ECONOMICS AND POLITICS 0954-1985
Volume 13 July 2001 No. 2

DO CRISES INDUCE REFORM? SIMPLE EMPIRICAL TESTS OF CONVENTIONAL WISDOM

ALLAN DRAZEN AND WILLIAM EASTERLY[1]

We find evidence for the crisis-induces-reform hypothesis at extreme values of the inflation rate and the black market premium. Episodes of extremely high inflation or black market premiums are followed by periods of better performance than episodes of moderately high inflation or black market premiums. We fail to find similar evidence of the crisis hypothesis when crisis is measured as a high current account deficit, a high budget deficit, or a negative per capita growth rate. The pattern of foreign aid disbursements may help explain the results. Foreign aid is reduced at extreme values of inflation or the black market premium, while it is actually increased for more extreme values of the current account deficit and the budget deficit.

1. INTRODUCTION

THE HYPOTHESIS that an economic crisis often induces policy change (or, even stronger, that crisis is not only sufficient to induce reform but also necessary; that is, that significant reform takes place only in the wake of a crisis) has become, in the eyes of many, the new orthodoxy.[2] As Tommasi and Velasco (1996, p. 197) write,

> That economic crises seem either to facilitate or outright cause economic reforms is part of the new conventional wisdom on reform.

On the other hand, it has been argued, that, like much of conventional wisdom, this view is a tautology. Rodrik (1996, p. 27) puts the problem clearly:

> Reform naturally becomes an issue only when policies are perceived not to be working. A crisis is just an extreme case of policy failure. That reform should follow crisis, then, is no more surprising than smoke following fire.

Though the view that reform follows crisis is widely accepted in some form (Rodrik's point is not that the statement is incorrect, only that it is tautological), it has been subject to almost no empirical testing. The absence of empirical testing may reflect several points. First of all, once an argument enters the realm

[1] We would like to thank John McLaren and two anonymous referees for helpful comments, and participants in seminars at the World Bank and Inter-American Development Bank. Views expressed here should not necessarily be taken as those of the World Bank. Corresponding author: William Easterly, MSN MC3–301, World Bank, 1818 H St NW, Washington DC 20433, USA.

[2] See, for example, the articles in Nelson (1990) or Williamson (1994).

of "conventional wisdom," the belief is that there is no further need to subject it to empirical testing. Second, Rodrik's argument that the hypothesis is tautological implies that there is nothing to test. In the next section, we argue that once one carefully distinguishes between different versions of the argument that "reform follows crisis," the most sensible version is both non-tautological and testable.

Rodrik also suggests another reason why the hypothesis has not been empirically tested – not that it is trivial, but that it cannot be rejected:

> . . . the hypothesis is virtually non-falsifiable: if an economy in crisis has not yet reformed, the frequently proffered explanation is that the crisis has not yet become "severe enough".

This statement contains a good bit of truth. However, as with the argument that it is tautological that reform follows crisis, the validity of this argument depends very much on what one means by the statement that "crisis follows reform." Whether one can learn anything from empirical testing of the crisis hypothesis depends, first of all, on carefully specifying what is the hypothesis to be tested.

The plan of the paper is as follows. In the next section, we define what we mean by the crisis hypothesis with an eye towards addressing the above questions. In section 3, we summarize some underlying theoretical arguments on the relation between crisis and reform. The fourth section, the heart of the paper, contains the empirical tests. We find that the behavior of inflation and the black market premium is consistent with the crisis hypothesis, while the behavior of the current account deficit (including foreign aid as financing), the fiscal deficit (both as a percent of GDP), and growth rates does not give support to the crisis hypothesis according to our simple test. (When foreign aid is excluded – i.e. treated as income – the current account deficit shows evidence that is weakly consistent with the crisis hypothesis.) In section 5 we consider some explanations for these difference between variables. The final section concludes.

2. WHAT IS THE "CRISIS HYPOTHESIS"?

What exactly does it mean to say that "reform follows crisis"? To see that the crisis hypothesis is not trivial in the sense of "smoke following fire," one must distinguish between two related, but distinct, "crisis arguments." One is that reform is induced by bad rather than good times. The other is that things need to get *very* bad (and not just bad) to induce reform. Variants of both arguments have been taken as representing the crisis hypothesis, though in our opinion the term should be reserved for the latter argument. To clarify why, let us examine the first crisis argument more closely.

In discussing whether reform is more likely in good or bad times, one should first differentiate when the need for policy change is *perceived* or *decided upon* from when the change is actually *undertaken*, realizing that "reform follows

crisis'' concerns the former. For example, the first quote from Rodrik above correctly concerns the perception that a policy change is necessary. Once the decision to enact a policy change is made (where this includes the removal of the political constraints that previously blocked an agreement to change policy), the timing of the actual implementation of the reform will depend on a number of factors, including the urgency of the economic situation. Though failure may be necessary to realize the need for change, this change need not *occur* in times of crisis, but may be postponed till more favorable times.[3] Hence, a ''reform follows crisis'' argument concerns the decision to enact a reform rather than the timing of its implementation.

If one then defines ''bad'' times as those when policies are not working, then the first crisis argument given at the beginning of this section is indeed largely tautological. The decision to undertake a policy change aimed at a large improvement generally requires that the situation is in need of improvement. (''Generally'' because it is conceivable that policy-makers take advantage of special circumstances to significantly further improve an already good situation. It is conceivable, but uncommon.) More generally, simple mean reversion in the crisis indicator would lead to a negative outcome being followed by a better one.

In contrast, the second crisis argument is definitely not tautological. To say that bad times are necessary to perceive the need to change policy does not, in itself, answer the question of why the economic situation often needs to get *very* bad, that is, why there must be a ''crisis'', before a reform is adopted? Why is it ''business as usual'' until times get really bad? It is the *magnitude* of the deterioration in the status quo which makes the argument non-trivial. Hence, we define the ''crisis hypothesis'' as follows: why do times need to get *very* bad (and not just bad) to induce reform? Stating it in this way makes clear that the hypothesis is *not* a tautology. That smoke follows fire is not surprising; that small or medium-size fires produce no smoke, while large fires do, would be surprising, suggesting that there is indeed something to be explained.

One is then led to Rodrik's second objection, namely that if one argues that reform follows severe deterioration, one runs the risk of formulating a hypothesis that cannot be rejected. To say that deterioration must be ''severe enough'' does in fact make the argument non-falsifiable: crisis is certainly a *sufficient* condition for reform, if crisis is defined as a deterioration severe enough that reform will necessarily follow.[4] However, if we interpret the crisis hypothesis as arguing that reform follows (or is much more likely to follow) only *extremely* bad economic situations, but not ''medium'' bad situations, then the hypothesis is both non-trivial and falsifiable.

This characterization of the crisis hypothesis is of more than just conceptual interest; our empirical tests flow straight from it. Rather than testing a specific

[3] Though I realize that my roof leaks only if it rains, do I fix it immediately, or wait till the next sunny day?

[4] To the extent that one argues that a crisis is *necessary* to induce a reform, the hypothesis is too easily falsifiable (unless one similarly defines any situation leading to reform as a crisis).

theory or mechanism, our aim is to test directly whether very bad economic performance is more likely to lead to a significant improvement than will less bad performance. To preview our results, the answer will depend on which indicators are used to measure economic performance. The crisis hypothesis as we have formulated it receives strong support if we consider the behavior of inflation or the black market premium, and less support if we consider the current account, government budget deficits, or the growth rate of GDP. Before turning to the empirical results, we quickly summarize some theories underlying crisis hypotheses in order to see where these differences may come from.

3. THEORIES OF CRISIS AND REFORM

In this section, we summarize some basic theories of the relation between crisis and reform. The point of this section is not to be a summary of the literature, but to highlight some arguments that may help us better understand the differences we find in the empirical behavior of macroeconomic variables we find in the next section. Why is extreme deterioration associated with improvement for some variables but not others? A more complete summary of theories of crisis and reform may be found in chapters 10 and 12 of Drazen (2000), on which this section draws.

One argument on why a crisis may be necessary to induce significant reform concentrates on *perception* of the need for reform. Only when things get quite bad do we realize there is a permanent problem that whose solution requires a major policy change, rather than a problem that is transitory (and hence, requires no major action) or one that can be solved without far-reaching policy changes. The perception argument is not simply that the current situation is unacceptable, but that deterioration makes clear the need for a major policy change in a way that less extreme deterioration would not. This may be a simple salience argument (e.g. Akerlof, 1991). It may be that severe dislocations induce practitioners to change their view of the world (e.g. Harberger, 1993). By significantly changing priors of individuals, groups, and policy-makers, a crisis may make them more likely to overcome ideological skepticism and accept proposed policy changes.

Perception by itself is most likely to provide the link between crisis and reform when it is relatively easy both to agree on what sort of change is needed and to effect the change. Hyperinflation may be needed to realize the money supply process is out of control, but once the realization has been made, correcting the problem presents no economic problems and often few political problems as well. Severe social disturbances may be necessary for a widespread perception that wealth is distributed too unequally for the social good; in contrast to hyperinflation, however, correcting the problem is another story entirely. Differences across indicators in the effect of economic deterioration may reflect differences in the possibility of affecting those variables via policy changes.

Differences across indicators in the effect of economic deterioration may also be due to differences in costs that such deterioration implies. The higher the cost

associated with deterioration, the more likely a crisis is to induce a reform. For example, very high inflation may impose costs far greater than a high current account deficit, which in the short run means expenditures higher than output. The former would generate far more pressure for change than the latter.

These two arguments may be combined into one when we consider the *net* cost of reforming, that is the cost of changing policy relative to remaining with the status quo. Another way of saying that it is easier to affect some variables (like inflation) via feasible policy changes than others (like severe poverty) is that the costs of doing so are lower. Hence, hyperinflation may exhibit a pattern consistent with the hypothesis that crisis induces reform either because the costs of reducing hyperinflation are low and because the benefits are high (that is, the cost of remaining with the status quo are high).

Differences in the effect of economic deterioration may also reflect differences in even agreeing on what needs to be done, due to the nature of the reform. Rodrik (1996), for example, argues that macroeconomic stabilization, on the one hand, and liberalization and structural (i.e., microeconomic) reform, often lumped together in discussion of reform, should be seen as distinct. Among other things, there is far less consensus on what constitutes appropriate structural reform than on the need for macroeconomic stability. Lack of consensus means the perception that current policies do not work does not easily translate into the enactment of new policies.

A second argument for the role of crisis in enabling reform follows the models of uncertainty about individual benefits, pioneered by Fernandez and Rodrik (1991). In Labán and Sturzenegger (1994), socially beneficial policy change is delayed by high uncertainty about who will be the winners and who the losers from the change. What made reform possible was a sufficiently large deterioration of the status quo. The greater the uncertainty about the post-reform environment, the greater the deterioration of the status quo must be to induce acceptance of uncertain reform. Hence, a crisis, defined as an extremely large deterioration of the status quo may be necessary for reform.

Two other relevant theories of the relation between crisis and reform introduce political behavior of interest groups more explicitly. One argument focuses on the ability of powerful vested interests to block policy change. According to Olson (1982), economic success creates powerful groups with vested interests who may naturally be against further policy changes which would be to their detriment. Their ability to block socially beneficial reform leads to the society becoming "sclerotic," as Olson terms it. If reform requires a significant weakening of the power of some interest groups, only a severe economic deterioration may be sufficient to weaken their power and bring about reform. A crisis is thus necessary to "reshuffle" interest groups in sclerotic societies. Arguments along these lines are common in the reform literature; see, for example, Nelson (1990, 1994).

Another type of model focuses not on the dissolution of powerful interest groups, but on their willingness to suspend their selfish interests and act in ways

that are socially beneficial. When the benefit from a policy change has the aspect of a public good as in the war of attrition model of Alesina and Drazen (1991), each individual perceives the possibility of large gain at small cost if someone else acts. There will thus be a large incentive to inaction. Given the large gap between the gain from acting and from someone else acting, a large deterioration of the status quo may be necessary to induce an interest group to accept a disproportionate part of the cost of a policy change and thus enable a reform to be enacted.

This result forms the basis of the argument of Drazen and Grilli (1993) on the benefit of crises in inducing policy reform. Since higher pre-reform distortions induce an earlier expected reform, the government, by inducing a crisis, can hasten the expected date of reform. Intuitively, if what is blocking a reform is the inability to gain consensus on how the burden of reform is to be divided among interest groups, a crisis can hasten agreement by increasing the distortion associated with the status quo, thus raising the cost of not agreeing to reform. A crisis will make each interest group more amenable to reform, and hence shorten the expected delay in adopting a reform. In fact, as Drazen and Grilli show, if the expected date of reform is significantly brought forward, lower pre-reform utility implies that expected welfare can actually rise due to the crisis.[5]

This argument should be seen as complementary to the first theory on the role of perceptions of the need for a change in policy. A crisis may not so much induce interest groups to see the world differently in intellectual terms, but to realize that their political interactions must change. As Williamson (1994, p. 19) puts the point,

> . . . a sufficiently acute crisis may also create a consensus that the old order has failed and needs to be replaced, leading individuals and groups to accept that their special interests need to be sacrificed (along with those of other special interest groups) on the altar of the general good.

Like arguments based on perceptions of the need to reform, more explicitly political theories suggest that a severe economic deterioration need not have the same effect on all economic variables in terms of their subsequent performance. Even controlling for the severity of a crisis, the effect of crisis in inducing reform depends critically on what sort of reform is being considered. For example, Nelson (1994) has argued that severe crisis

> generates a strong popular desire for a take-charge government with a plausible plan to contain the emergency. . . . But it is much less obvious that crisis generates acceptance of permanent structural reforms. Whether it does so depends in good part on the perceived causes of the crisis.

[5] This argument depends on the specific role that a crisis plays in a war-of-attrition model and should not be taken to imply that crises increase welfare in general in economies where necessary reforms are delayed.

Moreover, though the causes of a crisis (mismanagement, corruption, temporary shocks) very much dictate the response to the crisis, the nature of reform is crucial as well. Some reforms are induced by crisis that simply creates the perception of the value of a change. Others require sufficient severity to induce significant restructuring. Still others may not be induced by even very severe crises, not because of the lack of a sufficiently severe crisis, but because of the difficulty of the reform itself.

Foreign assistance complicates the picture. The relation between crisis and reform may depend on the effect of crisis on the receipt of foreign assistance. Aid that encourages reform may be more likely given in times of crisis. Aid may also be conditional on reform. Conversely, aid may soften the effect of a crisis and thereby make reform less likely. Moreover, the effectiveness of aid in inducing reform may depend on a country first passing through a crisis. The theoretical literature contains models where aid may either hasten or delay reform. In general, unconditional aid is seen as delaying reform, as in the work of Fernandez-Arias (1997), Svensson (1999a, 1996b), and Drazen (1999). This work suggests that economic deterioration that induces aid flows may not induce reform.

4. EMPIRICAL RESULTS

As indicated in the Introduction, empirical studies of the crisis hypothesis are rare. The basic problem is in finding measures of crisis and reform (especially the former) and designing a formal empirical test in a way that leads to convincing results. Our approach begins by defining a ''crisis'' as extremely low macroeconomic performance, as measured by different macroeconomic indicators relative to some reference level. Our simple test of the relation between ''crisis'' and ''reform'' is to consider the effect of poor macroeconomic performance at date t (the ''crisis'') on performance at some later date, taken to indicate the presence or absence of reform.

Our approach follows Bruno and Easterly (1996). They find that, measuring ''reform'' by subsequent inflation performance, high-inflation countries are more likely to undertake stabilization than moderate-inflation countries. That is, the correlation between inflation today and inflation tomorrow is not monotonically positive, but turns negative for high inflation. At high enough inflation (150–200 percent per year), there is a *negative* relation between current and inflation lagged five years. They take this as evidence of the crisis hypothesis, because it says that countries with extremely high inflation today will have lower inflation tomorrow than those countries in which inflation is only moderately high.[6]

We extend the Bruno–Easterly results in two main ways – we both consider a wider set of variables; and, for a given variable, we consider alternative ways of

[6] Bruno and Easterly (1996) found a possibly related result that growth after inflation crises (defined as 40+ annual inflation) is higher after the crisis ends than before the crisis.

how to define a reference level relative to which a crisis is measured. As in Bruno and Easterly, the basic idea is to look at an indicator of macroeconomic performance at t and at subsequent dates $t+s$ for $s>0$. Normally an indicator will be positively correlated across periods because policy indicators are persistent over time. The crisis hypothesis is represented by extremely poor performance at t implying not simply good performance at $t+s$, but *better* performance at $t+s$ than if performance at t was just moderately bad. If I_{it} is the crisis indicator in country i, then we interpret the crisis hypothesis as saying "for some extreme range of I_{it}, $I_{it} > I_{jt}$ implies $I_{it+s} < I_{jt+s}$." For ease of later exposition, we call this a "crossover."

The indicators that we use are inflation, the black market premium, the growth rate of GDP per capita, the current account deficit, and the public sector deficit. The results are striking in the two general patterns that we see: inflation and the black market premium illustrate behavior consistent with the crisis hypothesis, while growth rates, fiscal balance, and the current account balance do not (although we will see later that the current account balance excluding aid financing is mildly consistent with the crisis hypothesis). These results are robust to how the variables are measured.

4.1 Inflation

We begin with inflation. The data consist of yearly rates of CPI inflation over the period 1952–1996 for a sample of 156 countries. We consider two ways of organizing relative inflation performance. The first is to divide inflation episodes into percentiles according to the inflation rate in the country at t–5 and to consider median inflation in each percentile group five years later. Each percentile contains 42 observations. These results are presented in Table 1 and Figure 1, where we consider the 90th percentile (corresponding to median

TABLE 1 WORST INFLATION FIVE YEARS AGO COMPARED TO TODAY'S INFLATION, 1953–1996

Percentile of inflation, lagged five years	Median inflation, lagged five years	Median inflation, today
90	27	23
91	29	22
92	32	24
93	36	32
94	45	36
95	55	43
96	68	70
97	94	45
98	133	32
99	591	25

Note: Each percentile contains 42 observations.

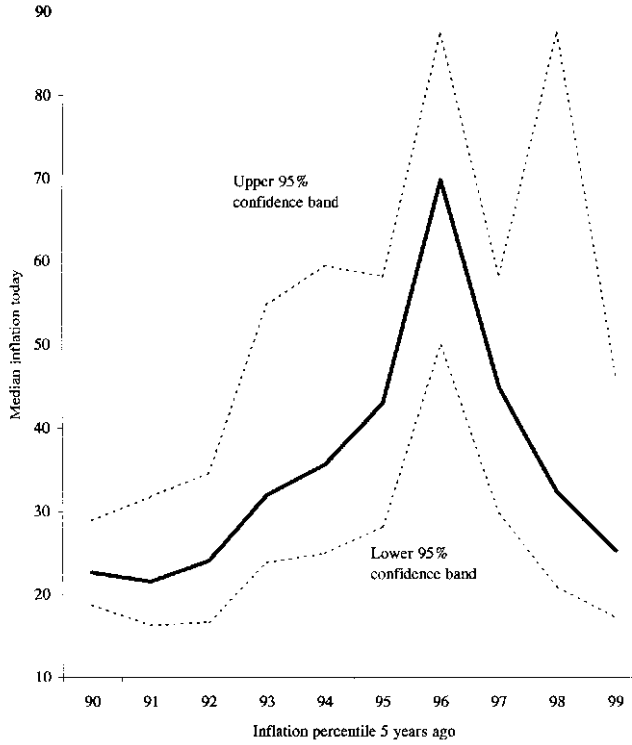

Figure 1. Median inflation today is decreasing in the percentile of inflation five years ago in the upper percentiles.

inflation of 27 percent per year) and above. We find that inflation today is increasing in inflation five years ago until one reaches the 96th percentile (which corresponds to median annual inflation of 68 percent). Above this level, inflation is decreasing in the level of inflation five years ago. The results on the 99th percentile (median inflation 591 percent) are particularly striking: the confidence band is fairly narrow and lies entirely below the confidence band on the peak at the 96th percentile. In other words, having 628 percent inflation today is likely to lead to lower inflation five years from now (21 percent) than is having 72 percent inflation (which predicts 66 percent inflation five years from now). These results echo those of Bruno and Easterly (1996) and support the crisis hypothesis.

Another way of organizing the data is to divide inflation episodes into a small number of groups and trace inflation performance over subsequent years. We divide inflation episodes in countries in our data into three groups: inflation from 40–100 percent in some years, 100–1,000 percent, and over 1,000 percent and then look at inflation performance in years since inflation first passed the respective critical level. Since there are only six episodes of 1,000+ inflation in the data 1952–1996 (excluding transition countries, on which more in a

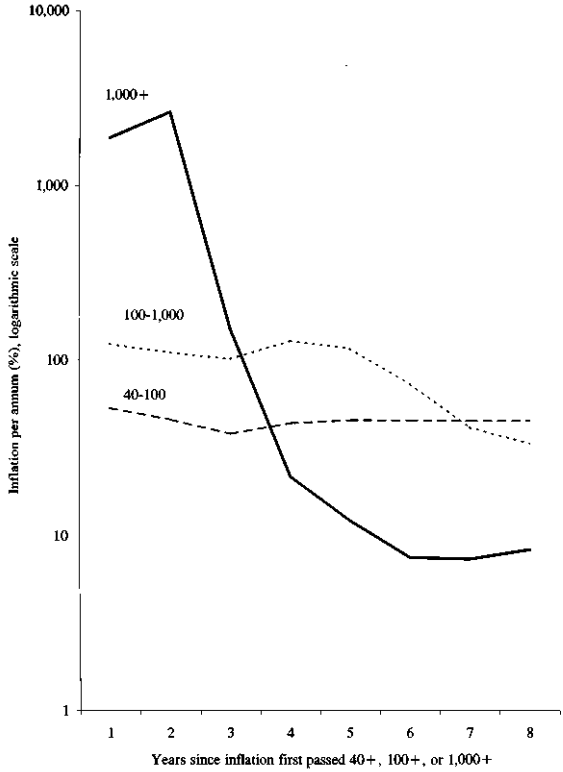

Figure 2a. Modern and historical cases of subsequent behavior of inflation after it reaches 1,000+, 100+, and 40+.

moment), we supplement the data with 12 historical episodes of 1,000+ inflation.[7] Looking at median inflation in each group (Figure 2a), we find that the median rate of inflation for countries in the first (initial inflation of 40–100 percent) group did not show a significant change over the subsequent five years. The median rate of inflation for countries in the second (initial inflation of 100–1,000 percent) group fell markedly over the subsequent five years, to a level that was insignificantly different from that of the first group. Median inflation for countries in the third (initial inflation above 1,000 percent) group, however, fell far more, to about 15 percent median inflation rate after five years and about 10 percent after eight years. After five years, the median inflation rate in the group with initial inflation above 1,000 percent was 33 percentage points *lower* than in the group with initial inflation between 100 and 1,000 percent, a difference that is statistically significant. This is our first example of a "crossover."

[7] The six modern countries are Argentina, Bolivia, Brazil, Indonesia, Nicaragua, and Peru. The historical episodes and start dates of 1,000+ inflation are: Austria 1922, China 1947, Germany 1922, Greece 1943, Hungary (1923, and again in 1945), Korea 1945, Poland 1923, Romania 1947, Russia 1919, and Taiwan 1948.

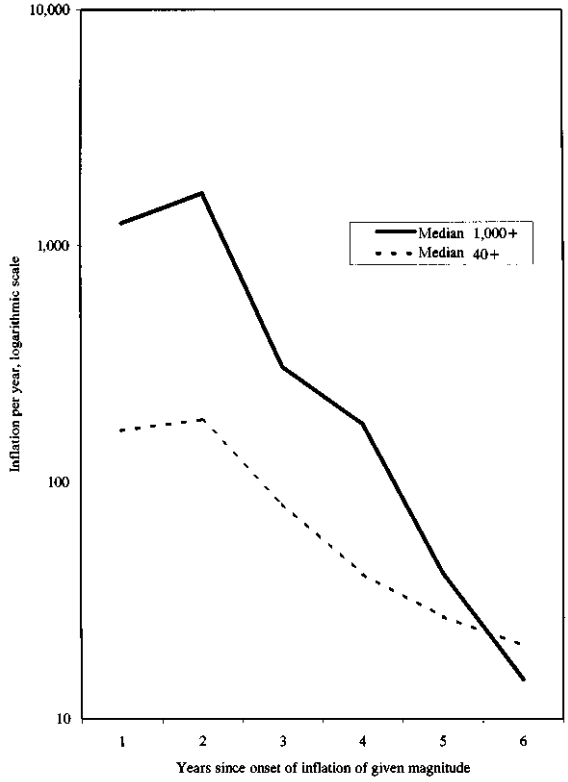

Figure 2b. Extreme and moderate inflations in transition countries.

We believe the transition countries should be considered separately, because of the special nature of their transition from planned economy to market. We use data on yearly inflation from 1989 to 1998 in 25 transition economies in Eastern Europe and the former Soviet Union (FSU). The pattern in the transition countries supports the idea that extreme inflation soon comes down to a low level, although the evidence for crossover is not as clear as it was in Figure 2a. Figure 2b shows the median inflation in 15 transition economies where inflation reached over 1,000 percent (mostly, but not entirely, FSU economies). This inflation had come down after six years to slightly below that of the 10 transition countries where inflation topped 40 percent but never topped 1,000 (mostly, but not entirely, Eastern Europe).

In addition to looking simply at absolute inflation performance over time, we considered inflation measured at a point of time relative to the country average over the sample and the global average in that year. On the basis of these observations of mean normalized log inflation, we considered an inflation ''crisis'' defined as an observation two standard deviations or more above the mean. Grouping inflation observations in this way, we then traced the standard

deviation of log inflation from its mean in prior and subsequent years. We compare the evolution of inflation to episodes in which inflation was initially 1.5, 1, or 0.5 standard deviations from the country and global mean (Figure 3). Organizing the data in this way tells a similar, though perhaps less clear, story. In terms of deviations from mean inflation, inflation 2 or more standard deviations above the mean falls rapidly (as does inflation at 1 and 1.5 standard deviations above the mean), falling below the mean after about seven years. There is not much of a "crossover" however. It is interesting that the evidence for "crossover" is much stronger with absolute inflation crises than it is with relative inflation crises.

How do other crisis variables affect inflation persistence? It could be that having another crisis indicator at a high level increases the pressure to stabilize. We examine the top decile of inflation lagged five years. We calculate the median black market premium for this sample (29 percent) and divide the sample into "high black market premium" (>29) and "low black market premium" (<29). Figure 4 shows the graph of inflation today versus inflation five years ago for the two samples. We see the same pattern as in Figure 1 for both samples, i.e. a hump shape where extreme inflation five years ago is associated with lower inflation today than is a less extreme level of inflation five years ago.

4.2. *Black Market Premium*

We next consider the black market premium, organizing the data as above both

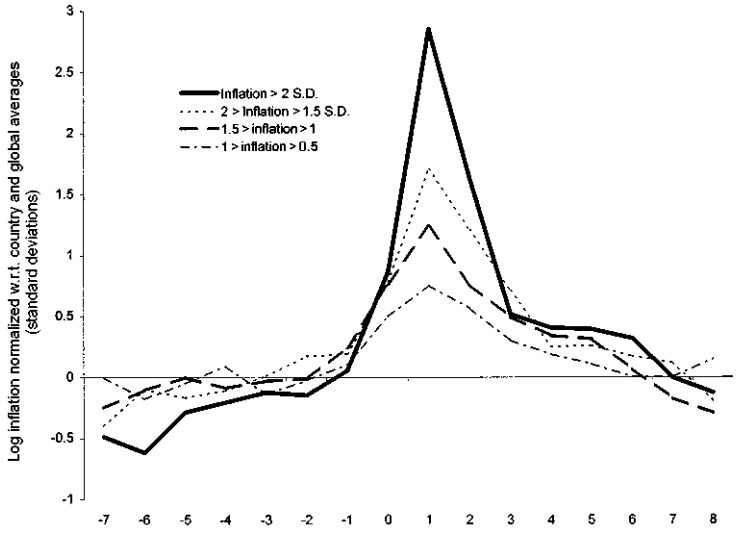

Figure 3. Mean normalized inflation with inflation crises at different standard deviations in year 1.

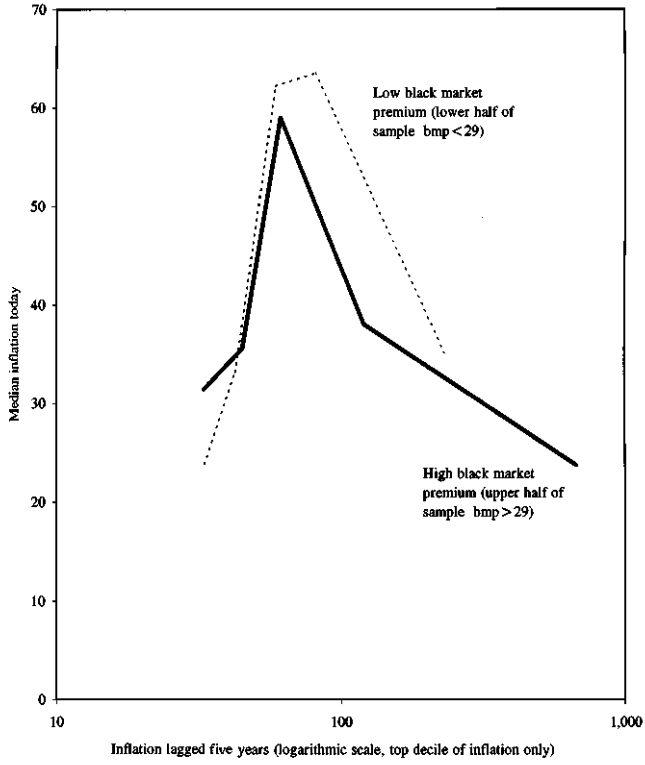

Figure 4. Inflation persistence at high and low values of black market premium.

by percentiles and into smaller groups. We use annual observations on the black market premium for a sample of 127 countries over the period 1960–1993. The behavior of the black market premium is similar to the behavior of inflation, exhibiting crossovers consistent with the crisis hypothesis.

Consider first the division of episodes of high black market premiums into percentiles according to the black market premium at t–5. We compare the median black market premium at t–5 with the median premium at t in each group. These results are presented in Table 2 and Figure 5, where we consider the 90th percentile (corresponding to median black market premium of 112 percent) and above. As with the inflation data, there is a crossover. The 99th percentile has a median black market rate after five years below that of a lower percentile (the peak is at the 96th percentile). The confidence bands are wide for the other percentiles besides the 99th (and the 97th), although the median at the 99th percentile still lies below the confidence band of the 96th percentile. The 96th percentile corresponds to a black market premium of 309 percent, which predicts a black market premium in five years of 317 percent. The 99th percentile corresponds to a black market premium of 1,038 percent (median

TABLE 2 WORST BLACK MARKET PREMIUM FIVE YEARS AGO COMPARED TO TODAY'S, 1965–1993

Percentile of black market premium, lagged five years	Median black market premium in percentile, lagged five years	Median black market premium, today
90	112	80
91	131	117
92	150	62
93	191	159
94	224	182
95	268	224
96	309	317
97	388	292
98	601	232
99	1,038	69

Note: Each percentile contains 32 observations.

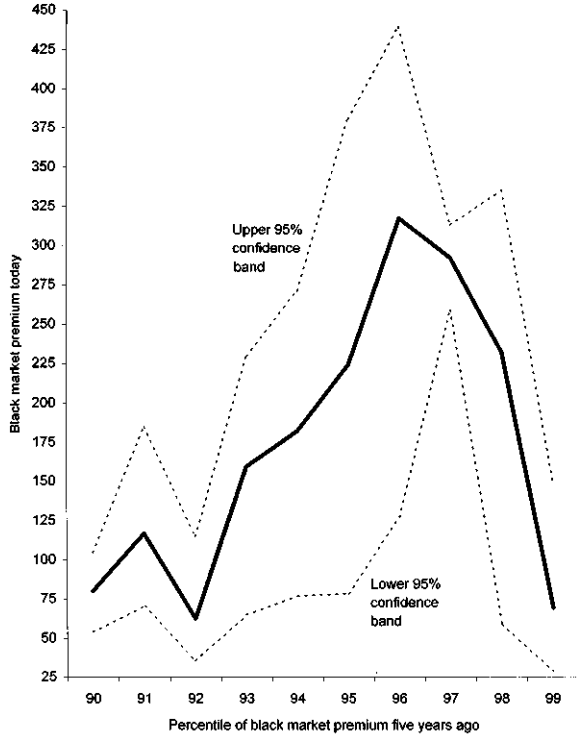

Figure 5. Median black market premium today is decreasing in the percentile of black market premium five years ago in the upper percentiles.

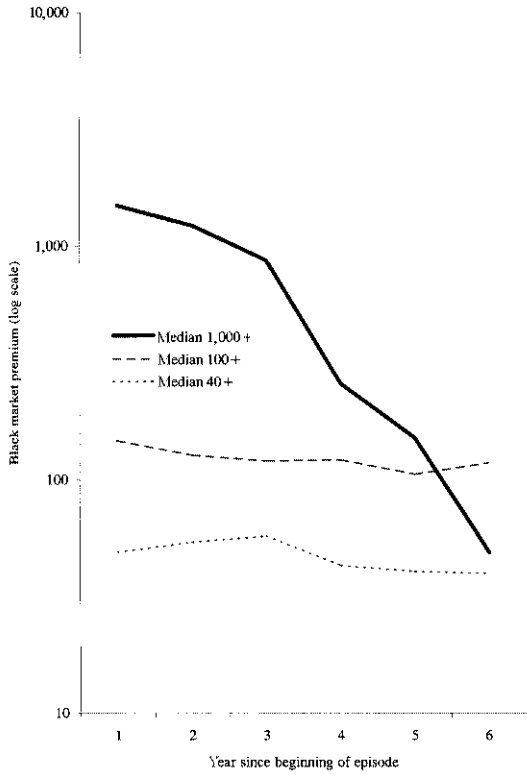

Figure 6. Black market premiums after they pass 1,000+, 100+, and 40+.

for the percentile), which predicts a black market premium in five years of 69 percent.[8]

We also organize the data on episodes of high black market premiums into three groups (a premium between 40 and 100 percent, 100 to 1,000 percent, and above 1,000 percent) and trace the behavior of the median in each group in the six years since the beginning of the episode (Figure 6). Episodes of a black market premium of over 1,000 percent at some date have a median black market premium significantly lower five years later than episodes of a black market premium between 100 and 1,000 percent. This crossover, like that of high inflation, also supports the crisis hypothesis.

4.3. Current Account Balance and Fiscal Deficits

In contrast to the results for inflation and the black market premium, we find no support for the crisis hypothesis for the three other indicators in either sort of

[8] These results, as well as those for inflation, can be taken as indicating that the crisis hypothesis really holds at only extreme levels like the 96th percentile and above of a crisis variable. This means that the crisis hypothesis is not that relevant for the majority of the sample with less extreme crisis indicators.

test. We begin by reviewing the evidence for the current account deficit (inclusive of deficits financed by foreign aid), expressed as a percentage of GDP, with annual data for a sample of 152 countries over the period 1970–1996.

In the percentile test, for example, there is a positive correlation between values at t–5 and five years later (Table 3 and Figure 7). There is mean reversion, but not "crossover." The 99th percentile of current account deficit as a percentage of GDP corresponds to a median deficit of 48 percent, falling to a still enormous median deficit of 34 percent after five years. The relation between the median deficit at t and five years later is fairly strongly monotonic as one moves down to the 90th percentile. Dividing countries into groups of initial current account deficit to GDP of 0 to −5 percent, −5 to −10 percent, and below

TABLE 3 WORST CURRENT ACCOUNT DEFICITS FIVE YEARS AGO COMPARED TO CURRENT ACCOUNT DEFICIT TODAY, 1970–1996

Percentile of current account balance/GDP, lagged five years	Median current account balance/GDP, lagged five years	Median current account balance/GDP, today
90	−13.9	−4.1
91	−14.9	−6.4
92	−16.3	−11.5
93	−17.6	−13.1
94	−19.1	−10.2
95	−20.6	−15.0
96	−23.1	−10.9
97	−26.9	−10.0
98	−35.3	−20.8
99	−48.0	−34.0

Note: Each percentile contains 21 observations.

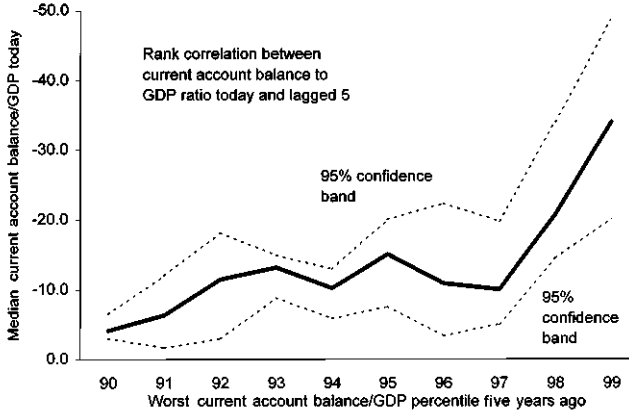

Figure 7. Median current account balance/GDP relative to current account balance/GDP percentile five years ago.

−10 percent displays analogous monotonicity when one considers median values over the eight years since the initial episode (Figure 8). Here too there is mean reversion, but no crossover.

The behavior of the fiscal deficit is qualitatively similar. We consider the public sector balance as a fraction of GDP for a sample of 84 countries over the period 1965–1994 (note that the sample is much smaller with public deficits, since we are using consolidated public sector deficits culled from World Bank and IMF reports). Table 4 and Figure 9 give the percentile results. Figure 10 shows the behavior of the median of episodes for three groups – fiscal deficit of 10 percent or higher, deficit of 5 to 10 percent, and deficit of 0 to 5 percent. In both ways of organizing the data on fiscal deficits, we observe significant mean reversion, but the absence of crossover.

4.4. Growth

As a last variable, we consider the correlation of growth behavior over time, using data on per capita growth for a sample of 169 countries over the period 1961–1996. Table 5 and Figure 11 do not show a clear pattern other than

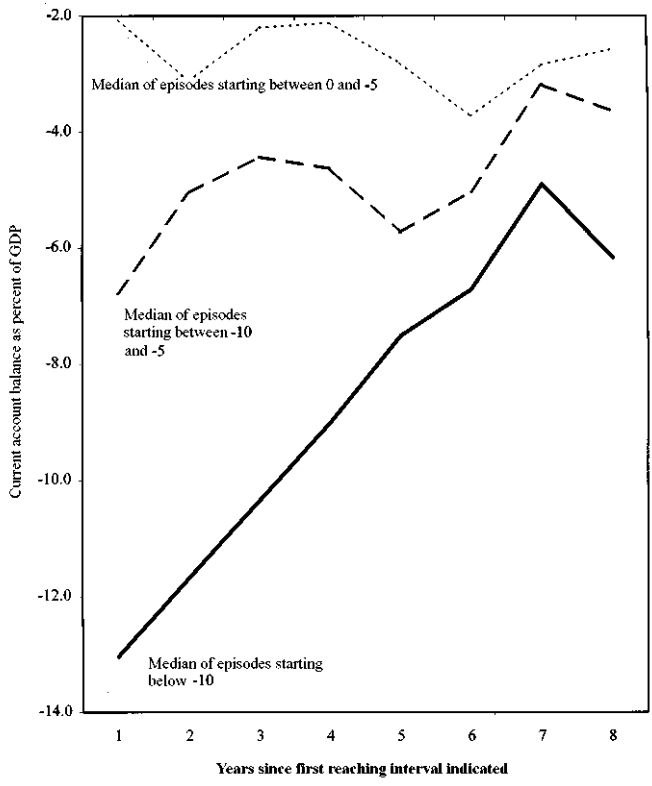

Figure 8. Current account balances, after starting at below −10, below −5, and below 0.

TABLE 4 WORST PUBLIC SECTOR BALANCES FIVE YEARS AGO COMPARED TO TODAY'S PUBLIC SECTOR BALANCE, 1965–1994

Percentile of worst public sector balance/GDP, five years ago	Median public sector balance/GDP five years ago, in percentile	Median public sector balance/GDP, today
90	−12	−6
91	−13	−8
92	−13	−8
93	−14	−7
94	−14	−11
95	−15	−7
96	−16	−9
97	−17	−10
98	−20	−12
99	−28	−16

Note: Each percentile contains 12 observations.

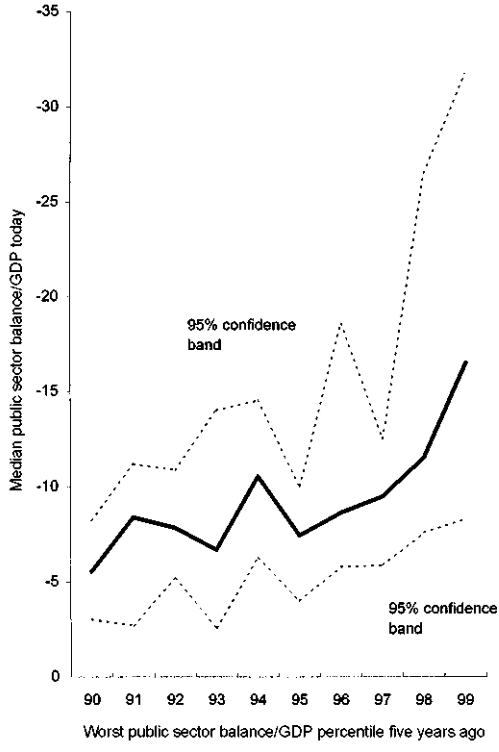

Figure 9. Median public sector balance/GDP relative to worst public sector balance/GDP percentile five years ago.

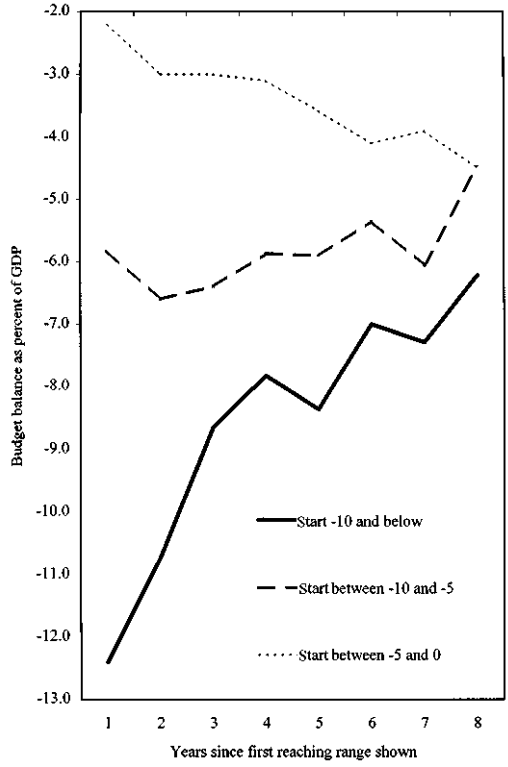

Figure 10. Budget balances after initial level of −10 or below, −10 to −5, and −5 to 0.

TABLE 5 WORST NEGATIVE GROWTH IN PREVIOUS FOUR YEAR PERIOD COMPARED TO CURRENT FOUR-YEAR PERIOD, 1961–1996

Percentile of worst growth per capita, logged four-year period	Median growth, lagged four-year period	Median growth, current four-year period
90	−2.9	−1.2
91	−3.5	0.0
92	−4.0	1.1
93	−4.5	0.3
94	−5.1	−0.7
95	−6.2	−2.7
96	−6.7	1.7
97	−7.8	0.4
98	−9.0	0.8
99	−14.3	2.0

Note: Each percentile contains 10 observations.

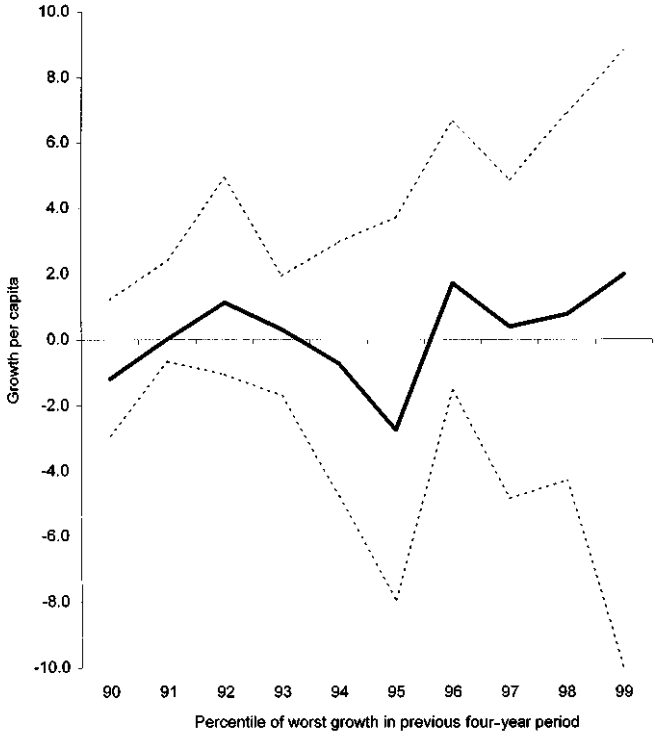

Figure 11. Growth in four-year period against percentile of worst growth in previous four-year period.

universal mean reversion. The worst negative growth outcomes are typically followed in the next four-year period by outcomes of positive growth, but there is extremely wide dispersion of outcomes. The behavior of growth rates after episodes of negative growth is quite interesting, as shown in Figure 12. Dividing countries with episodes of at least two years of negative growth into three groups (0 to -5 percent, -5 to -10 percent, and below -10 percent), we find a clear positive correlation between negative growth at t and median growth six to eight years later. The only difference is that the group of countries with the most negative (-10 percent or below) growth rates during the crisis show marginally better growth performance two to four years after the beginning of the crisis than those with less negative growth rates. However, this difference is not statistically significant and disappears after year four. We fail to find any evidence for crossover of growth outcomes, which means there is little support for the crisis hypothesis when per capita growth is the performance indicator.

5. EXPLAINING THE PATTERNS

Before we consider explanations for these patterns, some comments on the

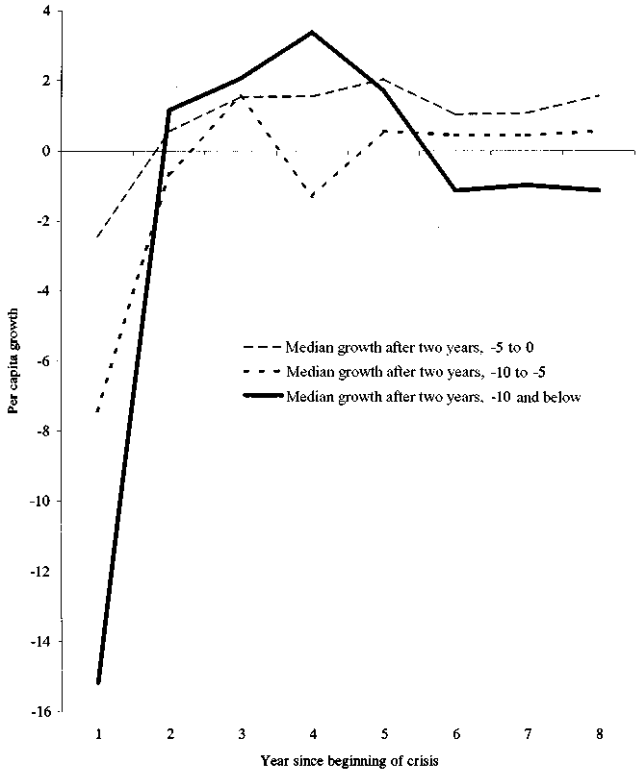

Figure 12. The aftermath of negative growth crises of two years below a given number.

nature of our tests are in order. The stark simplicity of our methodology is both its weakness and its strength. First, in our approach the same variable measures both crisis (when it indicates extremely poor performance) and reform (when there is a significant improvement). Ideally, one would like to measure reform more directly. However, as already indicated, finding more direct measures of reform is far from simple; following this route would so severely restrict our sample size as to make this a very different paper.[9] The virtue of our method is its simplicity, allowing us to look at a large group of countries without having to study the specific reforms that were undertaken in each one of them. The diversity of our results across indicators suggests that there is something to be learned from such a simple approach.

[9] Lora (1998) constructs policy indices in five areas (trade reform, reform of the domestic financial system, tax reform, privatization, and labor market reform) for 19 Latin American and Caribbean countries over the period 1985–1995, and the importance of political and economic factors for their adoption. One explanatory variable is crisis, measured as we do here by macroeconomic indicators. The strongest empirical result is the importance of crisis, consistently significant for the total reform index, no matter what other explanatory variables are included. See also Drazen (2000) for a fuller discussion of Lora's approach.

Second, one may ask whether we have adequately measured crises across such a broad range of countries and across time by our simple percentile technique. What constitutes a crisis will surely not be the same across countries (or across time in one country). Here too, the virtue of our approach is its simplicity. Any approach which is based on making a judgment in each episode as to whether poor macroeconomic performance reached crisis proportions is unworkable for a large sample and is almost sure to fail. One should judge these results not by whether we have really captured the essence of a crisis, but by what one can learn from them.

How then can one explain the difference in the behavior of the indicators? We suggest two, not inconsistent explanations, based on the theoretical work on the relation between crisis and reform discussed in section 3. First, the differing results for the different indicators may indicate a difference in the ease of adjustment of different variables (or, more generally, the net cost of reform, as discussed in section 3). That is, a "crisis" may be likely to lead to reform and economic improvement when the policy measures needed to improve are clear and relatively easy to implement, but have less of an effect for more difficult macroeconomic problems. More specifically, monetary variables, that is the rate of inflation or the black market premium, may be easy to adjust, real variables much more difficult. Reducing a high black market premium, for example, can happen overnight with a devaluation and exchange rate unification.

In contrast, it may be very hard to reduce a current account deficit of 48 percent of GDP to a current account deficit of 5 percent of GDP (or to eliminate a government budget deficit of 28 percent of GDP, the 99th percentile of public sector balance figures). Correcting imbalances that require implementing and sustaining structural reforms is neither easy nor quick. It is difficult to get agreement on structural reform; even when reforms are adopted and continued, results are far from immediate. Nelson (1994, pp. 473–474) puts it well:

> One particularly clear lesson of the 1980s is that basic market-oriented reforms take years to put in place and usually require still more time before they generate robust investment and growth. Many countries have launched promising efforts; far fewer have sustained those efforts long and vigorously enough to produce results. . . .
>
> Initial stabilization and adjustment packages typically include fiscal and monetary austerity measures, devaluation, and varying degrees of price and trade liberalization. These measures are administratively easy, in the sense that they can be decided and put into effect by a small circle of senior economic officials. Measures that are usually introduced later, such as financial sector reforms, the rationalization or privatization of large scale enterprises, opening of labor markets, and restructuring of social services and social security, are much more complex.

A second line of explanation concerns the relation between poor macroeconomic performance and foreign assistance. If poor performance induces higher foreign aid to cushion its effects, it need not correspond to a crisis. To further examine this possibility, we considered the relation between each of our

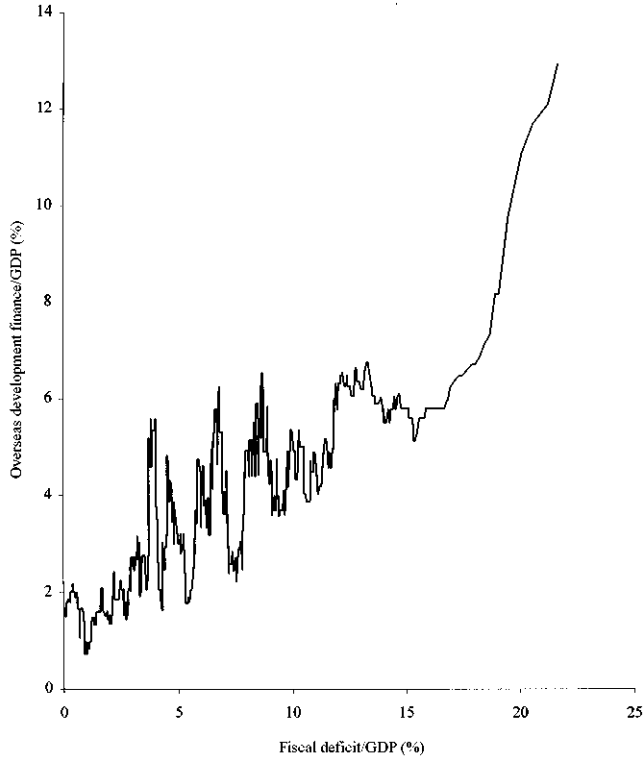

Figure 13. Overseas development finance/GDP and fiscal deficits/GDP.

indicators and Overseas Development Finance (ODF) relative to GDP, for fiscal deficit/GDP, current account deficit/GDP, per capita growth, inflation, and the black market premium, respectively. These diagrams may be understood as follows. We ranked the episodes by values of the indicator variable (for example, fiscal deficit as a percentage of GDP in Figure 13). Then we constructed a rolling median of 40 observations (100 observations for per capita growth rates), so that the first fiscal deficit observation in Figure 13 is the median of observations 1–40 for fiscal deficits and ODF, the second the median of observations 2–41, etc.

For the fiscal or the current account deficit, there is a very high correlation between these two indicators and foreign aid received as a percentage of GDP (Figure 14). For example, as the ratio of the current account deficit to GDP rises from 0 to 15 percent, aid rises almost monotonically from about 1 to about 6 percent of GDP. Countries with higher average current account deficits receive significantly more aid.[10] This suggests that high budget and current account

[10] In comparison, the relation between foreign aid received and per capita growth is far from clear, with no strong evidence that poor growth performance is correlated unambiguously with receiving more aid.

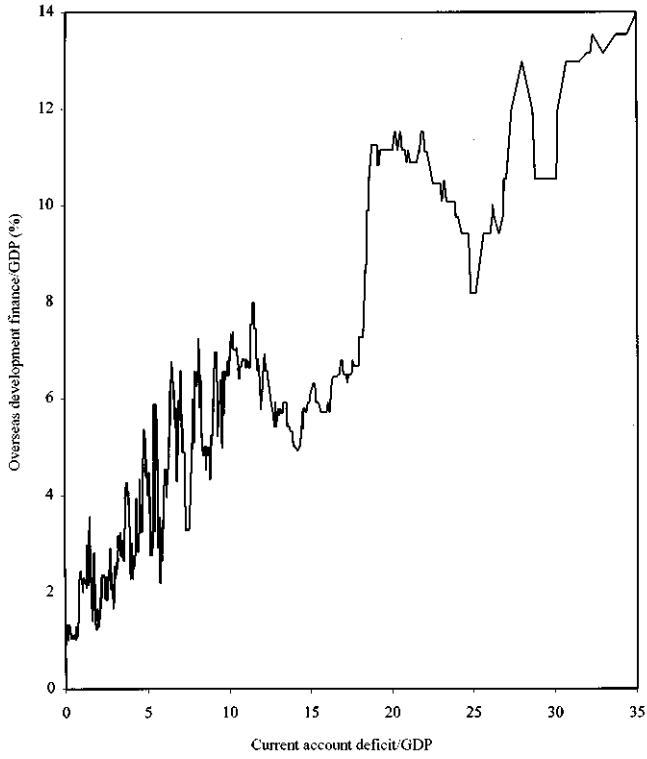

Figure 14. Current account deficits/GDP and ODF/GDP, rolling medians.

deficits may not be perceived as crises, because foreign aid cushions their effects. Hence, the failure of extreme values of deficits to induce reform may simply reflect that these episodes do not correspond to crises to the extent that the basic figures might suggest.

To test this further, we redefine current account deficits and budget deficits as excluding aid (that is, excluding the part of the deficits financed with aid – using data on aid as in Chang et al. 1999, which counts not only conventionally measured grants but also the grant element of foreign loans). The budget deficit excluding aid does not behave differently than the conventionally measured budget deficit – there is still no crossover (Figure 15). The current account deficit excluding aid does show some evidence of "crossover" (Figure 16), although the differences in year 8 are not statistically significant. There is definitely more rapid correction of extreme aid-adjusted current account deficits than of the conventional current account deficit, which supports the "aid delays reform" argument. The difference between aid-adjusted budget deficits and current account deficits may reflect the greater leverage of foreign lenders (for the current account deficit) over the government compared to domestic lenders (for the budget deficit) (Figure 1).

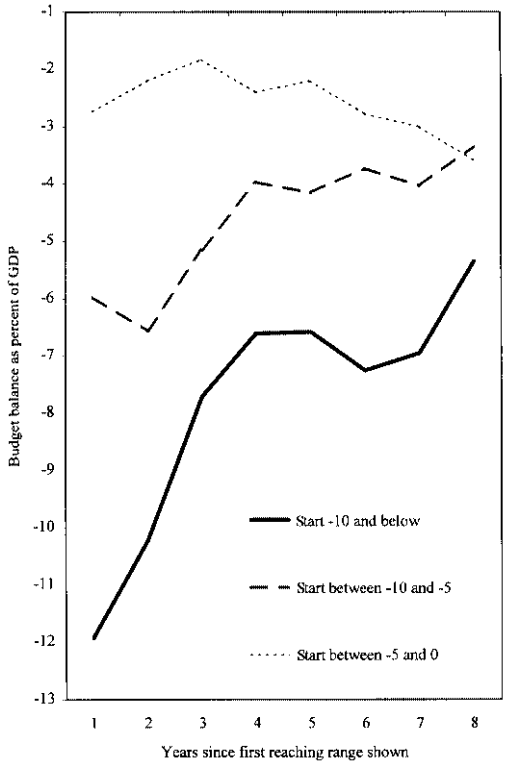

Figure 15. Aid-adjusted budget balances after initial level of −10 or below, −10 to −5, and −5 to 0.

In contrast, there appears to be a negative association between aid received and inflation, at least when inflation gets very large (Figure 18). That is, high inflation induces no increase in aid received and is instead correlated with lower aid, although the decrease is only unambiguous at very high levels of inflation. Unlike the fiscal or current account deficit, inflation crises are not mitigated by aid, giving one reason why there may be a difference in the behavior of the variables. The very high threshold of inflation that must be reached before aid clearly diminishes also helps explain why only very extreme levels of inflation fit the crisis hypothesis.

Likewise, there is a negative relationship between extremely high black market premiums and aid received (Figure 19), which may help explain why an extreme black market premium is a reform-inducing crisis. The continuation of aid to countries with high but not extreme black market premiums may help explain why only the very extreme black market premiums fit the crisis hypothesis.

6. CONCLUSIONS

We have found evidence for a strong version of the crisis-induces-reform

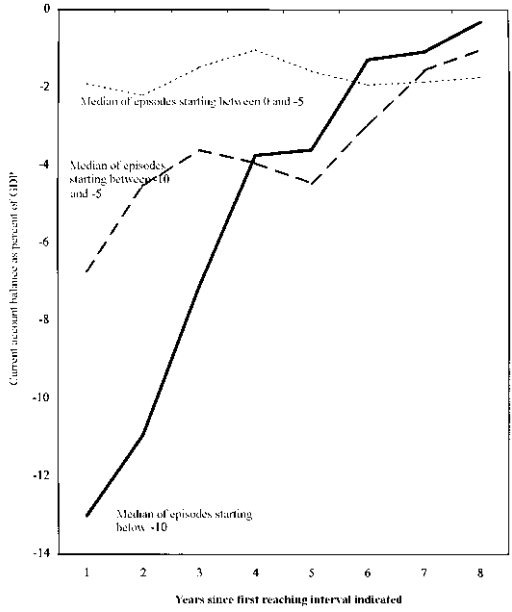

Figure 16. Current account balances excluding aid after starting at below −10, below −5, and below 0.

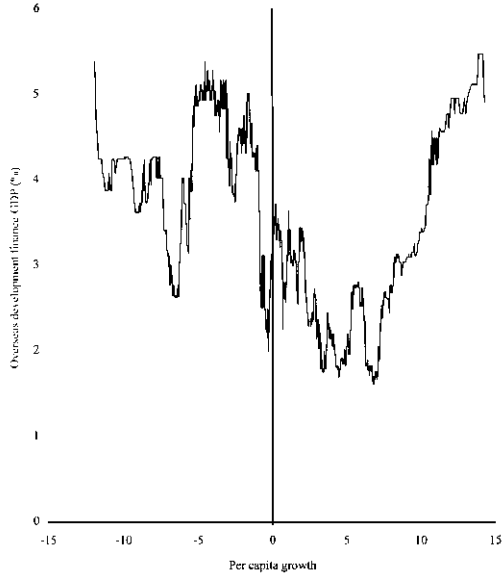

Figure 17. Overseas development finance/GDP and per capita growth in rolling medians of 100 observations.

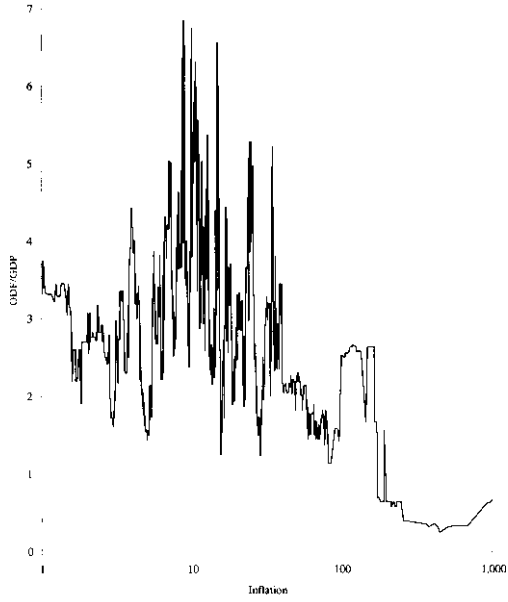

Figure 18. Inflation and ODF/GDP, rolling medians in groups of 40 observations.

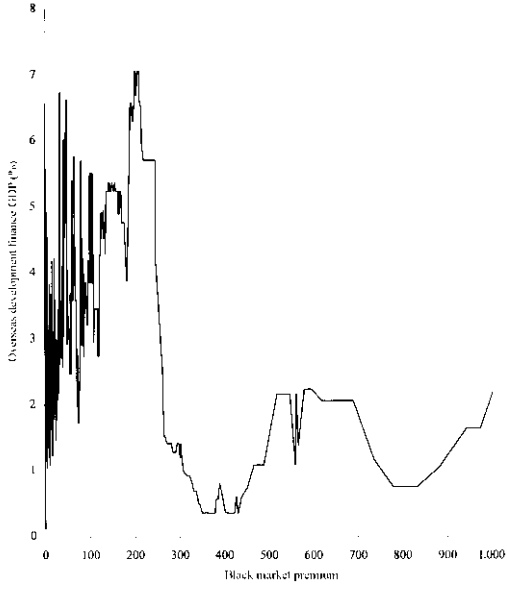

Figure 19. Overseas development finance/GDP and black market premium.

hypothesis at extreme values of the inflation rate and the black market premium. Episodes of extremely high inflation or black market premiums are followed by periods of better performance than that following episodes of moderately high inflation or black market premiums. We fail to find similar evidence of the crisis hypothesis when crisis is measured as a high current account deficit, a high budget deficit, or a negative per capita growth rate.

We speculate that the ease of correcting different type of crises may explain these different results. The pattern of foreign aid disbursements may also help explain the pattern of results. We also find that foreign aid is sharply reduced at extreme values of inflation or the black market premium, while it is actually increased for more extreme values of the current account deficit and the budget deficit. Reinforcing this point, the aid-adjusted current account deficit does show evidence for the crisis hypothesis.

Further research should examine more the mechanisms by which crises induce reform in some circumstances and not in others. We have not looked, for example, at whether having high values of multiple crisis indicators at the same time is more likely to induce a reform rather than just one. Research should also examine further the dynamics of reform or non-reform at less extreme values of crisis indicators. Our results have to do with only the most extreme values of inflation and the black market premium, following the tradition of a literature that pays most attention to these pathological cases. The policy implications of our results are complex. We think foreign aid should be reduced at extreme values of the current account deficit and the fiscal deficit, rather than increased as in current practice. Foreign aid could be reduced sooner in countries with high inflation or black market premiums, rather than reduced only at very high levels as we find now.

One implication we shy away from is that crises should be deliberately created so as to provoke reform. The damage created by very high inflation and black market premiums, and the uncertainty whether reform will follow (we find only increased likelihood, not certainty) gives us pause.

In conclusion, we find evidence for a strong version of the crisis-induces-reform hypothesis, but we find that it applies to some crisis indicators and not to others, and even then only at the most extreme values.

ALLAN DRAZEN
Tel-Aviv University and
University of Maryland

WILLIAM EASTERLY
The World Bank

REFERENCES

Akerlof, G., 1991, Procrastination and obedience. *American Economic Review* 81, 1–19.

Alesina, A. and A. Drazen, 1991, Why are stabilizations delayed? *American Economic Review* 81, 1170–1188.

Bruno, M. and W. Easterly, 1996, Inflation's children: tales of crises that beget reforms. *American Economic Review Papers and Proceedings* 86, 213–217.

Chang, C. C., E. Fernández-Arias and L. Servén, 1999, Measuring aid flows: a new approach. Policy Research Working Paper 2050 (World Bank).

Drazen, A., 1999, What is gained by selectively withholding foreign aid? Working Paper.

———, 2000, *Political Economy in Macroeconomics* (Princeton University Press: Princeton, NJ).

——— and V. Grilli, 1993, The benefit of crises for economic reforms. *American Economic Review* 83, 598–607.

Fernandez, R. and D. Rodrik, 1991, Resistance to reform: status quo bias in the presence of individual specific uncertainty. *American Economic Review* 81, 1146–1155.

Fernandez-Arias, E., 1997, Crisis, foreign aid, and macroeconomic reform. Working Paper (Inter-American Development Bank).

Harberger, A., 1993, The search for relevance in economics. *American Economic Review Papers and Proceedings* 83, 1–17.

Labán, R. and F. Sturzenegger, 1994, Distributional conflict, financial adaptation and delayed stabilization. *Economics and Politics* 6, 257–276.

Lora, E., 1998, What makes reform likely? Timing and sequencing of structural reforms in Latin America. Working Paper (Inter-American Development Bank, Office of the Chief Economist).

Nelson, J., ed., 1990, *Economic Crisis and Policy Choice: The Politics of Adjustment in the Third World* (Princeton University Press; Princeton, NJ).

———, 1994, The political economy of policy reform: panel discussion, in: J. Williamson, ed., *The Political Economy of Policy Reform* (Institute for International Economics; Washington DC) 472–477.

Olson, M., 1982, *The Rise and Decline of Nations* (Yale University Press; New Haven CT).

Rodrik, D., 1996, Understanding economic policy reform. *Journal of Economic Literature* 34, 9–41.

Svensson, J., 1999a, Foreign aid and rent-seeking. *Journal of Development Economics* forthcoming.

———, 1999b, When is foreign aid policy credible? Aid dependence and conditionality. *Journal of International Economics* forthcoming.

Tommasi, M. and A. Velasco, 1996, Where are we in the political economy of reform. *Journal of Policy Reform* 1, 187–238.

Williamson, J., 1994, In search of a manual for technopols, in: J. Williamson, ed., *The Political Economy of Economic Reform* (Institute for International Economics, Washington DC).