

WHAT EXPLAINS THE SUCCESS OR FAILURE OF STRUCTURAL ADJUSTMENT PROGRAMMES?*

David Dollar and Jakob Svensson

This paper analyses the causes of success or failure of adjustment programmes, using a new database on 220 reform programmes. We find that the success or failure of reform depends on domestic political-economy forces. A few donor-effort variables are also highly correlated with the probability of success. However, once these effort variables are treated as endogenous, there is no relationship between any of them and the success or failure of reform. These results have clear implications for how the international community should approach policy-based aid.

‘It seems clear that the lending cum conditionality process works well only when local polities have decided, largely on their own, possibly with outside technical help, to address their reform needs, effect certain policy changes sequentially, and approach the international community for financial help in getting there’ (Ranis, 1995).

Development assistance shifted to a large extent in the 1980s from financing investment (roads and dams) to promoting policy reform. This reorientation arose from a growing awareness that developing countries were held back more by poor policies than by a lack of finance for investment. The development community has had nearly 20 years of experience now with policy-based or conditional lending. There have been a large number of studies of adjustment lending, almost all of which take a case study approach. Ranis’s conclusion above – that policy-based lending works if countries have decided on their own to reform – is echoed by other studies. Our objective in this paper is to look more systematically at the causes of success or failure of adjustment programmes, using a unique database covering 220 reform programmes.

We approach this work with two hypotheses, not mutually exclusive. The first hypothesis is the one noted above: that success or failure of reform depends largely on political-economy factors within the country attempting to reform. Our analysis includes several variables that capture elements of the domestic political economy: ethnic fractionalisation, whether leaders are democratically elected, length of tenure, and others, which have been identified in the recent political-economy literature on reform (see Rodrik, 1996; Tommasi and Velasco, 1995; for recent surveys). It should be pointed

* The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries that they represent. We thank Paul Collier, William Easterly, and a number of seminar participants for comments, and Pablo Zoido-Lobaton for excellent research assistance. We also thank two anonymous referees for very constructive comments. Financial support from the World Bank’s Research Support Budget (RPO 681-70) is gratefully acknowledged.

out upfront that, even if reform depends primarily on domestic factors, policy-based aid may still be useful. In this case one should view conditional aid primarily as a commitment technology: it provides an opportunity for reformers to tie their own hands, in the same way that membership in the World Trade Organisation commits a government to good policy and insulates it from special-interest politics.

The data that we have cannot be used to discern whether policy-based aid is an effective commitment technology. To determine this would require a study of reform programmes supported by adjustment loans compared to reform programmes not supported by adjustment loans.

What we can do with our data is look at the important question of whether the donor community's effort increases the probability of success or failure of a reform programme, given that it is supported by adjustment assistance. Thus, a second hypothesis to consider – not mutually exclusive with the first – is that factors under the control of the donor community influence the success of adjustment programmes, after controlling for the domestic political-economy factors.

To assess this hypothesis, we exploit data from the World Bank's operation evaluation department (OED). The data base includes not only data on reform outcome, but also detailed information on variables under the World Bank's control, including the resources devoted to analytical work prior to reform, the resources devoted to preparation and supervision of adjustment loans, the number of conditions, and the sequencing of conditions (prior actions versus first, second, or third tranche conditions). While it would be preferable to have detailed behavioural data for all major donors, such data do not exist. However, the World Bank variables can serve as good proxies of overall donor behaviour for at least two reasons. First, the World Bank typically plays a key role in adjustment lending design and attempts to co-ordinate the donor community's actions. Second, the incentive system within the World Bank is in many respects similar to those of many bilateral donor organisations (see Mosley *et al.*, 1995; Svensson, 1997). Thus, *a priori* one would expect that many donor control variables are highly correlated across donor agencies.

In examining this second hypothesis it is important to recognise that the donor-effort variables are likely to be endogenous. We instrument for these in a two-stage probit regression. The search for good instruments reveals some interesting additional information about how the donor community (World Bank) allocates resources among activities.

We find considerable support for the first hypothesis, that domestic political-economy factors influence strongly the success or failure of reform programmes supported by adjustment loans. We find no evidence that any of the variables under the World Bank's control affect the probability of success of an adjustment loan. It is possible of course that in exceptional cases the donors' effort affects reform. What this kind of econometric work identifies is what is true on average or in general. There are a number of countries – Kenya or Zambia, for example – in which donors supported a series of mostly failed adjustment programmes. Our work suggests that these were not fertile grounds

for reform, that there are observable indicators that could have predicted this, and that the donor community working harder was not going to transform Kenya or Zambia into a successful reformer. These results have clear implications for how the international community should approach policy-based aid. They suggest that the role of donors is to identify reformers not to create them. Development agencies need to devote resources to understanding the political economy of different countries and to finding promising candidates for support. The key to successful adjustment lending is to find good candidates to support. Donors naturally may want to take some risks, and these results can help assess risks. But the results also show that adding more conditions to loans or devoting more resources to manage them does not increase the probability of reform. In fact, the World Bank devotes far more resources to the failed programmes. Once a bad loan is made, there is a tendency to put a lot of resources into salvaging it, and our evidence shows that this is fruitless.

There is a large opportunity cost to managing policy-based lending badly, and it comes in three forms. First, almost all adjustment loans disburse fully, even if policy conditions are not met (Svensson, 1999). Thus, poor choices about which reform programmes to support lead to disbursement of large amounts of aid into poor policy environments. Burnside and Dollar (1997) have shown that aid promotes growth only in a good policy environment, so that the channelling of resources into poor policy environments that accompanies failed adjustment programmes has a high cost. Second, the World Bank devotes more administrative resources to failed programmes than to successful ones, and we show that these resources have no impact. Deininger *et al.* (1998) have shown that the Bank's administrative resources have a high return in investment projects, so that using these resources on low-probability reformers has an opportunity cost. Finally, our results support the view that the best justification for policy-based lending is as a commitment technology for sincere reformers. However, the effectiveness of this technology is undermined if adjustment loans are given indiscriminately. In the data set, more than one-third of adjustment programmes failed. Such a failure rate may undermine the potential usefulness of the instrument. One reason, for example, that reformers might welcome a commitment technology is to convince private investors that policy change is permanent. However, if one-third of adjustment programmes fail (and in most cases money is still disbursed), then this instrument is not a very good signal and not much of a commitment technology. To increase the utility of adjustment lending, the donor community needs to be more selective and discerning in providing this kind of assistance.

The remainder of the paper is organised as follows: The next section describes what kind of data one would want ideally for this study and what data are actually available. Section 2 then outlines the empirical specification and estimation technique. Section 3 provides the main empirical results on explaining success of reform. Section 4 is concerned with explaining donor behaviour. The paper ends with a brief concluding section.

1. Measuring Reform Outcome and Donor Inputs

Existing studies of aid and policy reform rely almost exclusively on a case study approach. No doubt one reason for the lack of more systematic, quantitative studies is that it is hard to find the data that one needs for such a study. On the one hand, we need a measure of policy reform to use as a dependent variable. On the other hand, in order to explain reform outcome, we need variables that capture aspects of the domestic political economy as well as measures of donor inputs. It is well known that the aid business is highly fragmented with many different agencies that report differing amounts of information in different formats.

Our solution to the data problem is to use information from the World Bank covering more than 200 loans designed to support specific reform programmes, an approach that has advantages and disadvantages. The advantages are that these data provide a consistent measure of whether the reform programme succeeded or failed as well as details of a wide range of World Bank inputs supporting the reform programme. The disadvantages of this approach are that the measure of success is subjective and that the inputs of other donors are omitted from the analysis. We are not too worried about the second disadvantage. The major donors in the world are the major shareholders of the World Bank, and it seems safe to assume that the World Bank's inputs are an acceptable proxy for the donor communities' inputs. The more serious concern is whether the reform outcome measure of the World Bank is an unbiased assessment of whether or not reforms have been carried out.

The specific reform outcome measure that we use is a zero-one dummy reflecting failure or success of each reform programme as determined by the Operations Evaluation Department (OED) of the World Bank. There are three reasons why we think that this is an acceptable measure of success. First, the objective of OED evaluation is not to look narrowly at whether loan conditionalities were met or not; rather, the evaluators make a judgement as to whether or not the larger objective of reform has been met (has trade become more liberal, have enterprises actually been privatised?). This distinction is important. It is possible that a government could meet a number of unimportant conditionalities in a policy-based loan, without meaningful reform taking place. The objective of OED evaluation is what we are looking for in a dependent variable: an assessment of whether reform has taken place, not a narrow accounting of whether certain conditionalities were met.

Second, while there is clearly a subjective element to such an assessment, OED's independence within the World Bank means that there is no necessary bias in the results. OED is independent of the Bank's senior management; it has a budget allocated directly by the Board of Directors and reports to them. In our sample, 36% of the reform programs are judged by OED to have failed, that is; not to have met their objectives. In Africa, historically about one-half of reform programmes supported by adjustment lending have been judged by OED to be a failure. That assessment overall is consistent with the case study

literature that has highlighted the poor policy performance of many African countries.

The third reason to think that the OED measure is an acceptable proxy for reform is that it is highly correlated with improvements in observed economic indicators such as the rate of inflation or the extent of budget surplus. In Table 1 we present a series of regressions of changes in inflation and changes in the budget surplus over the life of a reform programme, explained by the OED measure of success or failure and one other control variable; initial GDP per capita. Specifically, we measure the changes as

$$\Delta_{t3-t0} INFL = (\text{yearly inflation rate three years } (t3) \text{ after closing year of reform programme}) - (\text{yearly inflation rate starting year } (t0) \text{ of the programme})$$

$$\Delta_{t3-t0} BB = (\text{budget balance/GDP three years } (t3) \text{ after closing year of reform programme}) - (\text{budget balance/GDP starting year } (t0) \text{ of the programme})$$

and similarly for two ($t2$) and one ($t1$) years after closing year of the reform programme. The OED sends evaluators to the country after the loan is completed to make an assessment of whether reforms were successfully carried out. And in this exercise we are relating this assessment to changes in inflation and the budget.

As is evident from Table 1, the reform programmes measured to be successful are on average associated with a reduction in the inflation rate three

Table 1
Effects of Reform on Inflation and Budget Balance

Dep. Var.	$\Delta_{t3-t0} INFL$	$\Delta_{t2-t0} INFL$	$\Delta_{t1-t0} INFL$
Outcome	-0.326 (-3.76)	-0.252 (-3.72)	-0.183 (-2.83)
Initial GDP per capita (log)	-0.110 (-2.15)	-0.061 (-1.52)	0.015 (0.40)
No. observations	159	160	160
R ²	0.12	0.10	0.05
Dep. Var.	$\Delta_{t3-t0} BB$	$\Delta_{t2-t0} BB$	$\Delta_{t1-t0} BB$
Outcome	0.035 (2.25)	0.029 (2.52)	0.013 (1.36)
Initial GDP per capita (log)	0.026 (2.47)	0.001 (0.18)	0.008 (1.45)
No. observations	129	132	144
R ²	0.09	0.05	0.03

Note: Estimation by OLS. t-statistics in parenthesis. Each regression includes a constant not reported here.

years after the reform programme ended of almost 33 percentage points. The impact of the reform measures is reduced by evaluating the reform outcome closer to the end date (year). Thus, two years after the reform programme ended, the inflation rate was on average 25 percentage points lower than when the reform was initiated (given that it was successful), and one year after the reform programme ended, the inflation rate was on average 18 percentage points lower than when the reform was initiated. The effects of successful reform on the central government's budget balance are similar. The yearly budget balance as a share of GDP is 3.5 percentage points higher three years after a successful reform. Again, the impact is lower, the shorter the lag between outcome year and the end of the reform period.

Thus, the OED assessment is significantly correlated with some of the proximate objectives of reform that can be measured, notably inflation reduction and budget control. While this increases our confidence that the OED measure is actually capturing policy reform, it also raises the question, why not simply use objective measures such as change in inflation as a measure of the extent of policy reform. There are, however, three problems with that approach. First, the reform measures differ across countries. Some reform programmes are very specifically aimed at reducing high inflation. Others aim at liberalising trade in an environment of already stabilised inflation. Reduction in inflation would not be a good measure of success of a trade liberalisation programme. We can get a large number of observations by including in the analysis all different types of reform programmes. But then it becomes impossible to find a single quantitative measure of success.

A second problem with using as the measure of policy reform an outcome variable such as change in inflation is that this change is partly driven by policy and partly driven by exogenous shocks, and it is difficult to disentangle these shocks from policy effects. Third, there is a lag between policy change and change in outcome, which is also difficult to capture (see Atkeson and Kehoe, 1997). It is precisely because it is difficult to find a good measure of policy reform, that there are so few systematic cross-country empirical analyses of the determinants of reform.¹

Our conclusion then is that the OED assessment of success or failure of reform is imperfect, but preferable to alternative measures of policy reform. It has the advantage that it can be linked up to a lot of specific information about World Bank inputs into supporting reform programmes whose achievements are then assessed on a consistent basis. In analysing our empirical work with this measure, we simply need to be mindful that there may be some bias in the measure, and we will consider how this may pollute any results.

¹ Despite the aforementioned shortcomings of using changes in inflation and fiscal balance as measures of the extent of policy reform, it is worth mentioning that using $\Delta_{t3-t0} INFL$ and $\Delta_{t3-t0} BB$ as dependent variables yields similar results as those reported in Table 3 with the OED measure as regressand, although slightly weaker [results available upon request], with respect to the set of political variables.

2. Explaining Success of Adjustment Programmes: Specification

Can the success or failure of adjustment programmes be explained by political economy variables? Do variables under control of the donor community, and in particular the World Bank, have any effect on the success rate of adjustment programmes? These are the primary questions that we address in this section. We started with 272 World Bank adjustment loans completed during the period 1980–95. For 182 of these loans we have been able to assemble data on several political-institutional factors, other exogenous variables (such as initial per capita GDP and population), and variables under the World Bank's control.

Our model can be outlined as follows. Let y_i^* be the probability of success of adjustment programme i . This probability is not directly observable. Instead we observe a zero-one indicator of success, y_i . Let \mathbf{p}_i be an $n \times 1$ vector of political-economy variables reflecting country conditions at the time of approval of adjustment loan i ; \mathbf{b}_i be a $k \times 1$ vector of variables associated with adjustment loan i under the donors' control; \mathbf{z}_i be an $k \times 1$ vector of exogenous variables correlated with the donors' behaviour but that do not influence success or failure of reform; and ε_{yi} (a scalar) and ε_{bi} (a $k \times 1$ vector) mean-zero error terms. Then the model can be expressed as

$$y_i^* = c_y + \mathbf{b}_i' \boldsymbol{\delta}_y + \mathbf{p}_i' \boldsymbol{\beta}_{yp} + \varepsilon_{yi} \quad (1)$$

$$\mathbf{b}_i = \mathbf{c}_b + \boldsymbol{\lambda}_b' \mathbf{z}_i + \boldsymbol{\beta}'_{bp} \mathbf{p}_i + \varepsilon_{bi} \quad (2)$$

where c_y is a scalar, $\boldsymbol{\delta}_y$ and \mathbf{c}_b are $k \times 1$ vectors, $\boldsymbol{\beta}_{yp}$ is a $n \times 1$ vector, $\boldsymbol{\lambda}_b$ is a $m \times k$ matrix, and $\boldsymbol{\beta}_{bp}$ is a $n \times k$ matrix of coefficients.

There are several issues in trying to estimate these equations. If the World Bank control variables were independent of the error term in (1), then we could use probit to estimate the zero-one indicator of success. However, it is likely that the error terms in (2) are correlated with the error term in (1). An exogenous shock that reduces the probability of success is likely to call forth more preparation and supervision resources. Thus, in order to estimate these relationships it will be necessary to find good instruments, that is; exogenous variables that are correlated with Bank effort but that do not influence success or failure of reform (\mathbf{z}_i). We will argue that there are such instruments and use a two-stage procedure to estimate equation (1). We will also estimate equations for two of the most important Bank-effort variables, the amount of preparation resources and the amount of supervision resources.

The theoretical literature on policy reform suggests some variables that one should include in an empirical analysis of the likelihood of successful reform, including measures of political instability (see for example Persson and Svensson, 1989; Tabellini and Alesina, 1991; Svensson, 1998), measures of polarisation and social division (see for example Alesina and Drazen, 1991; Labán and Sturzenegger, 1994; Velasco, 1993), the length of tenure of the government (see for example Tommasi and Velasco, 1995; Cukierman and Liviatan, 1992), extent of pre-reform crisis (see for example Drazen and Grilli,

1993; Velasco, 1993; Rodrik, 1994), and whether the government is democratically elected (see for example Haggard and Webb, 1994; and more generally Banks and Sundaram, 1993; Besley and Coate, 1995). We were able to collect data for a range of variables that capture these influences. Data sources and description of variables are reported in Appendix A.2.

3. Explaining Success of Adjustment Programmes: Results

Before proceeding, it is useful to take an initial look at some of the data. We have almost all of the necessary data for 182 adjustment loans (65 failures, or 36% of the sample; and 117 successes, or 64%). It can be seen in Table 2 that successful adjustment loans are associated with governments that were democratically elected (50% of successes compared to 32% of failures). Also, political instability (measured here by the average number of government crises) is highly correlated with failed adjustment. Two variables that we will use in a non-linear fashion are ethnolinguistic fractionalisation and length of time that a government has been in power.²

What is striking in Table 2 is that the World Bank-related variables are remarkably similar for successful and failed adjustment loans. Number of

Table 2
Features of Successful and Failed Adjustment Programmes

	Successful	Failed
Country characteristics		
Democratically elected	50.4%	32.3%
Political instability (average no. of government crises during reform period)	0.08	0.23
Ethnolinguistic fractionalisation	0.48	0.51
Length of time the incumbent has been in power prior to the reform	7.5	7.8
World Bank related variables		
Preparation staff weeks	141	128
Supervision staff weeks	69	101
Number of conditions	45	44
Loan size (million \$)	160	153
Sample information		
Number of loans	117	65
Features of successful and failed adjustment programmes (small sample)		
Budget surplus prior to the reform	-0.043	-0.059
Inflation prior to the reform	27%	34%
Income inequality	44.0	43.5
Terms of trade shock	1.92	1.54

² The political economy literature suggests that ethnic fractionalisation and length of tenure affect the probability of successful reform, but does not exactly identify the functional form of this relationship. The quadratic form chosen yields the best results.

conditions and loan size are nearly identical. Successful loans get about 10% more preparation resources (measured in staffweeks) than failed loans. The most striking difference is that failed loans get about 50% more supervision staffweeks. We will show that this relationship is endogenous; once the World Bank has made a bad loan, it puts a lot of resources into trying to salvage it. The interesting question will be whether those supervision resources make any difference, after controlling for this endogeneity.

Table 3 reports a series of probit regressions that attempt to explain the probability of success. For comparison, we also report the results of using a linear probability model in Table 6. Regression 1 has only the political-economy variables: success is associated with democratic government and with political stability. Ethnic fractionalisation and length of time that the incumbent has been in power enter non-linearly: the basic message is that high degrees of fractionalisation are bad for policy reform, and that long-term incumbents are not likely candidates for reform. The turning points for the length of tenure and ethnic fractionalisation vary between 15–21 years and 0.44–0.49, respectively, depending on specification. These relationships are pretty strong and the basic story is a plausible one. A recently elected government that launches reform has a 95% chance of success, *ceteris paribus*, compared to only about a 65% probability of success for an authoritarian leader in power already for 13 years (Fig. 1). That high probability of success for an elected reformer, however, can be undermined by political instability and ethnic division. An interesting finding is that the marginal impact of a democratically elected government (about 20% higher probability of success) is quantitatively independent of the degree of ethnic fractionalisation, as shown in Fig. 2. Thus, democratically elected governments have a higher probability of successfully implementing reforms, irrespective of underlying degree of ethnic fragmentation. Using only the political-economy variables, regression 1 predicts correctly 75% of the observations.³

In regression 2, we add some additional exogenous variables: initial per capita GDP, population, and regional dummies. Note that the predicted ability of the model goes down from 75% to 73% and that none of these variables has much relationship with outcomes. This is important because we are going to use these variables as instruments. It is interesting that adjustment loans do tend to be less successful in low-income countries and in Africa. But regression 2 indicates that those associations arise from the fact that low-income countries and African countries have characteristics that are not conducive to reform. With political-economy variables in the equation, there is no significance to the African regional dummy or to initial per capita GDP.

In regression 3, we add Bank-related variables to the probit regression, recognising that there is an endogeneity issue that has not yet been addressed. Some of these variables we are going to argue are exogenous: whether the

³ The prediction rule is $y = 1$ if the predicted probability > 0.5 , and 0 otherwise; that is, we predict a 1 if the model says a 1 is more likely than a 0.

Table 3
Probit Outcome Regressions

Dependent variable: Reform outcome

Regression No.	(1)	(2)	(3)	(4)	(5)	(6)
Observations	220	215	163	182	179	179
Countries	67	67	58	60	60	60
Constant	-0.098 (-0.32)	-0.753 (-0.34)	-0.002 (-0.01)	-0.762 (-0.72)	-0.366 (-0.25)	1.175 (0.93)
Ethnic fractionalisation	5.930 (4.16)	6.218 (4.00)	5.981 (2.82)	8.176 (4.40)	7.763 (4.04)	6.861 (3.74)
Ethnic fractionalisation (squared)	-6.513 (-4.27)	-7.00 (-3.89)	-6.509 (-2.90)	-8.501 (-4.32)	-8.046 (-3.79)	-7.212 (-3.57)
Political instability	-1.301 (-3.94)	-1.494 (-4.10)	-2.686 (-4.45)	-2.372 (-4.46)	-2.285 (-4.29)	-1.942 (-3.92)
Democratically elected	0.585 (2.61)	0.658 (2.71)	0.988 (3.17)	0.887 (3.11)	0.912 (3.09)	0.812 (2.80)
Time in power	-0.089 (-2.07)	-0.10 (-2.16)	-0.142 (-2.40)	-0.118 (-2.23)	-0.113 (-2.09)	-0.107 (-2.00)
Time in power (squared)	0.003 (2.15)	0.003 (2.21)	0.005 (2.28)	0.005 (2.17)	0.004 (2.02)	0.004 (1.88)
Preparation staff weeks			0.869 (1.82)	0.966 (2.31)	0.323 (0.24)	0.091 (0.08)
Supervision staff weeks			-1.342 (-2.41)	-1.410 (-2.92)	-0.869 (-0.67)	-0.934 (-0.84)
Finance conditions (%)			1.298 (1.86)	1.217 (1.84)	1.423 (2.02)	
Macro & fiscal conditions (%)			0.585 (0.59)	0.910 (1.05)	0.766 (0.89)	
Sectoral conditions (%)			1.640 (2.35)	1.386 (2.26)	1.161 (1.83)	
Trade conditions (%)			1.738 (2.19)	1.067 (1.70)	0.961 (1.46)	
Number of conditions			0.176 (0.68)			
Loan size (log)			-0.206 (-1.21)			
Expected reform period			-2.2 × 10 ⁻⁴ (-0.51)			
Prior analytical work (log)			0.065 (0.44)			
Sub-Saharan Africa		-0.175 (-0.44)				
Latin America & Caribbean		0.009 (0.02)				
East Asia		0.056 (0.12)				
Initial GDP per capita (log)		-0.213 (-0.98)				
Initial population (log)		0.144 (1.56)				
Predicted ability	0.75	0.73	0.83	0.80		

Note: Probit regressions. Regressions (5)–(6) are estimated by a two-stage procedure (Amemiya, 1978), described in Appendix A.1., with preparation and supervision specifications given in column 2, Table 4; and column 2, Table 5.

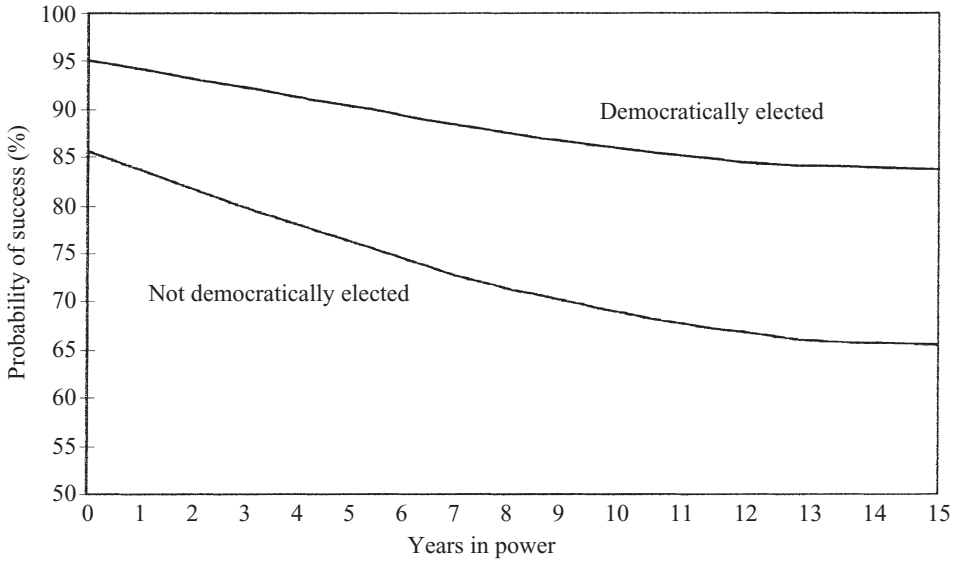


Fig. 1. Elections, Tenure, and Probability of Successful Reform.

Note: The probabilities are evaluated at the mean values of the explanatory variables.

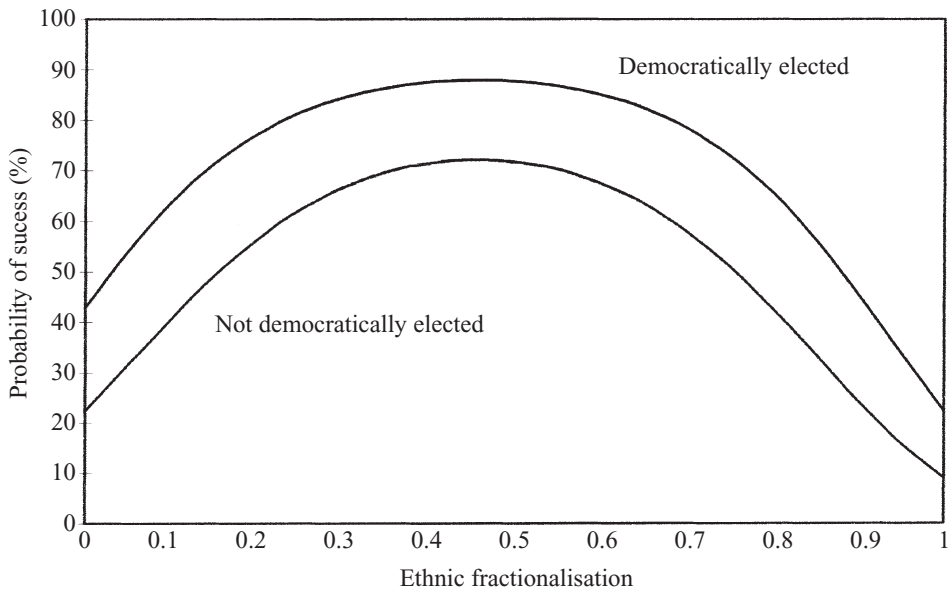


Fig. 2. Elections, Ethnic fractionalisation, and Probability of Successful Reform.

Note: The probabilities are evaluated at the mean values of the explanatory variables.

adjustment loan focuses on trade reform or sectoral reform depends on the nature of the policy problems in the country and the government's desire to attack particular problems. What is clearly under the Bank's (and more generally the donor agencies') influence is the amount of preparation staff-weeks; amount of supervision staffweeks; the staffweeks devoted to analytical work in the four years prior to the loan; the number of conditions; how conditions are allocated between upfront conditionality and first, second, and third tranches; the size of the loan, and the expected length of the reform programme.

It is difficult to instrument for all of these endogenous variables at the same time. We use the simple correlation and the partial correlation in the probit regressions to eliminate the variables that seem to have no relationship at all with outcome: number of conditions, loan size, prior analytical work, and expected length of the reform programme.⁴ Regression 4 shows the probit regression after these are removed. Of the remaining donor-related variables, the interesting story is that preparation is positively associated with outcomes and supervision, negatively associated. Once we control for these two variables, the number of conditions and the allocation of conditions play no role.

In regression 5, we instrument for preparation and supervision, using the two-stage procedure proposed by Amemiya (1978).⁵ The specifications for preparation and supervision are depicted in Table 4 (column 2) and Table 5 (column 2). Interestingly, once these Bank-effort variables are treated as endogenous, there is no relationship between any of them and the success or failure of adjustment programmes. In regression 6, we drop all of the Bank variables except preparation and supervision – for which we instrument – and again find no relationship.

The relationship between the political-economy variables and outcomes is stable throughout all of the regressions. This finding is consistent with the view that there are institutional and political factors that affect the probability of success of a reform programme. Given those factors, none of the variables under the World Bank's control affects success or failure of adjustment programmes. If endogeneity is ignored, there is a positive relationship between preparation and outcomes, and a negative relationship between supervision and outcomes. That these relationships disappear in the two-stage regressions indicates that the associations reflect how the World Bank allocates resources. In other words, preparation resources favour winners and supervision resources favour losers.

We also tried several other political variables in the outcome regression, including income inequality (as a proxy for distributional conflicts), terms of trade shocks, and the level of inflation and budget surplus prior to the reform

⁴ The measure of analytical work in the four years prior to the adjustment loan comes from Deininger *et al.* (1998). They find that that this variable has a strong association with success of investment loans; our regressions show that this is not the case for adjustment loans.

⁵ See Appendix for a brief description of the two-stage estimator.

(initial conditions).⁶ However, once we control for the vector of political variables neither of the additional regressors entered significantly at conventional levels.⁷ As the sample size is reduced by including these additional controls, we leave them out of the base specification. As shown in Table 2, however, successful reform episodes are associated with countries with better fiscal balance prior to the reform period. One explanation for why a policy variable such as initial fiscal balance does not provide additional information in the outcome regression is that it is driven by the same socio-political variables that affect the likelihood of success (see for example Easterly and Levine, 1997). It is interesting to note that this finding is consistent with earlier studies of the effects of structural adjustment programmes. Corbo and Rojas (1991) find that initial fiscal balance and terms of trade shocks are not significantly correlated with changes in GDP growth in a sample of both adjustment lending and no adjustment lending countries. Initial conditions do, however, play a role in determining whether or not a country seeks an adjustment loan, or more generally embarks on a donor-supported adjustment programme (see also Knight and Santaella, 1997). Thus, while macroeconomic conditions influence the demand for donor-supported adjustment programmes, the success of these programmes is largely determined by political factors.

In Section 1, we noted the possibility that the OED assessment of success or failure may be subject to various biases. Consider how any such bias may affect the main empirical results. We find that, after properly instrumenting, World Bank effort is not related to reform outcome. For that result to depend on a bias in the dependent variable, OED would have to systematically downgrade the ratings for reform programmes into which the World Bank has put a lot of effort – a bias that seems improbable. The same can be said for some of the findings on the political-economy variables: it is unlikely that OED has a systematic bias against ethnically diverse societies or countries with chronic political instability.

Where the bias problem may be relevant is with the findings on length of tenure and democracy. It is possible that, in rating the success of reform, there is a bias in favour of new governments and democratically elected ones. Our view is that there would not have been these kinds of systematic biases during the 1980–95 period from which our reform sample comes. If one looks at the examples of successful development policies cited in World Bank publications of that time (countries such as Chile under Pinochet, the East Asian miracle countries in the 1980s, and China), there is no tendency to single out

⁶ We also experimented with other measures of democracy. In particular, we constructed a democracy dummy from the Freedom House index, taking the value 1 if the index of political rights exceeded the average value and 0 otherwise. In a bivariate regression on the OED reform indicator, the Freedom House dummy enters significantly positive. We choose, however, to use the democratically elected dummy variable described in Appendix A.2. as our primary variable of interest, since the political agency literature briefly discussed in Section 2 stresses elections rather than the broader concept of political freedom as the driving mechanism.

⁷ Results available upon request.

democratic countries for praise.⁸ To the contrary, most of these good policy examples had authoritarian governments. A similar argument can be made about tenure. The World Bank has had deep involvement with many long-standing regimes, and one could easily imagine that any bias in the ratings of reform outcome would favour long-term clients. Thus, one should be aware of potential bias in interpreting the results, but overall we find the results plausible and consistent with other information.

4. Explaining Donor Behaviour

We turn next to an explicit examination of the World Bank's allocation of resources to help prepare and supervise reform programmes. Later in the section we consider whether World Bank behaviour is likely to be typical of the larger donor community. First, we look at the relationship between preparation resources and the political-economy variables (Table 4, regression 1). There is virtually no relationship, with only 4% of the variation in preparation resources explained by the political economy variables. Regression 2 shows a more completely specified equation for preparation resources. Note that preparation is strongly related to a number of variables that in turn have no relationship with outcomes (which is why we have adequate instruments to examine the relationship between preparation and success of reform). First, the World Bank allocates different amounts of resources to different regions, so that preparation resources tend to be low in East Asia and Latin America relative to Africa.⁹ (It is interesting that in the reform outcome equation, the political-economy variables are significant while regional dummies are not; whereas in the allocation of preparation resources, we find the opposite: regional dummies matter while most of the political-economy factors do not.) Second, there are more resources for large loans and for those with many conditions, though again these characteristics are unrelated to outcomes. Finally, resources go to low-income countries and to countries small in population.

There is a broadly similar story for the allocation of supervision resources (Table 5, regression 1). These resources favour loans that are large and have lots of conditions. Also, low-income countries and those small in population get more supervision resources. Unlike the preparation equation, regional dummies are no longer important.¹⁰ The regional departments of the World Bank have different amounts to prepare loans, but once these loans are approved the regions devote similar resources to supervising a loan of given characteristics. In the supervision equation we also have to consider that preparation may affect supervision. In studying World Bank-financed invest-

⁸ See for example World Bank (1993).

⁹ The F-statistic on the joint hypothesis that the coefficients on the regional dummies are zero is 3.21. Thus, we can reject the hypothesis at the 5% significance level.

¹⁰ The F-statistic on the joint hypothesis that the coefficients on the regional dummies are zero is 1.55. Thus, we cannot reject the hypothesis at the 5% significance level

Table 4
Preparation Regressions

Dependent variable: Preparation Staff Weeks

Regression No.	(1)	(2)
Observations	219	179
Countries	67	60
Constant	1.813 (21.58)	3.311 (4.38)
Ethnic fractionalisation	0.376 (1.00)	0.018 (0.04)
Ethnic fractionalisation (squared)	-0.327 (-0.82)	0.043 (0.10)
Political instability	-0.132 (-1.51)	-0.223 (-2.48)
Democratically elected	0.098 (1.67)	0.124 (1.98)
Time in power	0.013 (1.24)	0.004 (0.36)
Time in power (squared)	-3.4×10^{-3} (-0.95)	-3.7×10^{-3} (-0.99)
Finance conditions (%)		-0.149 (-1.07)
Macro & fiscal conditions (%)		-0.260 (-1.33)
Sectoral conditions (%)		0.002 (0.02)
Trade conditions (%)		-0.021 (-0.15)
Number of conditions		0.153 (3.29)
Loan size (log)		0.281 (5.29)
Structural adjustment loan		-0.145 (-2.16)
Sub-Saharan Africa		-0.080 (-0.78)
Latin America & Caribbean		-0.284 (-3.06)
East Asia		-0.148 (-1.39)
Initial GDP per capita (log)		-0.064 (1.04)
Initial population (log)		-0.147 (-3.90)
R ²	0.04	0.34
Adjusted R ²	0.01	0.26

Note: Estimation by OLS.

ment projects, researchers have found that more preparation resources lead to a smaller need for supervision resources. However, preparation and supervision are both associated with better outcomes in investment projects. Since there is no relationship between preparation and the success of adjustment programmes, it seems unlikely that a large amount of preparatory work would diminish the need for supervision.

In the OLS regression (Table 5, regression 1), there is a large, positive

relationship between preparation and supervision. This reflects the fact that the error terms in the preparation and supervision equations are certainly correlated. Anything unobserved that leads to higher (lower) than predicted preparation will almost certainly lead to higher (lower) than predicted supervision. The fact that the regional dummies seem to belong in the preparation

Table 5
Supervision Regressions

Dependent variable: Supervision Staff Weeks

Regression No.	(1)	(2)
Observations	179	179
Countries	60	60
Constant	2.685 (4.02)	3.272 (3.11)
Ethnic fractionalisation	-0.134 (-0.42)	-0.144 (-0.46)
Ethnic fractionalisation (squared)	0.213 (0.59)	0.254 (0.73)
Political instability	-0.029 (-0.39)	-0.017 (-0.18)
Democratically elected	-6.1×10^{-3} (-0.01)	-0.009 (-0.18)
Time in power	0.003 (0.29)	0.004 (0.48)
Time in power (squared)	-3.6×10^{-3} (-1.14)	-4.7×10^{-3} (-1.47)
Preparation staff weeks	0.339 (5.14)	0.364 (1.34)
Finance conditions (%)	-0.078 (-0.67)	-0.120 (-0.99)
Macro & fiscal conditions (%)	-0.323 (-1.97)	-0.256 (-1.41)
Sectoral conditions (%)	0.180 (1.65)	0.175 (1.59)
Trade conditions (%)	-0.141 (-1.25)	-0.141 (-1.23)
Number of conditions	0.074 (1.85)	0.077 (1.28)
Loan size (log)	0.210 (4.37)	0.220 (2.50)
Structural adjustment loan	-0.062 (-1.10)	-0.105 (-1.58)
Sub-Saharan Africa	0.093 (1.09)	
Latin America & Caribbean	0.020 (0.25)	
East Asia	-0.118 (-1.33)	
Initial GDP per capita (log)	-0.153 (-2.96)	-0.184 (-3.39)
Initial population (log)	-0.099 (-3.00)	-0.124 (-2.66)
R ²	0.50	
Adjusted R ²	0.45	

Note: Estimation by OLS [col. (1)], and 2SLS [col. (2)] with instruments given in regression 2, Table 4.

equation but not in the supervision equation means that we can use them as instruments to correct for this simultaneity problem. In the two-stage least squares regression (regression 2), the relationship between preparation and supervision is no longer significant.

It is beyond the scope of this paper to develop a full model of donor behaviour, but nevertheless, the results in this section can be related to some of the existing findings in the literature in order to draw more general conclusions about the donor community's actions. The allocation of development assistance is very much influenced by political and strategic considerations: aid is given disproportionately to former colonies and political friends of the major OECD countries (Alesina and Dollar, 1999). Once resources have been notionally allocated, there is a strong incentive for aid agencies to use these resources. A former chief economist of the Swedish aid agency writes:

'Both donor and recipient have incentive systems which reward reaching a high volume of resource transfer, measured in relation to a predefined ceiling ... In many administrations, both bilateral and multilateral, the emphasis is on disbursements and country allocations. Non-disbursed amounts will be noted by executive boards or parliamentary committees and may result in reduced allocations for the next fiscal year ... Results are measured against volume figures, with no regards for the quality ... Besides, when the time has come to evaluate the actual outcome, most of those responsible for the project on both sides will have been transferred.' (Edgren, 1996, p. 11)

The empirical results explaining the allocation of resources for preparing and supervising reform programmes can be fitted into the framework suggested by these other studies. There is at least some political element to the allocation of aid resources among regions and countries, and the incentives within agencies favour using these resources once they have been allocated. Thus, in some cases managers will prepare adjustment assistance for countries that are objectively low probability reformers. Once loans are approved, there is a tendency to use more resources to supervise loans to recalcitrant reformers. The results in the previous section suggest that neither type of resource increases the probability of successful reform. Nevertheless, supporting low-probability reform programmes makes sense for the managers in question, who have had their loans approved, who usually succeed in getting them disbursed even when reforms are not carried out (Mosley *et al.*, 1995; Svensson, 1999), and who expect to be doing something else when – years later – the support to the reform programme is judged to be a failure.

5. Conclusions

It is useful to open this section by noting what we have *not* investigated: whether or not the existence of policy-based aid affects the probability of successful reform. In reviewing major cases of reform in the post-war period, Sachs (1994) argues that in each case there was a committed reformist government, but that in addition financial aid played an important role in increasing the benefits of reform and hence sustaining political support for

the reform programme. That is a difficult hypothesis to test systematically: since virtually every modern reform programme has got financial aid, it is not possible to estimate what happens in the absence of aid. Suppose for the moment that Sachs is right. Our results are concerned with *other factors* that affect the probability of success, given that there is financial support for the reform programme; that is, our universe is a set of reform programmes supported by policy-based aid.

Our basic results can be illustrated nicely with the example of Zambia: in the 1980s, the World Bank approved four structural adjustment loans for this country, totalling \$212 million. All of these loans disbursed almost fully (less than 2% of the committed amount was cancelled). After they were completed, the independent Operations Evaluation Department within the World Bank rated three out of four as failures; that is, the reform measures supported by three out of four loans were not satisfactorily implemented. Other agencies that provided policy-based support to Zambia, such as the IMF or various bilateral donors, had similarly dismal experiences (Van de Walle and Johnston, 1996). Our results suggest that this outcome was largely predictable. Zambia at that time did not have conditions conducive to reform. The government had not been democratically elected. It had been in power for a long time in a country that is highly ethnically fragmented. Such a government is not a likely reformer.

More generally, we have shown that a small number of political economy variables can predict the outcome of an adjustment loan successfully 75% of the time. When variables under the World Bank's control – resources devoted to preparation and supervision or number of conditions – are added to the analysis, they have no relationship with success or failure of adjustment programmes.

While the results are based on data from the World Bank, we would argue that they have larger implications for how the international community should approach policy-based aid. The results do not necessarily imply that donors should stay away from high-risk environments. In these environments, the benefits of reform may be particularly high, and, as noted, the mere existence of policy-based aid may increase the probability of reform. But our results can help donors make a more informed analysis of benefits, costs, and risks. While the benefit of reform may be very high in particular cases, if the probability of success is close to zero, then the expected value of supporting the reform programme would be very low as well. Our results point to some observable political-economy variables that help assess the probability of success.

More importantly, donors should not expect to have a large impact on the probability of successful reform via such traditional approaches as adding more conditions or putting a lot of administrative resources into preparation and supervision of assistance. There may still be an important value to the administrative resources, but it would have to be through their effect on the *quality of the reform programme* rather than on the probability of announced reforms actually being carried out.

This work suggests a different mind-set for policy-based support. The role of

Table 6
Linear Probability Regressions

Dependent variable: Reform outcome

Regression No.	(1)	(2)	(3)	(4)	(5)	(6)
Observations	220	215	163	182	179	179
Constant	0.472 (4.77)	0.674 (1.01)	0.645 (1.79)	0.331 (1.22)	0.513 (1.30)	0.852 (2.54)
Ethnic fractionalisation	1.888 (4.29)	1.939 (4.09)	1.453 (2.82)	2.121 (4.47)	2.187 (4.23)	2.122 (4.20)
Ethnic fractionalisation (squared)	-2.073 (-4.42)	-2.196 (-4.00)	-1.586 (-2.91)	-2.203 (-4.40)	-2.275 (-3.99)	-2.235 (-4.05)
Political instability	-0.423 (-4.08)	-0.452 (-4.20)	-0.626 (-4.65)	-0.575 (-4.84)	-0.617 (-4.84)	-0.594 (-4.82)
Democratically elected	0.184 (2.66)	0.204 (2.68)	0.265 (3.40)	0.242 (3.23)	0.260 (3.25)	0.253 (3.15)
Time in power	-0.026 (-2.08)	-0.028 (-2.15)	-0.036 (-2.46)	-0.029 (-2.08)	-0.030 (-2.09)	-0.029 (-2.04)
Time in power (squared)	9.7×10^{-3} (2.30)	0.001 (2.28)	0.001 (2.55)	0.001 (2.19)	0.001 (2.18)	0.001 (2.08)
Preparation staff weeks			0.202 (1.72)	0.223 (2.07)	-0.009 (0.02)	-0.019 (-0.06)
Supervision staff weeks			-0.292 (-2.27)	-0.343 (-2.98)	-0.207 (-0.55)	-0.236 (-0.74)
Finance conditions (%)			0.325 (1.79)	0.295 (1.64)	0.359 (1.92)	
Macro & fiscal conditions (%)			0.116 (0.45)	0.205 (0.88)	0.191 (0.79)	
Sectoral conditions (%)			0.382 (2.08)	0.332 (1.95)	0.299 (1.66)	
Trade conditions (%)			0.402 (2.05)	0.238 (1.41)	0.247 (1.36)	
Number of conditions			0.042 (0.69)			
Loan size (log)			-0.076 (-1.76)			
Expected reform period			-9.7×10^{-5} (-0.84)			
Prior analytical work (log)			0.012 (0.31)			
Sub-Saharan Africa		-0.080 (-0.66)				
Latin America & Caribbean		-0.020 (-0.18)				
East Asia		0.025 (0.19)				
Initial GDP per capita (log)		-0.086 (-1.24)				
Initial population (log)		0.030 (1.12)				
R ²	0.17	0.20	0.34	0.32		
Adjusted R ²	0.15	0.15	0.28	0.26		

Note: Estimation by OLS (cols. (1)–(4)), and 2SLS (cols. (5)–(6)) with preparation and supervision specifications given in column 2, Table 4; and column 2, Table 5.

donors is to assess the benefits, costs, and risks of supporting particular reform efforts, taking as largely given the probability that the reform will be carried through. Staff inputs should be aimed at improving the quality of reform – that is, helping with the details of intended reform – rather than at cajoling governments to do things for which there is no political support. If donors work with the wrong model and believe that their conditions and staff inputs increase the probability of reform, then they are likely to overestimate the expected value of their aid and support too many low-probability reformers.

The general relevance of these findings can be seen by applying them to an important event that occurred after the end of our sample period, the East Asian financial crisis that developed in mid-1997. What do our results imply for international efforts to provide fast-disbursing finance to the crisis countries, conditional on policy reform? If you take our estimated model literally, it predicts that a reform programme initiated by President Suharto in the second half of 1997 was likely to fail. (The low probability comes from the fact that Indonesia had an authoritarian leader, in power for more than 30 years, in an ethnically diverse country.) On the other hand, reform programmes of newly elected governments in Korea and Thailand were likely to succeed. Furthermore, the low probability of a successful reform under the Suharto regime could not be altered by adding more conditions to loans or by applying more staff resources from international agencies. Working with the right model reduces the likelihood that resources will be wasted on reform programmes that ultimately fail.

The World Bank

Date of receipt of first submission: November 1998

Date of receipt of final typescript: December 1999

Appendix

A.1. Estimation of a Simultaneous Probit Model

The model with preparation and supervision as endogenous variables is

$$r_i = \mathbf{z}'_i \boldsymbol{\lambda}_r + \mathbf{p}'_i \boldsymbol{\beta}_{rp} + \varepsilon_{ri} \quad (\text{A.1})$$

$$s_i = \delta_{sr} r_i + \mathbf{z}'_{si} \boldsymbol{\lambda}_s + \mathbf{p}'_i \boldsymbol{\beta}_{sp} + \varepsilon_{si} \quad (\text{A.2})$$

$$y^*_i = \delta_{yr} r_i + \delta_{ys} s_i + \mathbf{p}'_i \boldsymbol{\beta}_{yp} + \varepsilon_{yi} \quad (\text{A.3})$$

where y^*_i is the probability of success of adjustment programme i . This probability is not directly observable. Instead we observe a zero-one indicator of success, y_i , s_i and r_i are supervision and preparation, respectively, of programme i . All other variables are defined in Section 2, where \mathbf{z}'_{si} is a subset of \mathbf{z}'_{ri} , and where \mathbf{p}'_i includes a constant.

In reduced vector form

$$\mathbf{r} = \mathbf{X}\boldsymbol{\Pi}_1 + \mathbf{u}_1 \tag{A.4}$$

$$\mathbf{s} = \mathbf{X}\boldsymbol{\Pi}_2 + \mathbf{u}_2 \tag{A.5}$$

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\Pi}_3 + \mathbf{u}_3, \tag{A.6}$$

where \mathbf{X} is a $t \times (n + m)$ matrix of predetermined variables. Let $\boldsymbol{\alpha}'_y = [\delta_{yr}, \delta_{ys}, \boldsymbol{\beta}'_{yp}]$, and $\boldsymbol{\delta}'_y = [\delta_{yr}, \delta_{ys}]$. We assume that $\varepsilon_{ri}, \varepsilon_{si}, \varepsilon_{yi}$ have a joint normal distribution with mean zero and covariance matrix

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Sigma}_{rs} & \boldsymbol{\Sigma}_{rsy} \\ \boldsymbol{\Sigma}_{yrs} & 1 \end{bmatrix} \tag{A.7}$$

where we have normalised $\sigma_y^2 = 1$.

A.1.1. *Estimation of $\boldsymbol{\alpha}'_y$*

The two-stage procedure proposed Heckman (1977), Nelson and Olson (1977) and others (see Lee, 1981), would be to estimate $\boldsymbol{\Pi}_1$ and $\boldsymbol{\Pi}_2$ by OLS and $\boldsymbol{\Pi}_3$ by probit, then estimate (A.3) by probit after substituting $\mathbf{X}\hat{\boldsymbol{\Pi}}_1$ for \mathbf{r} and $\mathbf{X}\hat{\boldsymbol{\Pi}}_2$ for \mathbf{s} . That is

$$y^* = \delta_{yr}\mathbf{X}\hat{\boldsymbol{\Pi}}_1 + \delta_{ys}\mathbf{X}\hat{\boldsymbol{\Pi}}_2 + \mathbf{p}'\boldsymbol{\beta}_{yp} + \eta \tag{A.8}$$

Instead of estimating (A.8), Amemiya (1978) suggests one should solve by regression methods the structural parameters from the estimated reduced form parameters. Based on this principle, one can derive asymptotically more efficient estimators. The key to this result is to note that the structural parameters are related to the reduced form parameters according to

$$\boldsymbol{\Pi}_3 = \boldsymbol{\Pi}_1\delta_{yr} + \boldsymbol{\Pi}_2\delta_{ys} + \mathbf{J}_y\boldsymbol{\beta}_{yp} \tag{A.9}$$

where $\mathbf{X}\mathbf{J}_y = \mathbf{p}$. Amemiya shows that by exploiting equation (A.6) and (A.8), (A.9) can be written as

$$\hat{\boldsymbol{\Pi}}_3 = \hat{\mathbf{G}}\boldsymbol{\alpha}_y + \mathbf{v} \tag{A.10}$$

where $\mathbf{v} = [\hat{\boldsymbol{\Pi}}_3 - \boldsymbol{\Pi}_3] - \delta_{yr}[\hat{\boldsymbol{\Pi}}_1 - \boldsymbol{\Pi}_1] - \delta_{ys}[\hat{\boldsymbol{\Pi}}_2 - \boldsymbol{\Pi}_2]$, and $\hat{\mathbf{G}} = [\hat{\boldsymbol{\Pi}}_1\hat{\boldsymbol{\Pi}}_2\mathbf{J}_y]$.

The estimates suggested by Amemiya are generalised least squares, GLS, estimates given by

$$\hat{\boldsymbol{\alpha}}_y^G = (\hat{\mathbf{G}}'\hat{\mathbf{V}}^{-1}\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}'\hat{\mathbf{V}}^{-1}\hat{\boldsymbol{\Pi}}_3 \tag{A.11}$$

where $\hat{\mathbf{V}}$ is a consistent estimator of the asymptotic variance-covariance matrix of \mathbf{v} . Thus, to be able to estimate (A.11), we need a consistent estimator of $\hat{\mathbf{V}}$. It can be shown that

$$\hat{\mathbf{V}} = d(\mathbf{X}'\mathbf{X})^{-1} + \mathbf{V}_0 \tag{A.12}$$

where $d = \boldsymbol{\delta}'_y\boldsymbol{\Sigma}_{rs}\boldsymbol{\delta}_y - 2\boldsymbol{\delta}'_y\boldsymbol{\Sigma}_{rsy}$ and \mathbf{V}_0 is the variance-covariance matrix of $\hat{\boldsymbol{\Pi}}_3$ (see Rivers and Vuong, 1988). Initial consistent estimate $\boldsymbol{\delta}_y$ can be obtained from $\hat{\boldsymbol{\Pi}}_1, \hat{\boldsymbol{\Pi}}_2$, and $\hat{\boldsymbol{\Pi}}_3$. As for $\sigma_{jy}, j = \{r, s\}$, it may be consistently estimated by

$$\sigma_{jy} = T^{-1} \sum_{i=1}^T (y_i \hat{v}_{ji} \hat{f}_i^{-1}) \tag{A.13}$$

where $\hat{v}_{ri}[\hat{v}_{si}]$ is the least squares residual from (A.4) [(A.5)], and

$$\hat{f}_i = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\hat{\boldsymbol{\Pi}}_3'x_i)^2}. \tag{A.14}$$

The asymptotic variance-covariance matrix of α_y^G is

$$\mathbf{V}(\hat{\alpha}_y^G) = (\hat{\mathbf{G}}'\hat{\mathbf{V}}^{-1}\hat{\mathbf{G}})^{-1}. \quad (\text{A.15})$$

A.2. Variables Defined in the Political Economy Literature and Empirical Proxies

Variable	Definition and source
Reform outcome	OED evaluation on adjustment operations, binary (0, 1) with 1 = successful reform. OED bases its ratings of programme outcome on assessments of whether the reform design was appropriate in terms of reducing poverty and fostering growth in the private sector, and to what extent stated policy goals have been met. Source: OED, World Bank.
Democratically elected	Dummy variable taking the value 1 if the incumbent that signed the reform was put in power by a democratic election prior to the reform, 0 otherwise. Source: Europa Yearbook (various years).
Political instability	Average number of governmental crises during the implementation of the programme. Source: Banks (1994).
Ethnic fractionalisation	Index of ethnolinguistic fractionalisation, 1960. Measures the probability that two randomly selected people in a country belong to different ethnolinguistic groups. Easterly and Levine (1997).
Income inequality	Gini-coefficient. Source: Deininger and Squire (1996).
Time in power	Number of years the incumbent that signed the reform has been in power. Source: Europa Yearbook (various years).
Terms-of-trade shock	Average growth rate of dollar export prices times initial share of exports in GDP minus the average growth rate of import prices times initial share of imports to during implementation of reform. Source: WDI 1997.
Inflation	Inflation rate prior to reform date (yearly). Source: WDI 1997.
Budget surplus	Budget surplus prior to reform date (yearly). Source: WDI 1997.
Preparation staff weeks	Preparation staff weeks on reform programme (loan). Source: OED, World Bank.
Supervision staff weeks	Supervision staff weeks on reform programme (loan). Source: OED, WB.
Finance conditions	Proportions of conditions related to financial policy out of total conditions. Source: ALCID, World Bank.
Macro and fiscal conditions	Proportions of conditions related to macro and fiscal policy out of total conditions. Source: ALCID, World Bank.
Sectoral conditions	Proportions of conditions related to sectoral policy out of total conditions. Source: ALCID, World Bank.
Trade conditions	Proportions of conditions related to trade policy out of total conditions. Source: ALCID, World Bank.

Number of conditions	Number of conditions in loan agreement. Source: ALCID, World Bank.
Loan size (log)	Logarithm of World Bank loan amount in US\$. Source: OED, World Bank.
Expected reform period	Expected reform period (days). Source: OED, World Bank.
Prior analytical work	Prior analytical work four years prior to reform on reforming country. Source: Deininger <i>et al.</i> (1997).
Sub-Saharan Africa	Binary variable taking the value 1 for countries in Sub-Saharan Africa, 0 otherwise.
Latin America & Caribbean	Binary variable taking the value 1 for countries in Latin America & Caribbean, 0 otherwise.
East Asia	Binary variable taking the value 1 for countries in East Asia, 0 otherwise.
Initial population (log)	Initial population (log). Source: WDI 1997.
Initial GDP per capita (log)	Initial GDP per capita (log). Source: WDI 1997.

References

- Alesina, A. and Dollar, D. (1999). 'Who gives aid to whom and why?' Harvard University, mimeo.
- Alesina, A. and Drazen, A. (1991). 'Why are stabilisations delayed?' *American Economic Review*, vol. 81, pp. 1170–88.
- Amemiya, T. (1978). 'The estimation of a simultaneous equation generalised probit model.' *Econometrica*, vol. 46, pp. 1193–205.
- Amemiya, T. (1979). 'The estimation of a simultaneous-equation Tobit-model.' *International Economic Review*, vol. 20, pp. 169–81.
- Atkeson, A. and Kehoe, P. (1997). 'Industry evolution and transition: a neoclassical benchmark', NBER Working Paper Series, no. 6005.
- Banks, J. and Sundaram, R. (1993). 'Long-lived principals, short-lived agents.' University of Rochester, mimeo.
- Bates, R. H. and Krueger, A. O., eds. (1993). *Political and Economic Interactions in Economic Policy Reform*. Cambridge, MA: Blackwell.
- Besley, T. and Case, A. (1995). 'Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits.' *Quarterly Journal of Economics*, vol. 110, pp. 769–98.
- Branson, W. and Jayarajah, C. (1995). 'Structural and sectoral adjustment: World Bank experience, 1980–92.' Washington, DC, OED, The World Bank.
- Burnside, C. and Dollar, D. (1997). 'Aid, policies and growth.' The World Bank Policy Research Working paper no. 1777.
- Corbo, V. and Rojas, P. (1991). 'World Bank-supported adjustment programmes: country performance and effectiveness.' The World Bank Policy, Research, and External Affairs Working paper no. 623.
- Cukierman, A. and Liviatan, N. (1992). 'Dynamics of optimal gradual stabilisation.' *World Bank Economic Review*, vol. 6, pp. 439–58.
- Deininger, K. and Squire, L. (1996). 'A new data set measuring income inequality', *World Bank Economic Review*, vol. 10, pp. 565–91.
- Deininger, K., Squire, L. and Basu, S. (1998). 'Does economic analysis improve the quality of foreign assistance?' *World Bank Economic Review*, vol. 12, pp. 385–418.
- Drazen, A. and Grilli, V. (1993). 'The benefit of crises for economic reforms.' *American Economic Review*, vol. 83, pp. 598–607.
- Easterly, W. and Levine, R. (1997). 'Africa's growth tragedy: policies and ethnic division.' *Quarterly Journal of Economics*, vol. 112, pp. 1203–50.
- Edgren, G. (1996). 'A challenge to the aid relationship.' In *Aid Dependency*, Stockholm: Swedish International Development Authority.
- Haggard, S. and Webb, S. B., eds. (1994). *Voting for Reform: Democracy, Political Liberalisation, and Economic Adjustment*. New York: Oxford University Press.
- Heckman, J. J. (1977). 'Dummy endogenous variable in a simultaneous equation system.' *Econometrica*, vol. 46, pp. 931–59.

- Knight, M. and Santaella, J. A. (1997). 'Economic determinants of IMF financial arrangements.' *Journal of Development Economics*, vol. 54, pp. 405–36.
- Labán, R. and Sturzenegger, F. (1994). 'Distributional conflict, financial adaptation and delayed stabilization', *Economics and Politics*, vol. 6, pp. 257–76.
- Lee, L. F. (1981). 'Simultaneous equation models with discrete and censored dependent variables.' In (C. Manki and D. McFadden, eds.) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.
- Mosley, P. (1987). 'Conditionality as a bargaining process: structural adjustment lending, 1980–86.' Princeton Essays in International Finance no. 168.
- Mosley, P., Harrigan, J. and Toye, J. (1995). *Aid and Power: the World Bank and Policy-Based Lending*. New York: Routledge.
- Nelson, F. D. and Olson, L. (1978). 'Specification and estimation of a simultaneous-equation model with limited dependent variables.' *International Economic Review*, vol. 19, pp. 695–709.
- Persson, T. and Svensson, L. E. O. (1989). 'Why a stubborn conservative would run a deficit: policy with time-inconsistent preferences,' *Quarterly Journal of Economics*, vol. 104, pp. 325–45.
- Ranis, G. and Mahmood, S. A. (1992). *The Political Economy of Development Policy Change*. Cambridge, MA: Blackwell.
- Ranis, G. (1995). 'On fast disbursing policy-based loans.' Economics Department, Yale University, mimeo.
- Rivers, D. and Vuong, Q. H. (1988). 'Limited information estimators and exogeneity tests for simultaneous probit models.' *Journal of Econometrics*, vol. 39, pp. 347–66.
- Rodrik, D. (1994). 'The rush to free trade in the developing world: why so late? why now? will it last?' In (S. Haggard and S. B. Webb, eds.) *Voting for Reform: Democracy, Political Liberalization, and Economic Adjustment*. New York: Oxford University Press.
- Rodrik, D. (1996). 'Understanding economic policy reform.' *Journal of Economic Literature*, vol. 34, pp. 9–41.
- Sachs, J. (1994). 'Life in the economic emergency room.' In (J. Williamson ed.) *The Political Economy of Policy Reform*. Washington, DC: Institute for International Economics.
- Svensson, J. (1997). 'When is foreign aid policy credible? Aid dependence and conditionality?' *Journal of Development Economics* (forthcoming).
- Svensson, J. (1998). 'Investment, property rights and political instability: theory and evidence.' *European Economic Review*, vol. 42, pp. 1317–41.
- Svensson, J. (1999). 'Why conditional aid doesn't work and what can be done about it?' The World Bank, mimeo.
- Tabellini, G. and Alesina, A. (1990). 'Voting on the budget deficit.' *American Economic Review*, vol. 80, pp. 37–49.
- Van de Walle, N. and Johnston, T. (1996). *Improving Aid to Africa*. Washington, DC: Overseas Development Council.
- Tommasi, M. and Velasco, A. (1995). 'Where are we in the political economy of reform.' C. V. Starr Center for applied economics, NYU RR 95-20.
- Velasco, A. (1993). 'A model of endogenous fiscal deficits and delayed fiscal reforms.' C. V. Starr Center for applied economics, NYU RR 93-4.
- World Bank (1993). *The East Asian Miracle: Economic Growth and Public Policy*. New York: Oxford University Press.
- WDI (1997). *World Development Indicators*, World Bank, Washington D.C.