching, including their relative advantages.

Propensity score matching also differs from regression analysis in its greater emphasis on modeling the policy change (i.e., raising interest rates) rather than the specific functional form that links the policy change to the outcome (i.e., how raising interest rates affects GDP growth). This reduces challenges related to simultaneity and choosing an appropriate lag length. It also easily incorporates the use of a large set of variables to estimate the propensity score. This is useful when there is only vague theoretical guidance on the set of variables to be included in the model.

Propensity score matching, however, also has several disadvantages relative to regression analysis. The requirement for a sufficient number of "similar" observations for matching can be difficult to meet when using an annual, cross-country panel, as is typical in macro and international economics.[27] Satisfying the balancing assumption, as discussed above, is also difficult in some macroeconomic studies, showing the challenges in creating a set of untreated observations that are not significantly different than the treated observations. Even when all of these criteria are satisfied, different matching methodologies can yield different results, so it is important to check for the robustness across the various matching methods to ensure that results are not just an artifact of the matching methodology. Also, it is important to note that matching methodologies that involve panel data cannot effectively estimate effects over long periods of time because, even if the technique creates an accurate control group at the point of the treatment (policy change), this does not control for subsequent events that may make the treatment and control groups less similar over time.

Finally, two additional potential concerns related to this time dimension of the propensity-score methodology merit discussion.[28] As noted above, we lag the variables used in the first-stage logit estimates to reduce the likelihood that a policy that is undertaken affects the determinants of that policy (e.g., a major interest rate increase is not likely to have caused a change in GDP in the previous period). In the subsequent estimates of the ATTs, there is the possibility of endogeneity in the initial period (e.g., of correlation between an interest rate increase and GDP in the same period), but any such endogeneity is less likely to be an issue in the subsequent six quarters on which we will largely focus our discussion and evaluation of different policies. Also, if a country adopts an additional policy response after the initial

[27] This is one reason why propensity-score matching has been more widely adopted in labor economics, where papers often have larger samples of individuals that provide more degrees of freedom for effective matching.
[28] Most papers using propensity-score matching have focused on the cross-section effects of policies and not included a time-dimension.

treatment or in the same quarter, this is not directly controlled for in the analysis. Tables 1a and 1b, however, suggest that this does not occur frequently as the preponderance of major policy responses (as defined in the paper) do not occur within the same quarter or in the quarter before or after. The logit estimates used for the matching also control for some cases of the use of multiple policies because one set of variables controls for previous policy actions and is therefore part of the calculation used to match treated and control observations.

B. Results of Propensity-Score Matching

We use the estimates from the logit model in Table 2 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.[29] 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.

Before turning to formal tests of the performance of the different matching algorithms, it is useful to examine the "nearest neighbor" matches obtained to get an intuitive understanding of the types of matches obtained using this methodology. An examination of these nearest neighbors indicates that although some matches seem more intuitive than others, they are generally matched across time and to countries that are similar according to measurable statistics. For example, focusing on several well-known events, Korea's large devaluation in 1997Q4 is matched with the Czech Republic in 1997Q4, and Brazil's large devaluation in 2001Q1 is matching with Latvia in 2001Q1. Continuing with Brazil, Brazil's large interest rate increase in 2001Q3 is matched with Turkey in 1998Q2, Indonesia's interest rate increases in 1997 Q3 is matched with Argentina in 1997Q4, and Iceland's in 2007Q1 is matched with the Netherlands in 2008Q3. The matches for large reserve sales are often not as intuitive, undoubtedly reflecting the

[29] 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.

limited set of countries with large reserve stockpiles that are not depleted during periods that other countries with large reserve holdings decide to sell large shares of those reserve sales. For example, Singapore's large reserve sales in 2007Q1 are matched with China in 2007Q3.

Moving next to more formal tests of the matching technique, diagnostics of the different matching algorithms are presented in Table 3. The first row in each of the four panels, which correspond to the four policies, reports the number of observations in the treatment group and unmatched control group in the first two columns. The columns to the right report in parentheses the number of observations that are off-support after using each matching algorithm.[30] These statistics show that in most cases, almost all treatments are on support—with between zero and two observations off support for each of the policies and each of the matching algorithms. The only exceptions are with radius and kernel matching for major reserve sales, which have nine and eight off-support observations, respectively. Radius and kernel matching also have one more observation off support relative to the other techniques for large interest rate increases. The second row in each panel reports how many variables fail the independence tests for each policy response and matching algorithm. These independence tests evaluated if the matching was able to remove any significant differences in observable variables between the treated and control groups. This shows that the matching techniques all perform very well, with no cases in which any variables fail the independence test when evaluated at the 5% significance level. There is only one case when the test fails at the 10% level: when local-linear matching is used for reserve sales, two variables fail.

In order to further explore the performance of these independence tests, and the one case when it shows weaker results at the 10% significance threshold, Table 4 reports the underlying results of tests of the independence assumption for 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 treated and unmatched control groups for nine of the 13 variables used in the logit regression. The remaining columns of Table 4 report results of the same t-tests, except now for the difference in means between the treated

[30] The numbers in the treated group are less than the numbers in Tables 1a and 1b because not all of the variables used in the logit regressions are available for all countries for all years.

group and control groups after matching (using each of the five techniques). 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 matched control groups at the 95% significance level. (With local-linear matching, there is only a significant difference for 1 variable at the 90% significance level.)

Tables similar to Table 4 for the other policy responses (reserve sales, large increases in interest rates, and increased capital controls) are not reported but yield similar results. In each case, there are significant differences between the variables in the treated group and unmatched control group according to several variables, but these differences are removed after using each of the propensity-score matching techniques. These results are also summarized in the second row on each panel in Table 3. The last two rows in each panel in Table 3 report 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 relative to that of the control group. Different matching methods yield mean propensity scores closest to that for the treated group, with the local-linear matching yielding the closest scores for two of the four policy responses. Therefore, given that local-linear matching has a number of stronger theoretical advantages than the other techniques (as discussed above) and performs at least as well (if not better) than the other techniques according to each of the test criteria, we will focus on these results in the discussion below (but also include other results if they differ significantly).

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

This section uses the matched control groups to analyze the average treatment effect on the treated (ATT) of large reserve sales, sharp currency depreciations, major increases in interest rates, and new controls on capital outflows on real GDP growth, inflation, and unemployment. We use bar graphs to present the ATT on outcome variables at each quarter over a one-and-a-half-year period, beginning with the quarter that the major policy response (the treatment) occurs and over each of the next six quarters. This allows us to evaluate both immediate and medium-run effects. (We do not consider longer-term effects because the matching algorithms become less accurate.)

The heights of the bars in the graphs represent the effects of each of the policies relative to the counterfactual, estimated as the change in each of the outcome variables relative to the average over the two years before the policy change occurred. For example, after a policy change such as a major increase in interest rates, we calculate real GDP growth during the quarter of the policy change (which we denote as t=0), relative to average GDP growth over the previous eight quarters (t-1, t-2, … t-8). For the next six post-treatment periods, we calculate the change in GDP growth in the respective periods relative to average GDP growth over the same eight quarters before the initial major policy change.[31] This 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 *ceteris paribus* approach in that it does not incorporate any adjustment for changes in post-treatment covariates. The black line on each graph is the fitted line for the average treatment effect. The color of the bars in the graphs indicates the statistical significance of these results; black indicates that the ATT for that quarter is significant at the 95% confidence level (or greater), medium-blue indicates significance at the 90% to 95% level, and light blue indicates that the effect is statistically insignificant. The standard errors used to calculate these levels of significance must be calculated through bootstrapping methods because the propensity scores are estimated.[32]

A. Base Case Results

Figures 2 through 4 present bar graphs summarizing the estimated effects of each of the four policies on GDP growth, unemployment, and CPI inflation, respectively. Figure 2 indicates that major interest rate increases, major currency depreciations, and adding controls on capital outflows all initially have negative and significant effects on GDP growth. There is no evidence of a significant effect of reserve sales on GDP growth. Increasing interest rates initially has the strongest effect, with an estimated decrease in GDP growth of three to four percentage points in the three quarters after interest rates were increased, although fading sharply afterwards. Adding new controls on capital outflows also has a large significant effect, peaking at a reduction in GDP growth of almost six percentage points in the second quarter after the controls were added, with the negative effect still significant (although falling to just under four percentage points) after five quarters. Large currency depreciations are estimated to initially have a smaller effect, peaking at GDP growth about three percentage points lower in the second quarter, and this effect quickly fades and even appears to reverse so that GDP growth increases by about two percentage

[31] As a robustness check, we also consider the ATT of GDP growth without reference to the pre-treatment values; see Section F below. These results are very similar to the base case results presented in this section.
[32] See Lechner (2002) for the appropriate methodology. We use 500 repetitions for the bootstrap.

points in the sixth quarter after the depreciation. This estimated impact of a large depreciation 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. This result also suggests that a policy that is typically believed to be expansionary over time may initially have contractionary effects.

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 generate a significant impact on unemployment at the 95% confidence level or greater. (In fact, the only result that is significant at the 10% level is possibly a negative effect of a major depreciation on unemployment for the period $t = 0$). This lack of significant estimates, despite the significant results found for GDP growth, is consistent with the generally lagged and muted response of unemployment to GDP. Furthermore, another 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 policies and GDP affect unemployment.

Figure 4 shows the estimated effects of each of the four policies on CPI inflation. Inflation appears to increase after major increases in interest rates, major currency depreciations, and increased controls on capital outflows. Inflation appears to decrease after major reserve sales. The only estimated effect that is significant, however, is the rise in inflation beginning three periods after a major increase in interest rates. The magnitude of most of these estimated effects on inflation is also large.

The results presented in Figures 2, 3, and 4 show that different policy responses to crises can yield very different outcomes. Large currency depreciations appear to be the only policy response 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 an increase in inflation. Benefits from any of these policies 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. Extension: ATTs for Emerging and Non-OECD Economies

As a first extension of our analysis, we repeat the base-case estimation focusing on a subsample of countries that excludes high-income OECD economies (based on World Bank classifications, which vary each year). High-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.

The ATTs from this exercise estimating the effect of large reserve sales, major depreciations, substantial interest rate increases, and new capital controls on growth, unemployment and inflation for emerging markets (and non-OECD economies) are presented in the twelve graphs in Figure 5 in which the rows represent treatments and the columns represent outcomes. We continue to focus on results based on local-linear matching to facilitate comparability with the base-case results and also as this continues to perform as well as or better than other techniques according to the tests discussed above. Many of the results correspond to those for the larger sample, which includes high-income OECD economies. Major increases in interest rates, new capital controls, and large depreciations continue to be correlated with higher inflation and lower GDP growth relative to the counterfactual—at least initially.

But there are also noteworthy differences between the results presented in Figure 5 and those presented in Figures 2 through 4. Many of the estimated effects of these policies are larger in magnitude for the emerging markets than the full sample. For example, capital controls are now estimated to be correlated with inflation that is 8% higher (instead of 4.5% for the full sample) and major depreciations are correlated with inflation that peaks at 3.5% higher (versus 1.5% in the base case). The estimated effects of capital controls and major depreciations on GDP growth are also larger than in the full sample. Some of the estimates that were significant for the full sample, however, are less precisely estimated for the sample of emerging markets and non-OECD economies. For example, the initial negative effect of depreciations on growth is no longer significant. The negative effect of capital controls on growth is also no longer significant. These results may indicate that the diversity across emerging markets makes it more difficult to estimate a consistent effect of these policies, or that these policies are less effective in emerging markets. The larger magnitude of the estimated effects, however, also indicates that policies can have large—albeit imprecisely estimated—effects on growth, inflation, and unemployment in emerging markets.

C. Extension: ATTs for Each Crisis Period

As an additional extension, we estimate the effects of each of the four policies separately for each crisis window, the "1990s crises" from 1997 to 2001 and the "2000s 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.[33]

One difference between this exercise and the base case is that the matching algorithms perform less well than for the full time period. Each of the techniques has 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 large reserve losses during the 1990s crises. Between seven observations (for nearest neighbor, five-nearest neighbor and local-linear matching) and 10 or 11 observations (for radius and kernel matching) are off support when predicting major currency depreciations in the 2000s window. In contrast, no more than two observations were ever off support in the base case, which combined the two crisis periods using any of the matching methods (except for radius and kernel matching predicting reserve losses). The matching algorithms also more often fail the tests of the independence assumption.

With these important caveats that the propensity score matching methodology yields inferior performance in these more limited samples, Figures 6.A, 6.B, and 6.C graph the ATTs for GDP growth, unemployment, and inflation, respectively, for just the period from 1997 to 2001 (the left column in each of these figures) and just the 2007-2011period (the right column). In general, the estimated direction of the effects of the different policy responses agrees with those reported in the base case and other sensitivity tests. During the 1990s, the effects of various policies on inflation appear to be somewhat larger (relative to the base case and the 2000s), which is consistent with the theory that inflation expectations may have been less firmly

[33] Claessens and Kose (2014) discusses similarities and differences across crises.

anchored during this period. Major depreciations are no longer estimated to provide any significant boost to growth over the longer term as found in the full sample, but depreciations are estimated to generate a significant and sustained reduction in unemployment during the 2000s.

The most noteworthy differences in results across the two periods and relative to the base case for the full sample are the effects of major increases in interest rates. Higher interest rates are positively and significantly correlated with unemployment in the 1990s, but negatively and significantly correlated in the 2000s. Higher interest rates are also no longer correlated with a significant decline in GDP growth in the 2000s (although they continue to have a negative and significant relationship in the 1990s as found in the full sample). These different results for interest rate policy may reflect changes in monetary policy relationship related to the zero-lower bound or to the massive contraction in credit related to the global nature of the financial crisis. Given the small sample sizes and corresponding limits to the use of the propensity-score methodology with this limited variation, it is impossible to draw any strong conclusions from this set of results for the individual crisis periods.

D. 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.

To test for the impact of using a broader set of covariates in the logit regression, we use the full set of covariates listed on the left of Table 2 rather than the subset of variables that were significant at the 20% level (in a regression with the full set of covariates). Most of the results are similar to those in the base case in terms of direction and magnitude, but in some cases results which were previously significant at the five-percent level are no longer significant. For example, the estimated positive relationship between sharp increases and interest rates is no longer significant and some of the estimated effects of capital controls and depreciations on growth (in certain quarters) are both insignificant under some matching methodologies. This reduced significance is as expected as a larger set of covariates than in the base case would be expected to generate less precise estimates with fewer significant estimated ATTs.

We have also attempted to estimate an augmented version of this 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 included in the base case, however, and their inclusion reduces the sample size and, importantly, the number of treated observations to the point that the matching methodology no longer works well.

E. Sensitivity Tests: Major Policy Responses Defined Based on Looser Thresholds

In Section II, we discussed the criteria used to identify changes in reserve holdings, exchange rates, and interest rates as "major" policy responses. In order to focus on large and infrequent policy responses, we established a threshold to qualify as a "major" response based on the criteria that it occurred in only 5% of the country-quarter observations during the two crisis windows. As an additional sensitivity test, we use looser criteria for a policy to qualify as a "major" policy response and instead set the threshold at 10%.

Applying this looser threshold to the same sample of countries, and making no other changes to the sample construction or additional criteria to qualify as a major policy change, yields the resulting cutoffs. Major reserve sales are defined as quarters during which international reserves (relative to GDP) fall by at least 7.8 percentage points compared to the previous year (versus 13.4 percentage points in the base case).[34] Large depreciations are defined as quarters in which there is at least a 16.1 percent depreciation over the previous year in the country's exchange rate versus the U.S. dollar (versus 22.6 percent in the base case). 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 133 basis points over the past year (versus 200 basis points in the base case).[35] All other requirements and criteria are unchanged. The threshold to qualify as using new controls on capital outflows is unchanged, as any additional controls qualified as adopting this policy response and there was no numerical threshold.

After establishing these looser definitions of major policy responses, we re-estimated the logit regression with the full set of covariates listed on the left of Table 2. Then, following the methodology in the base case, we included only variables that were significant at the 20% significance level to calculate the final logit and corresponding propensity scores. We then used

[34] Details on sources and definitions for the data are in the Appendix.
[35] 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.

each of the four matching methodologies to estimate the impact of the four policy responses on growth, unemployment and inflation. The results are very similar to those in the base case, and in many cases strikingly so (and therefore not reported). One noteworthy difference is that the magnitude of the estimated effect is slightly smaller in some cases, such as for the effect of different policies on inflation. This is not surprising given that weaker policy responses are now included as treatments. The only other noteworthy change is that depreciations are correlated with a small and significant increase in unemployment after about a year based on local-linear matching and five-nearest neighbor matching, but this result reverses and suggests a small and significant decrease based on kernel matching (and is insignificant at the 5% level based on radius matching).

F. Sensitivity Tests: Real GDP Growth without Reference to Pre-Treatment Values

The ATT effects for real GDP growth presented in the four panels of Figure 2 control for pre-treatment differences in real GDP growth between the treated observations and their respective controls, that is:

$$\text{ATT}=\{\text{E}[G_{1,i,After}|D_i\text{=1}] - \text{E}[G_{0,i,After}|p(X),D_i\text{=1}]\} - \{\text{E}[G_{1,i,Before}|D_i\text{=1}] - \text{E}[G_{0,i,Before}\ |p(X),D_i\text{=1}]\}.$$

This is appropriate for considering the relative performance of the treated and control observations relative to a trend, as represented by the two-year average of GDP growth before the treatment. An alternative specification is one in which the ATT represents the direct difference in real GDP growth between the treated observations and their respective controls, without reference to the respective trends, that is:

$$\text{ATT}' = \{\text{E}[G_{1,i,After}|D_i\text{=1}] - \text{E}[G_{0,i,After}|p(X),D_i\text{=1}]\}.$$

One could argue that this is the more appropriate specification, since the ATT includes $\text{E}[G_{1,i,Before}|D_i\text{=1}] - \text{E}[G_{0,i,Before}\ |p(X),D_i\text{=1}]\}$, but the focus of the analysis should not be altered by pre-treatment differences between the treated and control groups.

As an additional robustness test, we use this alternate specification and measure the ATT′ for real GDP growth that does not consider real GDP growth relative to trend. In other words, instead of testing for differences between the treated and untreated groups in growth rates relative to each country's trend, we simply compare the growth rates across these two groups. The results using local-linear matching are shown in the four panels of Figure 7, and can be compared to the four bar graphs presented in Figure 2 (which are based on changes in growth relative to trends). The two sets of graphs using the different growth measures are very similar based on their patterns and magnitudes. Which bars are significant, however, varies somewhat across specifications (as often occurs with different matching techniques).

More specifically, major increases in interest rates and increased controls on capital outflows continue to reduce growth, especially in the quarters immediately after the policy is adopted, although these results are more often significant in the original specification in Figure 2. Major currency depreciations continue to initially slow growth, and then to raise growth over time (following a J-curve effect), but the later boost to growth is more often significant in the newer specification in Figure 7. Results based on large reserve sales are insignificant in each case.

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 that they adopt these policies at specific times in response to changes in specific variables. Propensity-score matching takes these econometric issues seriously when evaluating the effects of policy responses to crises.

The results indicate that none of the traditional responses to crises will yield significant and meaningful improvements in GDP growth, unemployment, and inflation relative to the counterfactual. Instead, large interest rate increases, major reserve sales, large currency depreciations, and new controls on capital outflows are ineffective at improving these three outcome variables and often imply substantial tradeoffs. More specifically, sharp increases in interest rates and new controls on capital outflows appear to significantly reduce GDP growth, have no consistent effect on unemployment, and are correlated with higher inflation. Large depreciations also appear to initially reduce GDP growth and be correlated with higher inflation, but there is some evidence that they can yield a lagged improvement in growth and reduce unemployment, especially during the 2000s. Major reserve sales generally have insignificant effects and even the estimated direction of these effects fluctuates across time for each of the key outcome variables. There is some evidence indicating that all of these policy responses have larger, but more imprecise, effects on emerging markets and that changes in interest rates had different effects during the 1990s than the 2000s. But both of these results should be interpreted cautiously due to the limitations resulting from small sample sizes.

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, Òscar Jordà, and Guido Kuersteiner. (2013). "Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited." NBER Working Paper 19355.

Angrist, Joshua and Guido Kuersteiner. (2011). "Casual Effects of Monetary Shocks: Semiparametric Conditional Independence Tests with a Multinomial Propensity Score." *Review of Economics and Statistics* 93(3): 725-747

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, Wenjie Chen, and Kathryn Dominguez. (2011). "Foreign Ownership and Firm Performance: Emerging Market Acquisitions in the United States." Unpublished mimeo.

Chari, Anusha and Peter Blair Henry. (2015). "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. (2014). "Financial Crises: Explanations, Types and Implications." In Claessens, Stijn, Ayhan Kose, Luc Laeven and Fabian Valencia, eds., *Financial Crises: Causes, Consequences and Policy Responses*, International Monetary Fund.

Claessens, Stijn, Ayhan Kose, Luc Laeven and Fabian Valencia, eds. (2014). *Financial Crises: Causes, Consequences and Policy Responses*, International Monetary Fund.

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

Dehejia, Rajeev, and Sadek Wahba. (2002). "Propensity-Score Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1):151-161.

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 (2015). "Capital-Flow Management Measures: What are they Good For?" *Journal of International Economics*, forthcoming.

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.

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

Imbens, Guido, "Matching Methods in Practice: Three Examples," *NBER Working Paper* #19,959, March 2014.

Jordà, Òscar, and Alan Taylor. (2013). "The Time for Austerity: Estimating the Average Treatment Effect on Fiscal Policy." NBER Working Paper 19414.

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 (2): 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.

Levchenko, Andrei, Romain Rancière, and Mathias Thoenig. (2009). "Growth and Risk at the Industry Level: The Real Effects of Financial Liberalization." *Journal of Development Economics* 89: 210-222.

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 and Donald Rubin. (1985). "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *The American Statistician* 39(1): 33-38.

Rubin, Donald and Neal Thomas. (1992). "Characterizing the Effect of Matching Using Linear Propensity Score Methods with Normal Distributions." *Biometrika* 79 (4): 797-809.

Shambaugh, Jay (2004), "The Effect of Fixed Exchange Rates on Monetary Policy", *Quarterly Journal of Economics* 119(1): 300–351.

Table 1a: Policy Changes, 1997 – 2001

	RS	ER	IR	CFM	Total
Foreign Reserve Sales (RS)	9	2	6	0	18
Exchange Rate Depreciation (ER)		14	8	1	27
Interest Rate Increase (IR)	1	1	17	2	35
Capital Flow Management (CFM)				13	18

Reports instances of major policy changes (see below). Numbers in diagonal cells represent policy change that occurred with no other policy change between one quarter before and one quarter after that change. Off-diagonal elements represent pairs of policy changes that occurred within the same three quarter period (i.e. one policy was matched by another that occurred between one quarter before and one quarter after that one). Overlay boxes represent triplets or quadruplets of policy changes that occurred within plus or minus one quarter.

Table 1b: Policy Changes, 2007 - 2011

	RS	ER	IR	CFM	Total
Foreign Reserve Sales (RS)	33	6	2	2	44
Exchange Rate Depreciation (ER)	1	13	2	1	24
Interest Rate Increase (IR)		1	9	2	17
Capital Flow Management (CFM)				9	15

Reports instances of major policy changes (see below). Numbers in diagonal cells represent policy change that occurred with no other policy change between one quarter before and one quarter after that change. Off-diagonal elements represent pairs of policy changes that occurred within the same three quarter period (i.e. one policy was matched by another that occurred between one quarter before and one quarter after that one). Overlay boxes represent triplets or quadruplets of policy changes that occurred within plus or minus one quarter.

Notes on Thresholds for Defining Major Policy Changes: Major policy changes defined as occurring in only 5% of the country-quarter observations in the full sample of crisis years (1997-2001 and 2007-2011). Large reserve sales are based on the reserve-to-GDP ratio. The sharp depreciation and large interest rate changes are only counted in quarters in which annual inflation did not exceed 10% or more. New capital controls are new controls on capital outflows (as described in the text). Countries are not included in the criteria used to evaluate large reserves sales, exchange rate depreciations, or interest rate increases during years for which they are members of the euro zone. We include an exclusion restriction such that a major policy change cannot occur more than once in four consecutive quarters.

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

		Reserve Sales	Currency Depreciations	Interest Rate Increases	New Capital Controls
Global	Global risk	0.030	0.071***		-0.054
Measures		(0.023)	(0.021)		(0.028)
	U.S. interest rate	-0.469***	-0.355		
	(ch)	(0.169)	(0.227)		
	Commodity price		2.760	4.321***	
	Index (log)		(1.826)	(0.999)	
	1990s crisis		1.963	4.866***	0.614
	dummy		(1.417)	(0.966)	(0.459)
Domestic	Real GDP growth		-0.072***		
Vulnerabilities	(ch)		(0.025)		
	Capital outflows,		-6.682***	3.053***	-1.319
	% of GDP (ch)		(1.964)	(1.156)	(0.914)
	Capital inflows,	0.713	6.231***		
	% of GDP (ch)	(0.607)	(2.334)		
	Current account			-1.276*	
	% of GDP			(0.750)	
	Inflation (ch)	0.085			-0.043
		(0.063)			(0.030)
	Private credit, %				0.010
	of GDP (ch)				(0.017)
	Commodity index *	1.240**		1.567***	1.042**
	Comm.. exporter	(0.505)		(0.402)	(0.416)
	Income per capita		-0.438*		
Country	(log)		(0.250)		
Characteristics	Institutions index	0.514		-2.491**	-1.146
		(1.619)		(1.099)	(0.762)
	Reserves as %	2.122***	-2.897***	-0.227	
	of GDP	(0.301)	(0.819)	(0.326)	
	Peg dummy		-2.667**		
			(1.175)		
	Openness	-0.553***	-0.277**	-0.350**	-0.061
		(0.205)	(0.140)	(0.139)	(0.165)
Recent Policy	Reserves as %	-21.022***	-9.427***	-4.465***	
Changes	of GDP (ch)	(5.073)	(3.141)	(1.650)	
	ER vs. U.S.	-0.084***	0.085***		
	$ (pch)	(0.022)	(0.022)		
	Interest rate vs.	-0.164**		0.133***	
	U.S. rate (ch)	(0.071)		(0.039)	
	Capital controls				0.930*
	(ch)				(0.572)
Observations		***1079***	***1060***	***1539***	***1454***
Pseudo R^2		***0.41***	***0.39***	***0.22***	***0.08***

Notes: Covariates included in each regression if they are significant at the 20% level (or less) in regressions that use full set of covariates. "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
Summary Statistics for Different Matching Algorithms

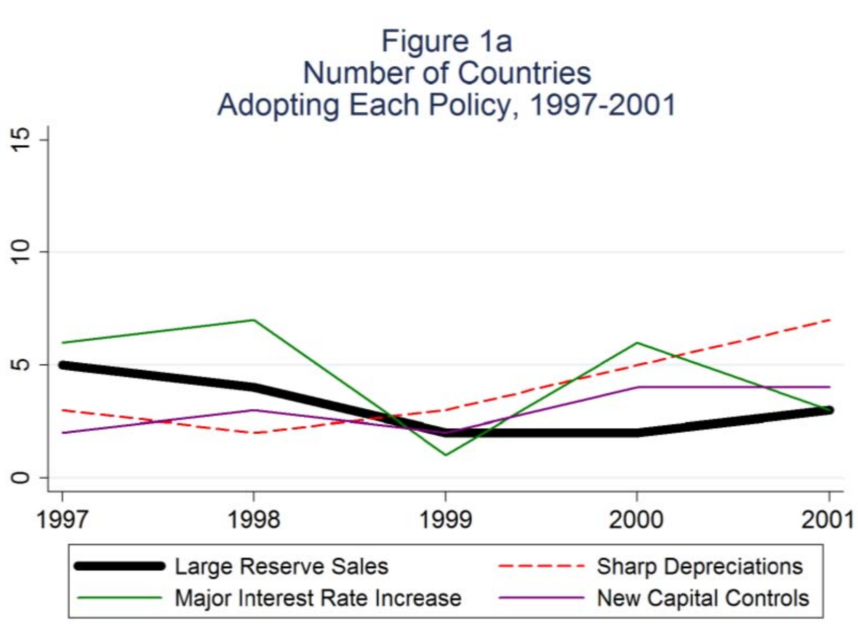
	Treatment Group	Unmatched Control Group	Matched Control Group Based on Matching Algorithm: Nearest Neighbor (no replacement)	5 Nearest Neighbors	Radius	Kernel	Local-Linear
Major Reserve Sales							
Observations (off support)[1]	38	1041	(2)	(2)	(9)	(8)	(2)
Variables failing independence test[2]			0	0	0	0	0
Mean Propensity Score[3]	599.6	392.7	459.9	431.7	424.1	424.2	472.1
Mean Absolute Bias[3]		34.7	12.3	11.6	7.8	8.7	19.3
Sharp Currency Depreciations							
Observations (off support)[1]	39	1021	(2)	(2)	(2)	(2)	(2)
Variables failing independence test[2]			0	0	0	0	0
Mean Propensity Score[3]	405.8	440.6	421.6	428.4	411.5	415.3	426.8
Mean Absolute Bias[3]		57.2	13.9	8.3	10.2	11.5	14.9
Large Interest Rate Increases							
Observations (off support)[1]	37	1566	0	0	(1)	(1)	0
Variables failing independence test[2]			0	0	0	0	0
Mean Propensity Score[3]	404.0	407.5	408.9	477.6	466.2	463.6	484.4
Mean Absolute Bias[3]		40.8	12.6	4.1	9.3	9.5	14.6
New Controls on Capital Outflows							
Observations (off support)[1]	26	1428	(1)	(1)	(1)	(1)	(1)
Variables failing independence test[2]			0	0	0	0	0
Mean Propensity Score[3]	435.5	391.0	351.9	414.0	408.5	409.6	431.3
Mean Absolute Bias[3]		27.5	20.8	7.9	12.4	12.2	20.8

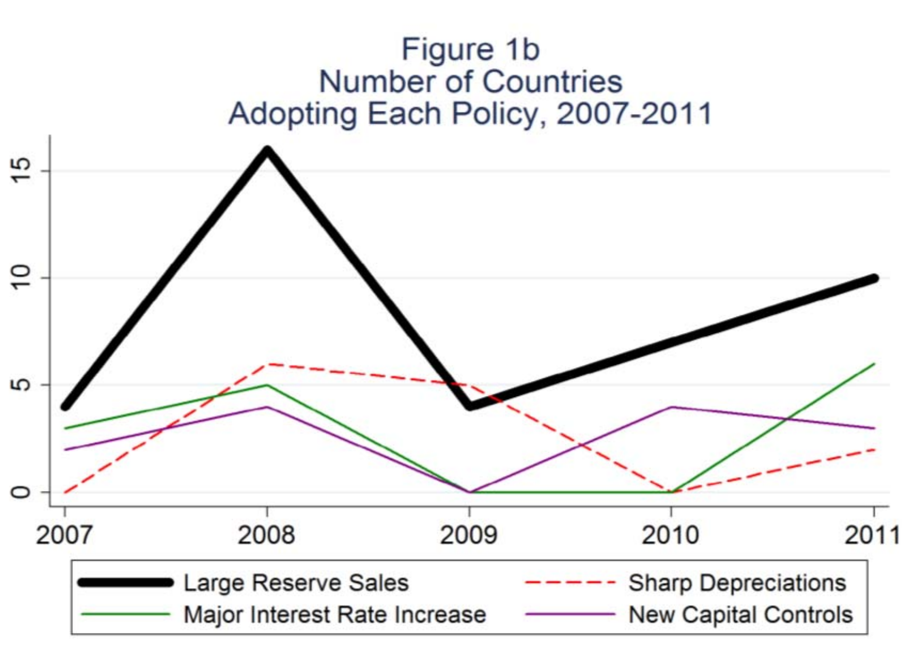
Notes: Results from matching based on regression results reported in Table 2. (1) Observations in the treatment and control groups reported in the first two columns. Observations that are off-support based on each matching method are reported in the right-hand columns in parentheses. (2) The number of variables that do not satisfy the independence test at the 5% significance level—as shown for the case of major depreciations in Table 4. (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 4
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. Mean: Matched Control	5 Nearest Neigh. t-stat	Radius Mean: Matched Control	Radius t-stat	Kernel Mean Matched Control	Kernel t-stat	Local-linear Mean Matched Control	Local-linear t-stat
Global Risk	32.6	24.3	6.30***	30.6	0.89	30.8	0.82	30.6	0.91	28.7	1.88*
ΔU.S. interest rate	-1.19	-0.38	4.72***	-0.99	0.50	-0.93	0.65	-0.92	0.66	-1.02	0.42
Commodity prices	4.68	4.79	1.75*	4.66	0.33	4.68	0.07	4.68	0.05	4.65	0.48
1990's crisis dummy	0.51	0.44	0.91	0.59	0.65	0.54	0.22	0.54	0.21	0.62	0.93
ΔReal GDP growth	-5.93	0.43	4.55***	-5.10	0.43	-4.75	0.63	-4.53	0.75	-3.76	1.15
Capital Outflows	-0.045	0.022	1.80*	-0.050	0.11	-0.037	0.18	-0.030	0.35	-0.042	0.08
Capital Inflows	-0.055	0.022	2.20**	-0.06	0.05	-0.046	0.21	-0.038	0.40	-0.057	0.05
Income per capita	8.22	8.36	0.69	8.07	0.56	8.15	0.26	8.14	0.33	8.16	0.24
Reserves/GDP	0.39	0.73	2.81***	0.40	0.14	0.45	0.68	0.46	0.72	0.38	0.24
Peg dummy	0.027	0.24	3.17***	0.027	0.00	0.06	0.72	0.07	0.79	0.27	0.00
Openness	0.90	1.37	2.43***	0.70	0.58	0.64	0.74	0.62	0.80	0.26	1.75
ΔReserves/GDP	-0.038	0.053	3.91***	-0.040	0.08	-0.02	0.57	-0.02	0.68	-0.04	0.12
%ΔER/U.S.$	10.49	0.18	7.05***	10.96	0.20	9.91	0.26	10.17	0.14	12.18	0.76

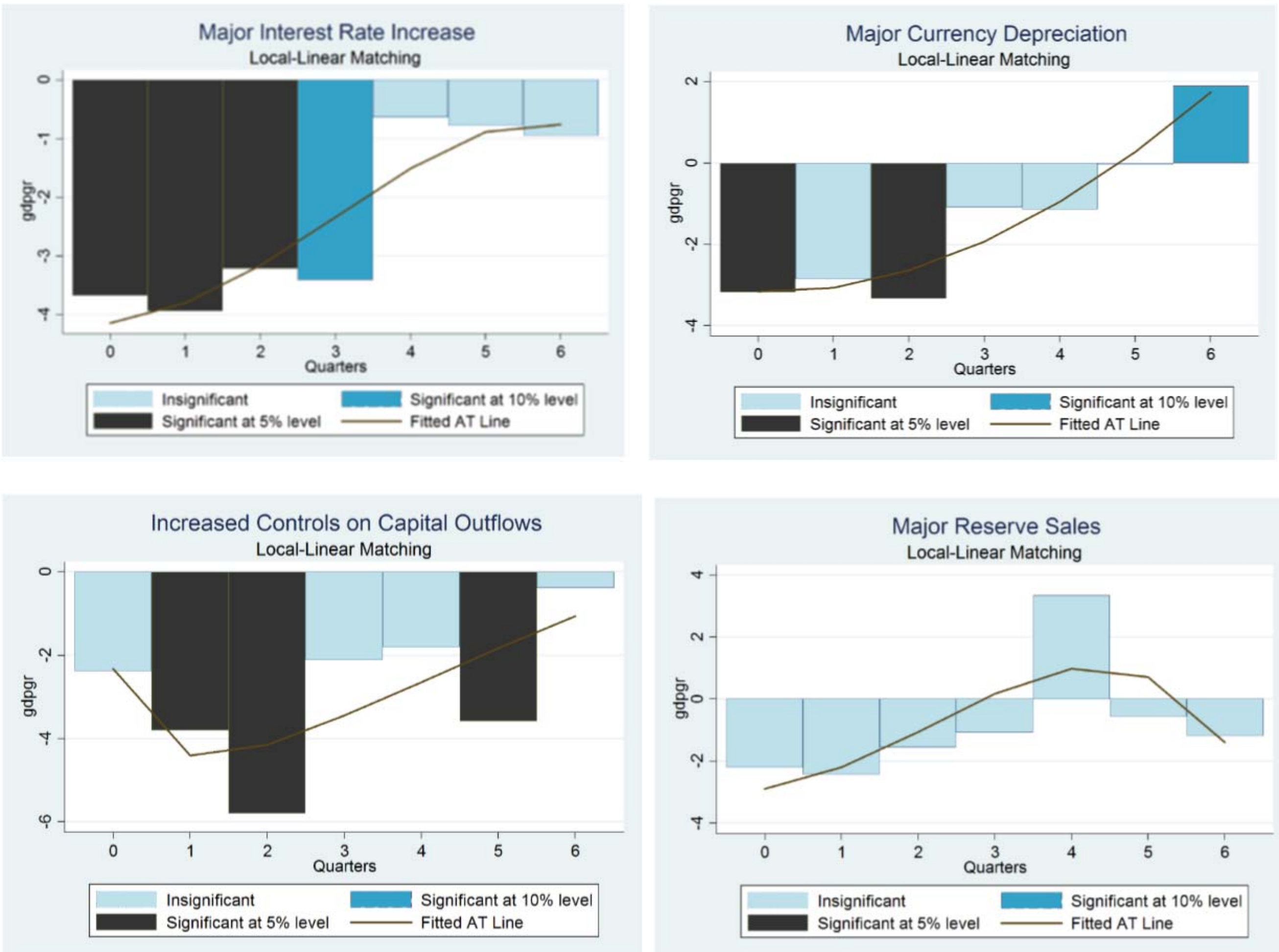
Notes: Reports difference in means between treatment and control groups, with control group created based on regression results reported in Table 2 and matching performed using algorithms listed at top of table. See Appendix for detailed variable definitions. * indicates significant at the 10% level, ** at the 5% level and *** at the 1% level.





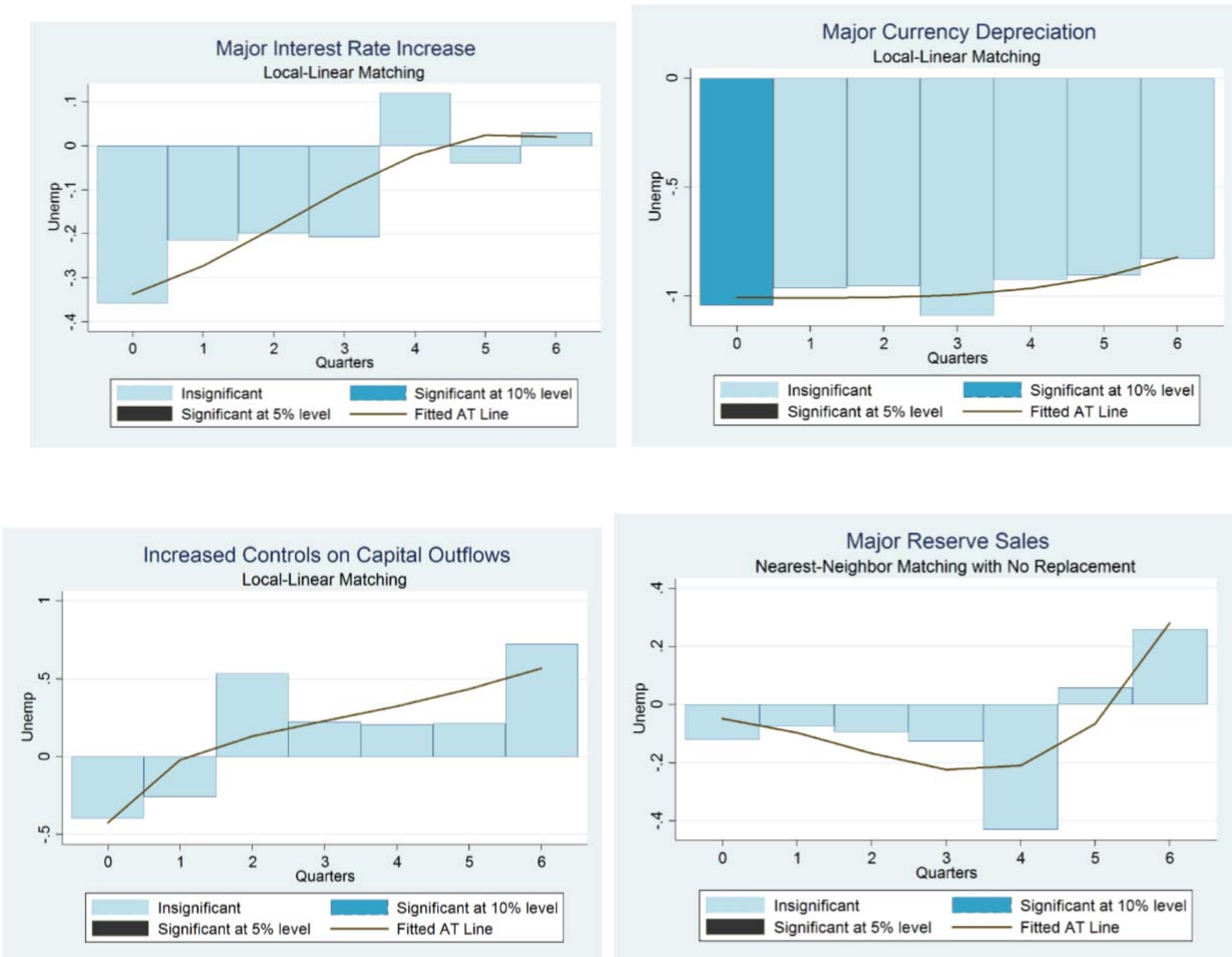
Notes: Reports annual time series of major policy changes. Criteria to qualify as a "major" policy change is set to equal 5% of the sample for reserve sales, sharp depreciations, and interest rate increases. Using this threshold, large reserve sales are at least a 13.4 percentage point decrease in reserves (relative to GDP) between one quarter and that quarter in the previous year. Sharp depreciations are at least a 22.6% deprecation in the nominal exchange rate versus the U.S.$ between one quarter and that quarter in the previous year (with annual inflation below 10%). Major increases in interest rates are at least a 200 basis points increase between one quarter and that quarter in the previous year (with annual inflation below 10%). 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). While a member of the euro zone, countries are excluded from the sample able to have large reserve sales, sharp depreciations, or interest rate increases.

Figure 2: Average Treatment Effects on Real GDP Growth using Local Linear Matching



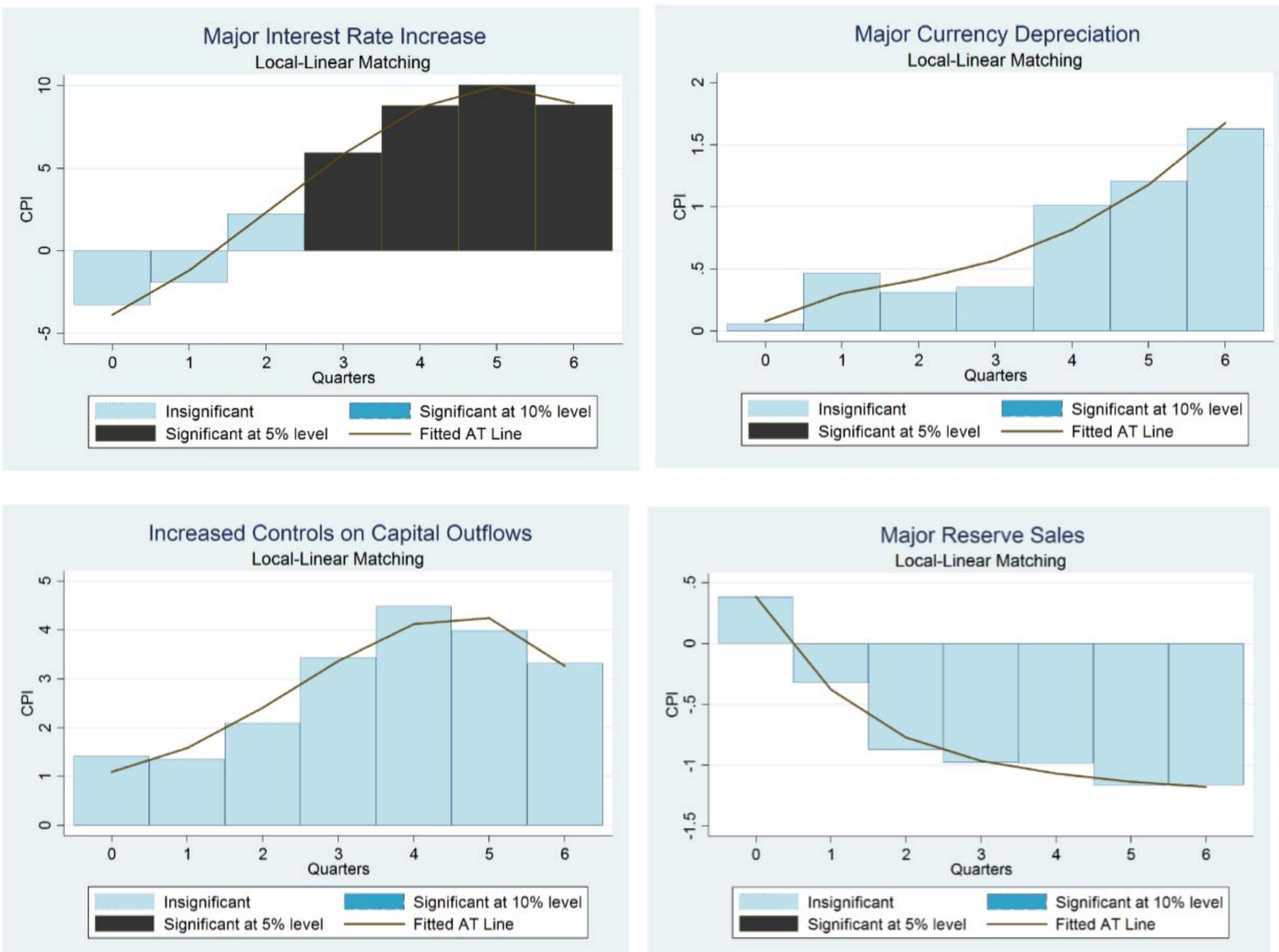
Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 3: Average Treatment Effects on Unemployment using Local Linear Matching



Note: ATE 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) using Local Linear Matching

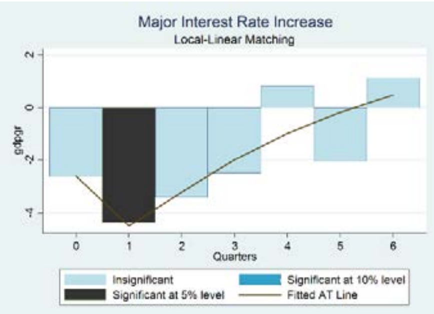


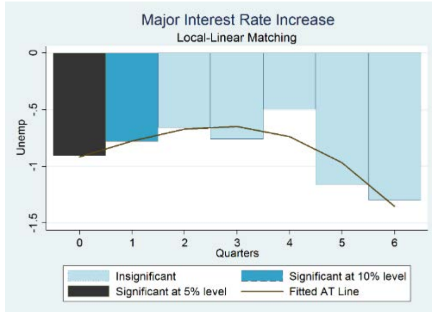
Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

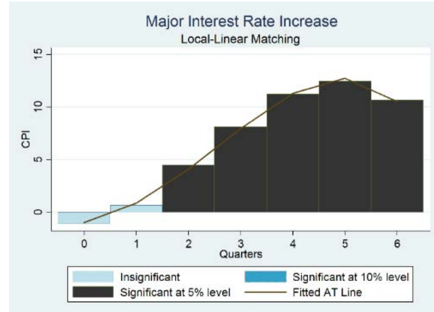
Figure 5: Average Treatment Effects for non-OECD and EMEs using Local Linear Matching

GDP Growth	Unemployment	Inflation

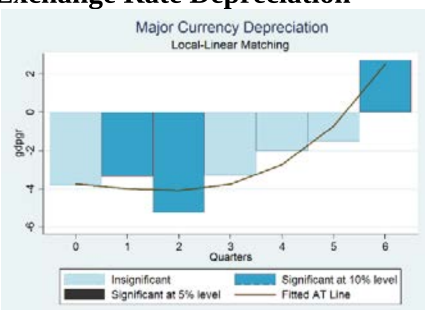
Interest Rate Increase

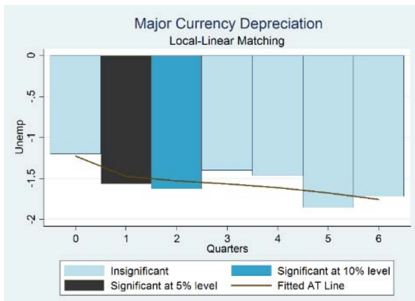


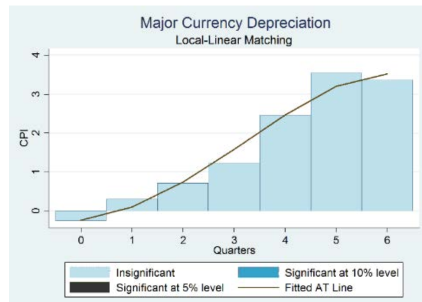




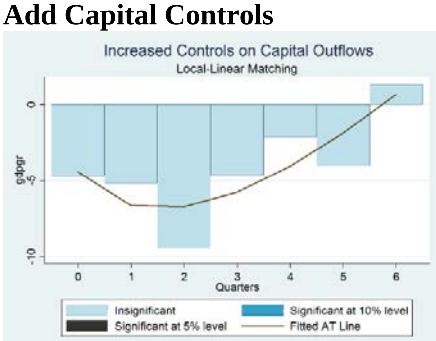
Exchange Rate Depreciation

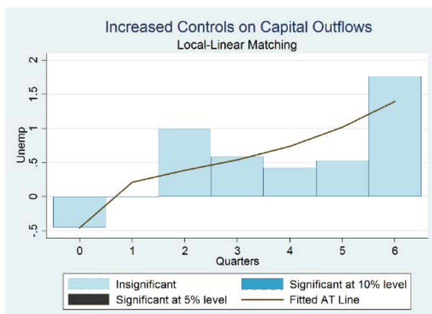


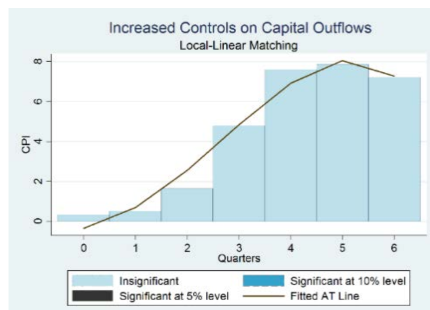




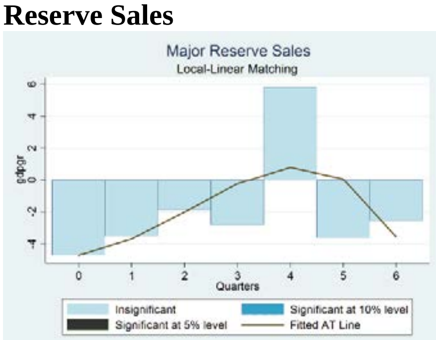
Add Capital Controls

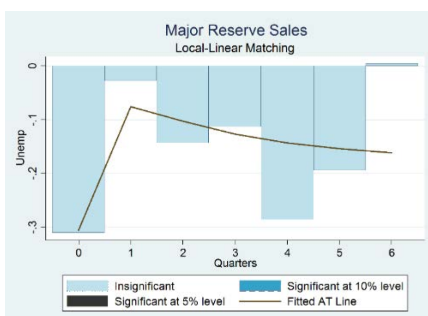


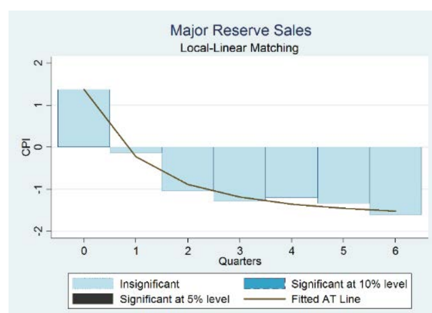




Reserve Sales



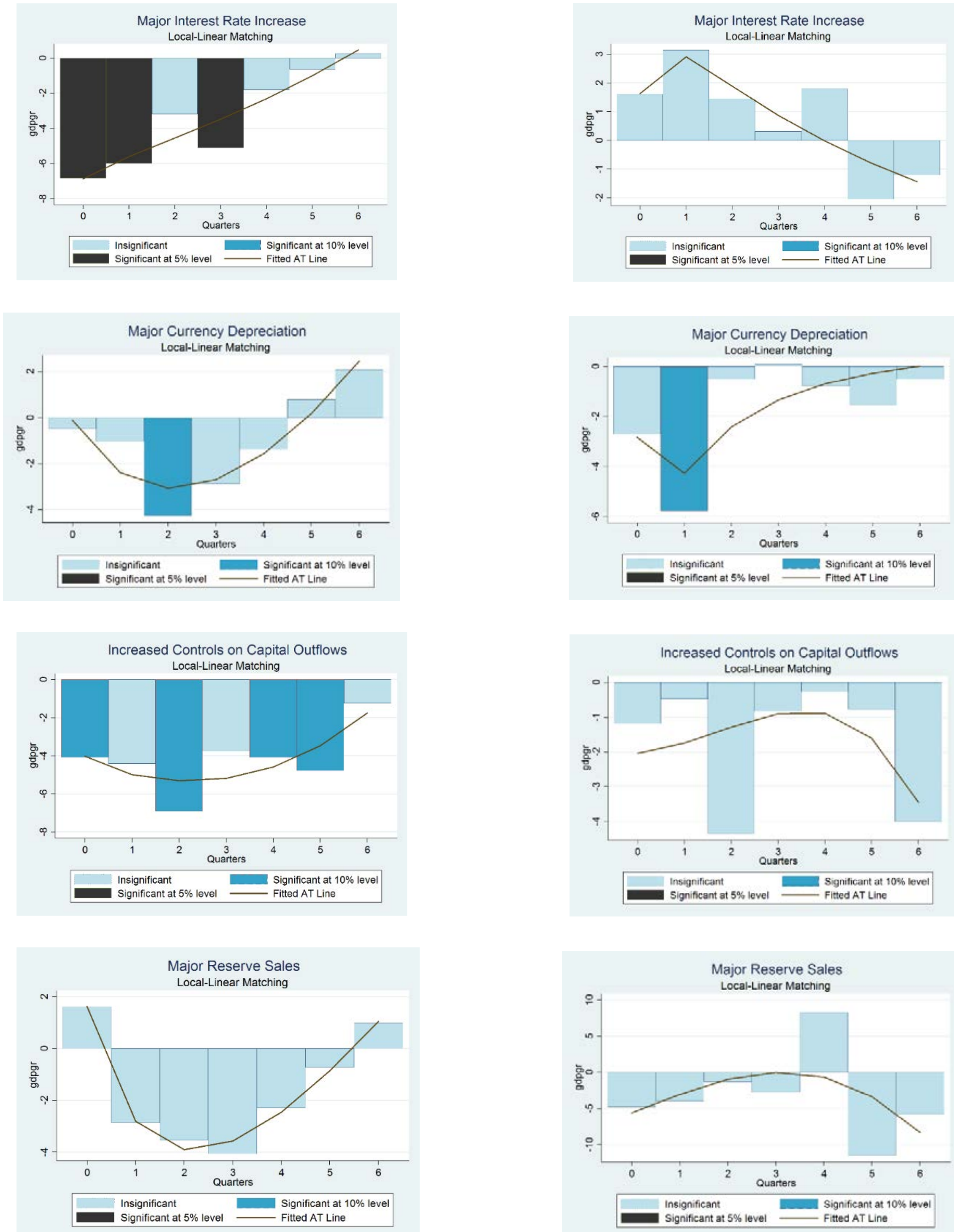




Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 6A: Average Treatment Effects for GDP Growth using Local Linear Matching

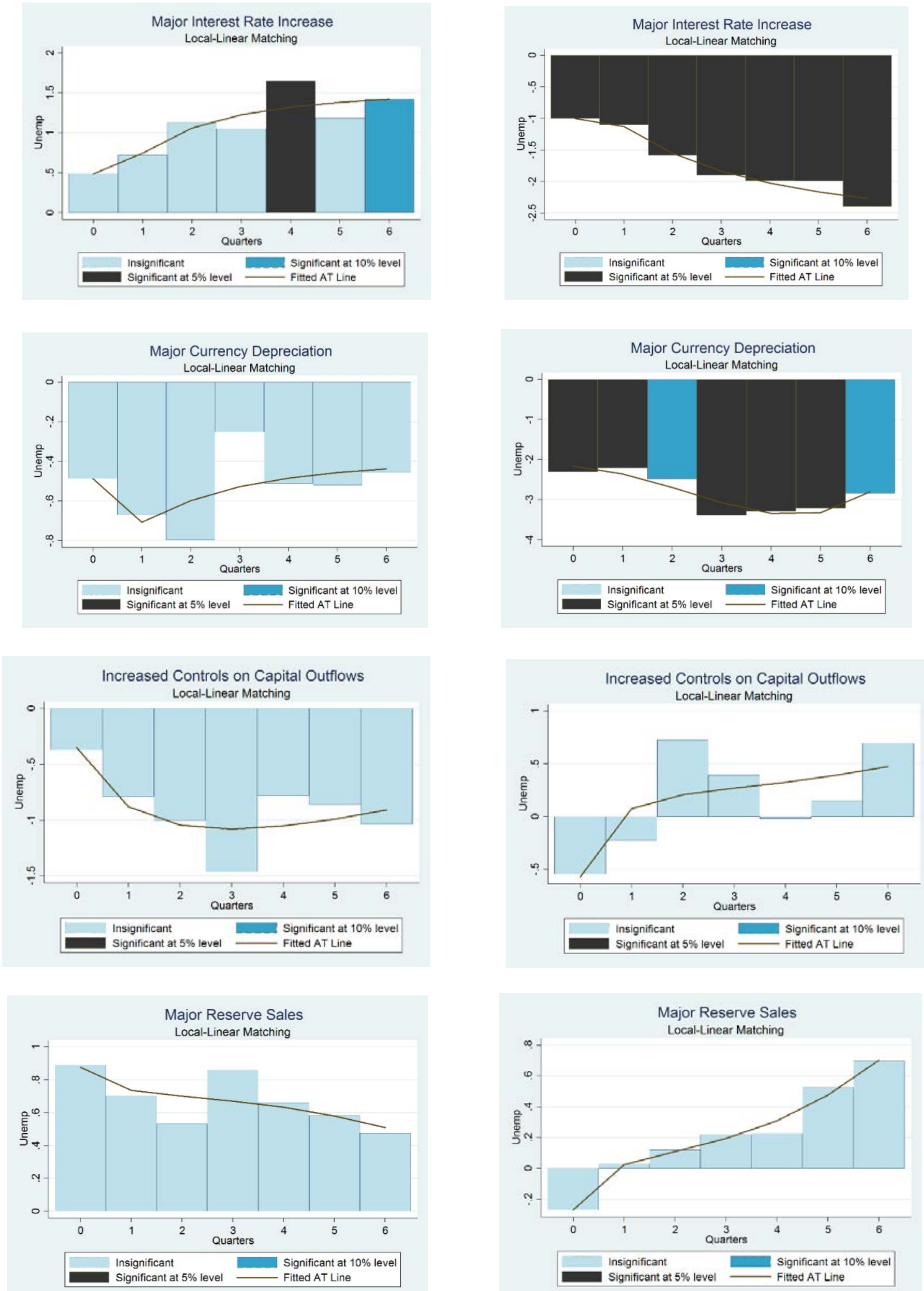
1997-2001 **2007-2011**



Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 6B: Average Treatment Effects for Unemployment using Local Linear Matching

1997-2001 **2007-2011**



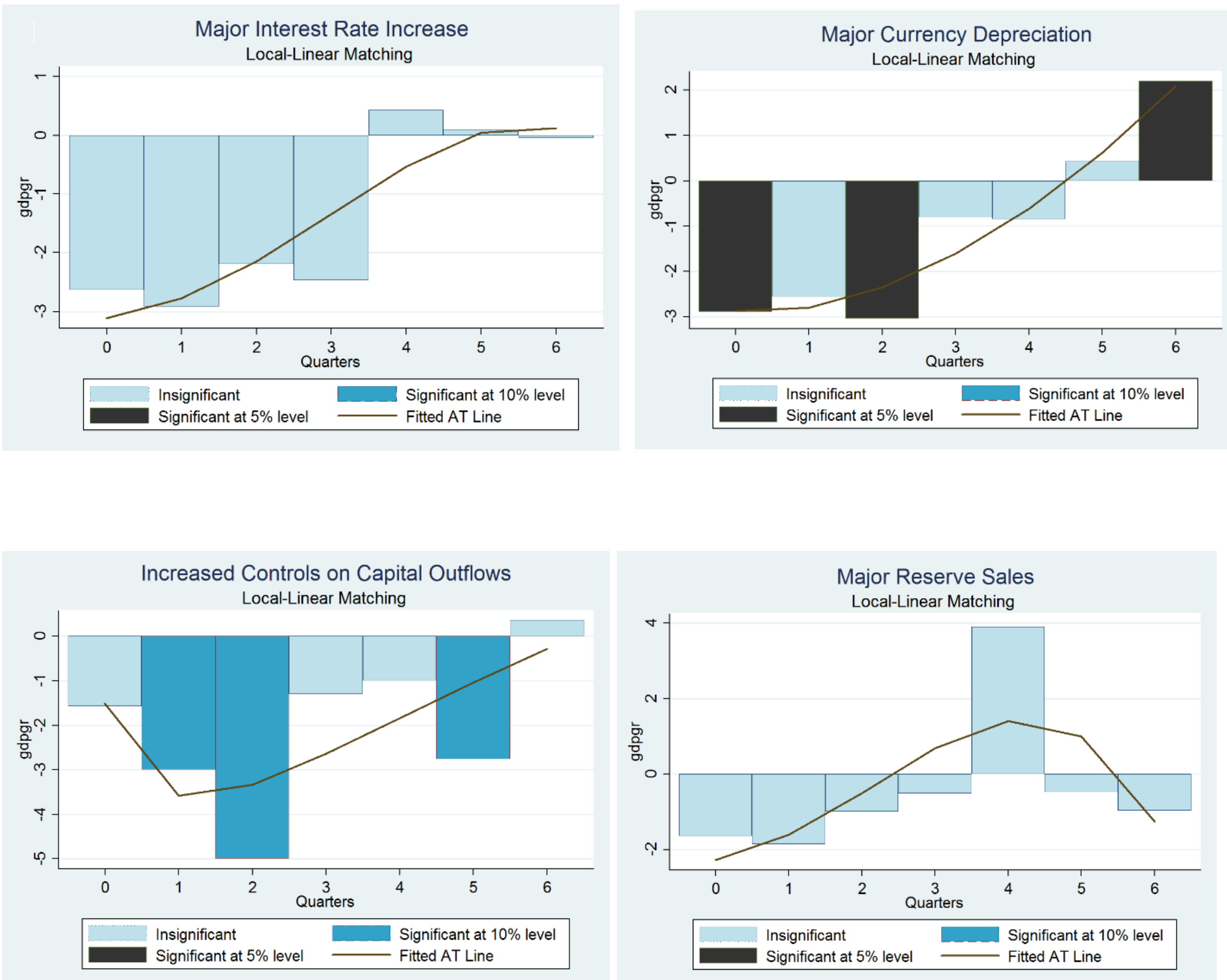
Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 6C: Average Treatment Effects for Inflation using Local Linear Matching

1997-2001 **2007-2011**

Note: ATE measured as the change in the relevant quarter relative to the two-year average before the treatment occurred.

Figure 7: ATT of Real GDP Growth, Not Relative to Average Pre-Treatment values



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. Millions of U.S.$. (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	Global Financial Data,

	using prices for a range of options on the S&P 100 index.	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 U.S.$; expressed as share of GDP. Capital outflows are reported with a positive sign (unlike BOP accounting). Data for 1995-2005 uses the old balance-of-payments definitions, so there is a series break between the two crisis periods in order to avoid inconsistencies in the changes calculated over time. (Q)	IMF, BOP as of 09/13 for data from 2005-2012; Forbes and Warnock (2012) for data from 1995-2003.
Current account balance	Current account balance in millions of U.S.$, 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.