POLICY RESEARCH WORKING PAPER 2681

On the Duration of Civil War

Paul Collier
Anke Hoeffler
Måns Söderbom

The duration of large-scale, violent civil conflict increases substantially if the society is composed of a few large ethnic groups, if there is extensive forest cover, and if the conflict has commenced since 1980. None of these factors affects the initiation of conflict. And neither the duration nor the initiation of conflict is affected by initial inequality or political repression.

The World Bank
Development Research Group
Office of the Director
September 2001

Summary findings

Collier, Hoeffler, and Söderbom model the duration of large-scale, violent civil conflicts, applying hazard functions to a comprehensive data set on such conflicts for the period 1960–99. They find that the duration of conflicts is determined by a substantially different set of variables than those that determine their initiation. The duration of conflict increases substantially if the society is composed of a few large ethnic groups, if there is extensive forest cover, and if the conflict has commenced since 1980. None of these factors affects the initiation of conflict.

The authors also find that neither the duration nor the initiation of conflict is affected by initial inequality or political repression. This finding is consistent with the hypothesis that rebellions are initiated where they are viable during conflict, regardless of the prospects of attaining post-conflict goals, and that they persist unless circumstances change.

This paper—a product of the Office of the Director, Development Research Group—is part of a larger effort in the group to study large scale violent conflict. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Paul Collier, mail stop MC3-304, telephone 202-458-8208, fax 202-522-1150, email address pcollier@worldbank.org. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The other authors may be contacted at anke.hoeffler@ox.ec.ac.uk, or mans.soderbom@ox.ec.ac.uk. September 2001. (29 pages)

On the Duration of Civil War

Paul Collier,[(1), (2)] Anke Hoeffler,[(2)] and Måns Söderbom[(2)]

[(1)]World Bank
[(2)]Centre for the Study of African Economies, University of Oxford

Prepared for the World Bank, Development Research Group/University of California, Irvine, Center for Global Peace and Conflict Studies, Workshop on Civil Wars and Post-Conflict Transitions, May 18-20, 2001.

1. Introduction

In understanding civil wars it is useful to distinguish between their initiation and their scale. In our previous research we have focused on explaining their initiation, we now address scale. The scale of a conflict has various dimensions, and here we are concerned only with one, namely, duration. Other pertinent dimensions are the geographic reach of conflict, and its intensity in terms of mortality and refugees. We consider duration because from a policy perspective it is arguably more important to know how to bring a war to an end than to know how to contain it.

In Section 2 we relate the duration of conflict to two alternative models of its initiation. Until recently, economic models of rebellion assumed that the benefits of rebellion accrued only upon victory; new models treat the benefits as accruing during the conflict. This difference has implications for whether conflict duration will be determined by substantially similar, or substantially different, factors from those that determine the initiation of conflict. In Section 3 we present the results of hazard function regressions of the duration of conflict based on data for the period 1960-99. We start from the same explanatory variables as that of the Collier-Hoeffler model of the initiation of conflict, and use statistical rather than theory-based criteria to refine the model. We compare the resulting model of duration with the model of conflict initiation and draw out implications for the process of rebellion. Section 4 concludes.

2. The Initiation and Duration of Conflict

Until recently it was conventional to model the benefits of rebellion as being contingent upon rebel victory. This was, for example, the assumption underpinning the celebrated model of Grossman (1991). Although Grossman's model was not temporal, in a simple extension Grossman (1995) introduced a discount rate into the analysis. Collier and Hoeffler (1998) assumed both that benefits accrued only upon victory, and that during the period of fighting the rebels incurred net costs. Hence, the longer the expected duration of war the more heavily discounted would be these benefits, and the higher would be the costs of fighting. The expected duration of a conflict would thus influence the decision whether to rebel, and conversely, the expected eventual pay-off from victory would influence whether a long conflict was worthwhile. Duration and initiation were therefore inter-dependent. For example, suppose that the extent of forest cover determines the probability of reaching a militarily decisive outcome but not the eventual pay-off from victory: forests lengthen wars but do not affect post-war outcomes. Conversely, suppose that natural resource rents determine the eventual pay-off but not the duration. Then we would expect to find that forestation would be associated with a longer duration of war and therefore with a reduced probability of war. Conversely, natural resource rents would be associated both with an increased probability of war and with wars of longer duration: high eventual pay-offs would justify long wars.

More recently, rebellion has been modeled as a `quasi-criminal' activity which is profitable *during* the conflict (Collier, 2000). In this model because the duration of a conflict is not a cost, the factors which determine duration, forestation in the above example, would not influence the decision to initiate conflict and so would not be correlated with conflict initiation. Conversely, factors which determine the initiation of conflict, natural resource rents in the above example, would not be correlated with the duration of conflict. Of course, there might potentially be a factor, for example government capability, which influenced both initiation and duration. High government capability might reduce the profitability of rebellion during conflict, so reducing the probability of war, and increase the prospects of a decisive military outcome, thus reducing duration. However, contrary to the implication of the Grossman-type models, duration and initiation would not be inter-dependent, arising from the logic of the decision calculus. They would simply have some explanatory variables in common.

In recent empirical work we have modeled the initiation of civil conflict, (Collier and Hoeffler, 2001). The dependent variable is the probability of conflict during a five year period, and the sample is in principle global for the period 1960-99, being reduced only by considerations of data availability. If the rebel decision calculus balances the eventual pay-off against costs linked to the expected duration of the conflict, a significant explanatory variable might be effecting either the eventual pay-off or the expected duration. If, at the other extreme, the rebel decision calculus focuses only on viability during conflict, then significant explanatory variables must be affecting that viability. A comparison of the empirical correlates of conflict duration with those found in the Collier-Hoeffler model of conflict initiation can thus to an extent constitute a test of whether the expected duration of conflict is important in the decision to initiate a rebellion. In turn, this illuminates the extent to which rebellion is determined by the anticipation of post-conflict benefits, relative to the benefits which accrue during conflict. If the expected duration of conflict is important in the rebel decision calculus the factors that are significant for the predicted probability of conflict initiation should include those that are significant for the predicted duration of conflict. Note, however, that the signs of the coefficients need not be the same. In the above example, natural resource rents would be associated with longer duration and a higher risk of conflict initiation, while forestation would be associated with longer duration but a *lower* risk of conflict initiation. Conversely, if the expected duration of conflict does not enter into the rebel decision calculus, then we would expect to find little correspondence between the sets of significant variables in the two regressions.

The empirical correlates of the initiation of civil war, as found in the Collier-Hoeffler model, can be grouped into economic, social, geographic and historical. Of these, the most important are the economic. The level, growth and structure of income are all significant influences on conflict risk: countries with low per capita income, slow growth, and a high share of primary commodity exports in GDP face considerably higher risks. Social composition also affects risk. Countries in which the largest ethnic group constitutes between 45% and 90% of the population have around double the risk of other countries. However, other than this `ethnic dominance' effect, ethnic and religious diversity actually reduces conflict risk: diverse societies are safer than homogenous societies. Geography matters: countries with dispersed populations, and those which are mountainous, face somewhat higher risks. Finally, history matters: once a country has had a conflict, it faces a temporarily higher risk of further conflict. Several other explanatory variables have been tested for inclusion in the model and rejected: for example, neither inequality nor political repression have a significant effect upon the risk of conflict initiation. Hence, in this paper we first investigate whether the variables of the Collier-Hoeffler model of conflict initiation provide a statistically reasonable explanation of conflict duration.

This comparison is not the only potential contribution of the Collier-Hoeffler model to the study of conflict duration. Our study of the initiation of conflict found that in the early post-conflict phase the risk of renewed conflict is markedly higher *as a result of the conflict*. The typical post-conflict country faces a 50% risk of renewed conflict within the first five years of reaching peace. Of this risk, around half is due to the factors inherited from before the conflict, the other half being attributable to some new risks which were generated as a result of the conflict. These conflict-acquired risks gradually decay if peace is maintained. In effect, our two snapshots of risk on the eve of an initial conflict and just after peace has been re-established tell us that during the conflict something happens which drastically but temporarily increases risk. In this paper we focus on how the conflict evolves year-by-year, and specifically how the risk of continued conflict changes. We might expect to see the same rising risk of continued conflict *during* the conflict as we find has occurred in comparing the before and after snapshots of risk.

Both peace and civil war are highly persistent states. Societies at peace have a high probability of remaining at peace. Societies in civil conflict have a high probability of remaining in civil conflict. In this sense the process of war duration cannot be the same as the process of war initiation: the initiation of war usually represents an unlucky drawing from a distribution, whereas the continuation of war is a high probability event. Thus, the continuation of war cannot be a drawing from the same distribution as that which generated it, and so war must radically change the distribution. The gradual decay of the conflict risk in post-conflict societies tells us that the persistence of peace is itself dynamic: the longer it lasts the more likely is it to persist. Conflict persistence potentially has the same property, and we investigate whether this is the case.

3. Empirical Analysis

Data

The data on which the analysis is based are from Collier and Hoeffler (2001, 2001a). However, while that analysis compared countries in which civil conflicts were initiated with those which remained at peace, the present analysis necessarily only compares among countries in which conflicts were initiated. Here we provide a brief summary of the data. Please refer to the appendix for a full list of the 52 war observations used in our baseline model and a full description of the variables.

On average the civil wars in our sample are just over seven years long. The shortest wars only lasted one month while the longest lasted over 30 years. In Table 1 we present some descriptive statistics of the potential determinants of war duration. In the first column we list the descriptive statistics for the entire sample and compare these with wars of a duration of four years or less (column 2) and with wars longer than four years (column 4). We chose this comparison because in our regression model we include dummy variables for biannual intervals. All variables were measured in the year the conflict began or, if these data were not available, in the year closest to the start of the conflict.

We begin the description of the data with the economic variables. GDP per capita (in 1985 US dollars) is around $1762 for the entire sample. Countries with shorter wars initially had a higher per capita income ($2277) while countries with longer wars were characterized by lower incomes ($1413). This difference between the different sub-samples is also apparent when we compare the male enrolment rates in secondary schools. At the start of the conflict in countries with shorter wars a much higher proportion of young men were enrolled (43 percent) than in countries with longer wars (25 percent). However, the structure of the different war economies were similar. The average share of primary commodity exports in GDP was about 13.5 percent for all war observations. Countries with shorter wars are characterized by a slightly higher share of primary commodity exports (14.5 percent).

We now turn to some socio-political variables. Using data from the Polity III data set we measure the level of democracy as the openness of the political institutions on a scale of 0 (least open) to 10 (most open). Both short and long wars are characterized by the same low levels of democracy (average values of about 2.3). However, in terms of ethnic and religious diversity we find that the averages of the sub-samples are different. We measure diversity on a scale of zero to 100. The diversity measures are the probabilities that two randomly drawn individuals do not belong to the same group. Thus, a value of zero characterizes perfect homogeneity and 100 complete heterogeneity. The ethnic diversity measure is based on the data from Atlas Naradov Mira and the religious diversity measure is based on the World Christian Encyclopedia. Countries with shorter wars are characterized by much lower levels of ethnic and religious diversity (44 and 31, respectively) than countries with longer wars (57 and 37, respectively).

In terms of geographic variables we also find that the countries have different characteristics. Observations with shorter wars are more mountainous (28.2 percent of the country), have a lower coverage of woods and forests (13.7 percent of the country) and have a more dispersed population (0.52). Longer war observations are less mountainous (22.7 percent of the country) but have a higher coverage of woods and forests (30.8 percent of the country) and have a more concentrated population (0.61). Lastly, countries with shorter wars are located in less conflict-ridden regions. On average only 0.3 neighbors of the countries with shorter wars were characterized by internal warfare. For countries with longer wars 0.4 neighbors were experiencing a conflict at the beginning of the country's civil war.

To summarize, the descriptive statistics show considerable variation between the countries with shorter wars and the ones that experienced longer wars. We now apply hazard functions to model the duration of civil conflict.

Constructing the Likelihood Function

The econometric specification is a continuous time hazard model of the monthly transition rates from war to peace. To illustrate the specification, consider a country i

which in period t experiences a war. The transition rate from war to peace is represented by the hazard function, which we specify as

$$h_{it} = \exp(X_{it}\beta + \mu_i)h^B(D_{it}), \tag{1}$$

where X_{it} is a vector of explanatory variables, β is a coefficient vector, μ_i is a random effect, permanent for each country, h^B is the baseline hazard, and D_{it} is the duration of the spell. The associated survivor function, which measures the unconditional probability that a duration is longer than D, is equal to

$$S_{it} = \exp\left(- \int_0^{D_i} \exp(X_{it}\beta + \mu_i)h^B(u)du \right). \tag{2}$$

and the likelihood that a period of length D is completed (δ=1) or censored (δ=0) at time t can then be written $L_{it} = S_{it}(h_{it})^{\delta_i}$. Because some countries have had several spells of war over the sample period, we write country i's contribution to the likelihood from R spells equal to the product $\prod_{l=1}^{R_i} L_{it}$.

For the baseline hazard we adopt a specification which allows for stepwise estimation, which is flexible. By dividing the time axis into W intervals by the points $c_1, c_2, ..., c_W$ and assuming constant hazard rates within each interval, we can write the hazard function as

$$h_{it} = \exp(X_{it}\beta + \mu_i)\exp(\lambda d_{wit}), \qquad c_{w-1} < D_{it} \le c_w, \; w = 1,2,...,W \tag{3}$$

with $c_0 = 0$ and $c_W = \infty$, and where d_{wit} is a dummy variable equal to one if $c_{w-1} < D_{it} \le c_w$, and zero otherwise. This model, termed *piecewise exponential*, will thus estimate W baseline hazard parameters, $\lambda_1, \lambda_2, ..., \lambda_W$. The fact that the baseline hazard is allowed to vary freely between intervals imposes few restrictions on duration dependence. In estimation we include an intercept in each hazard model, and therefore exclude the first "duration dummy". Hence, negative (positive) coefficients on remaining duration dummies means that the hazard is lower (higher) than in the first interval.

To control for unobserved country-specific heterogeneity we include a random effect μ_i. It is not difficult to think of unobservable factors which affect the hazard in this context, and it is well known that omission of such factors will lead to spurious negative duration dependence and biased parameter estimates for the explanatory variables (see e.g. Lancaster, 1990, chapter 4). To obtain the unconditional likelihood it is necessary to integrate over the distribution of μ. We follow Heckman and Singer (1984) and adopt a non-parametric strategy in which the (unknown) distribution of μ, $g(\mu)$, is approximated by a discrete multinomial distribution with two points of support $\mu = \mu^*$, where

$$\mu^* = \begin{cases} \mu_1 & \text{with probability } P \\ \mu_2 & \text{with probability } 1-P \end{cases}.$$

Since we include an intercept we need to normalise one of the support points to zero. Hence one support point and one probability parameter will be estimated in addition to the other parameters of the model. The individual contribution to the likelihood function can thus be written

$$(4) \qquad L_{UH} = P \cdot \prod_{l=1}^{R} L_{lit}\left(\cdot \middle| \mu^* = -0.5\mu_2\right) + (1-P) \cdot \prod_{l=1}^{R} L_{lit}\left(\cdot \middle| \mu^* = 0.5\mu_2\right).$$

This specification is the basis for the regression analysis to which we now turn.

Results

The hazard function predicts the probability[1] of peace in each phase of the conflict. Figure 1 depicts this probability, year-by-year during the conflicts in our sample, before we introduce any other explanatory variables. The translation from the hazard function to a single summary number for the duration of conflict is non-trivial since the hazard function describes a vector of probabilities. An increase in the `hazard' of peace will evidently shorten the expected duration of conflict, but the relationship is not linear. Here we convert the hazard function to the mathematical expectation of the duration of conflict. While the regression results directly report the hazard function, we report in the text the effects of important variables upon the expected duration of conflict, these having been calculated case-by-case.

We begin by using the variables of the Collier-Hoeffler model of the initiation of conflict to explain the duration of conflict. The only deviation from the variables used in that

[1] Strictly speaking it is inappropriate to use the term "probability", as the hazard in theory is unbounded above. "Risk" or "rate" are more appropriate terms. In practice, with discrete data, the distinction is less important.

model is to add dummy variables for various phases of the conflict. Specifically, we use as a base the first two years of the conflict and add dummy variables for the three subsequent two-year periods and a dummy variable for those conflicts which lasted longer than eight years. The Collier-Hoeffler model has two variants: because per capita income and secondary educational enrolments are highly correlated, they cannot be included in the same regression, but each is significant when included separately. We therefore test both variants.

The results are reported in Table 2. The model performs remarkably badly in either variant. The only variables which are significant are the dummy variables for the phases of the conflict and these are obviously not part of the model of conflict initiation. This is at least *prima facie* evidence that the expected duration of conflict is not an important factor in rebel decisions to initiate conflict. However, such an interpretation is insecure unless the best-performing model of conflict duration turns out to be substantially different.

Table 3 presents our preferred 'baseline' model of conflict duration, reached after a series of iterations in which insignificant variables are deleted and variants of the economic, social, geographic and historical explanatory variables are then tested in turn. We show results without and with controls for unobserved heterogeneity, in columns [1] and [2] respectively.

We first consider the economic variables. The level of per capita income and the male secondary school enrolment rate are each significant when included separately (recall that they cannot be included together because they are too highly correlated). The results controlling for unobserved heterogeneity imply that if the proportion of males enrolled in secondary school is increased by 10 percentage points the 'hazard' of peace is increased by 42% and the expected duration of conflict is reduced by 18%.[2]

[2] Evaluated at mean values of the regressors.

Now consider the social variables. Ethnic fractionalization and its square are both highly significant. Thus, ethnic fractionalization has a non-monotonic effect: the duration of conflict is at its maximum when ethnic fractionalization is around 50 on its 0-100 range. This typically occurs when the society is made of a few large ethnic groups. Not only is ethnic fractionalization highly statistically significant, the effect is very substantial. Both a homogenous society and a completely fractionalized one have a hazard of peace which is around eight times higher than one with a fractionalization score of 50, reducing the duration of conflict by 77%. There is a statistically insignificant but similarly signed effect of religious fractionalization. Controlling for composition, the size of population is significant: more populous countries have longer wars. The effect is fairly large: doubling the population reduces the hazard of peace by 23% and so increases the duration of conflict by 14%.

Now consider the geographic variables. The greater the proportion of a country which is covered in forest the lower is the hazard of peace and hence the longer is the conflict. This is both statistically and economically significant. Comparing a country with no forest cover with the average cover of those countries with wars lasting more than four years, namely 31%, the hazard of peace is 58% lower and the duration of conflict consequently around 66% longer. The other geographic variable of significance is if a neighboring country has a civil war. This reduces the hazard of peace by 47% and so increases the duration of conflict by 41%.

Now consider the historical variables. The hazard function includes dummies both for the decade and for the phase of the conflict. The dummy for the 1990s is potentially problematic in a hazard function due to the truncation problem: for example, it is not possible for a war which commenced in 1996 to have a duration of more than four years given our period of observation. However, since our highest duration phase dummy is for wars that lasted more than eight years, any war which started during or prior to the 1980s has had the time to last more than eight years by the end of our period of observation. To correct for the truncation problem during the 1990s we add three interaction terms: the 1990 decade dummy interacted with each phase of the conflict other than the first. Thus, the coefficient on the 1990s decade dummy itself represents only the effect of the 1990s on the probability of peace during the first two years of war, for which the truncation problem is minimal. Both the decade dummies and the phase of conflict dummies are collectively significant. The decade dummies reveal that post-1980 the hazard of peace is far lower than prior to 1980. A conflict starting prior to 1980 had a peace hazard around eight times greater than one occurring post-1980 and consequently had an expected duration only 30% as long.[3] We investigated why wars that started post 1980 appear to be longer by interacting the post-1980 decade dummies with a variety of explanatory variables, but none were significant.

The conflict phase dummies show a U-shaped pattern for the hazard of peace. There is a relatively good chance that a conflict will end in the first two years. During the third and fourth year the prospects for the conflict ending are at their minimum, and thereafter they gradually revive. We investigated whether the characteristics which influence the duration of conflict are differentially important in different phases of conflict. For this taking each significant variable in turn, we added interaction terms between the variable and the phase of conflict dummies. Only one of these produced significant results: secondary school enrolment at the start of the conflict (Table 4). We found that while the higher is initial secondary school enrolment the shorter is the conflict, this effect gradually erodes as the conflict persists. A possible explanation for this, consistent with much case study evidence, is that during a long conflict the educated part of the population is the most likely to emigrate. Hence, while in the early phase of conflict education is a force for peace, its influence gradually diminishes because the educated leave.

We now compare the hazard function with our earlier model of the initiation of conflict. Both have similar overall levels of explanatory power of around 30%, which might be judged reasonably given the parsimony of the models, the poor quality of the data in

[3] The confidence interval is large however. Looking a the raw data, the average duration for wars that started prior to 1980, none of which were right-censored, was 78 months. In contrast, the average duration for wars that started in the 1980s was 105 months, hence more than 30% longer. This number is biased

conflict countries, and the highly idiosyncratic nature of the events to be explained. However, the two models are strikingly different. Only three variables are common to both models: neighboring wars, secondary education and population. It is unsurprising that a war in a neighboring country increases both the risk of conflict initiation and the duration of a conflict should it occur. Turning to case study evidence to illustrate what might be going on, there are numerous cases in which the same rebel organization operates in more than one country, gaining advantages from the fact that while it is united, the opposing governments may find it hard to cooperate effectively. Current examples are the RUF using its base in Sierra Leone to move into Guinea, the KLA using its base in Kosovo to move into Macedonia, and the Basque separatist movement using its safe havens in France to mount terrorist actions in Spain. However, case studies also suggest a different pattern in which hostile neighboring governments actively support rebel organizations in their neighbors as a form of indirect international war. An example would be Sudanese support for the LRA in Uganda and Ugandan support for the SPLA in Sudan. In the present analysis we cannot distinguish between these effects although both would give rise to the results which we find. Obviously, the policy implications are somewhat different: the former requires active cross-border cooperation by governments, whereas the latter requires an international peace negotiation between governments.

Recall that male secondary school enrolment and per capita income are separately significant in reducing both the risk of the initiation of war and its duration. Perhaps one variable proxies the other or both proxy some further correlate of the level of development, such as government effectiveness. Such an unspecific indicator provides little guidance for policy other than that developed societies will have both fewer and shorter wars.

We would expect the risk of the initiation of war to be correlated with population simply because larger countries in effect aggregate the risks of more localities. This aggregation effect alone would produce an elasticity of risk with respect to population of close to

downwards, as 20% of these wars were still ongoing at the end of 1999; that is, the true average duration is bound to be higher.

unity for small levels of risk. In our analysis of the initiation of conflict the coefficient on the logarithm of population is not significantly different from unity, so that this is the most straightforward interpretation of the result. The result on the duration of conflict may be due to a similar effect. In some countries more than one conflict is going on at the same time. For example, in Ethiopia during the 1980s there were up to six distinct simultaneous civil wars. Larger countries are more likely to have multiple civil wars. In our data, even if one conflict ends, if another conflict is continuing the country is still classified as at civil war. Hence, for both the initiation and the duration of conflict the population effect may be spurious.

There are eight variables which are significant in initiating conflict yet which do not effect its duration: the growth of income prior to conflict, the share of primary commodity exports in GDP, ethnic dominance, social fractionalization, the size of diasporas,[4] geographic dispersion, mountainous terrain, and the time which has passed since any previous conflict. Similarly, there are three variables which are significant in the duration of conflict but not in its initiation: ethnic diversity (as a quadratic relationship), forest cover, and decade dummies. These differences suggest that the likely duration of a conflict is not a substantial factor influencing the rebels who decide to initiate it. In turn, this implies that the initiation of rebellion is less influenced by the prospects of attaining post-conflict goals than by its short-term viability.

Three variables that dominate the political discourse on conflict are insignificant both in its initiation and its duration: prior inequality, whether of income or of land, prior political repression, and prior military expenditure. All of these variables are difficult to measure, and there are powerful other reasons for favoring both equity and democracy.

[4] We investigated the effect of diasporas on conflict duration. However, our data on diasporas covers only a limited number of countries. In our previous analysis of the initiation of conflict (Collier and Hoeffler, 2001a), the inclusion of diasporas as an explanatory variable approximately halved the sample size relative to that used in our baseline model. Nevertheless, with a sample of over 300 observations the main variables in the baseline model remained significant, and the diaspora variable was also significant. In the present analysis, because our sample is necessarily confined only to countries which have experienced a conflict, the reduction in sample size involved in the inclusion of the diaspora variable is much more problematic. We are reduced to only 107 observations and none of the variables in our baseline model remains significant. The diaspora variable is itself insignificant, but since all the variables that are significant in the baseline also become insignificant, nothing can be concluded from our attempt to analyze diasporas.

However, their lack of significant effects on conflict can be contrasted both with their significance in other economic relationships (even when measured with the same data), and with the fact that some other factors have such powerful effects on conflict. For our analysis of the effects of military expenditure we drew upon our previous analysis of its effects on the risk of conflict initiation (Collier and Hoeffler, 2001a). There, to avoid the endogeneity problem inherent in the analysis of the effect of military expenditure on conflict, we instrumented military expenditure using various measures of external threat. We used these same instruments in the present analysis.

The change from those variables that are important for conflict initiation to those which affect its duration permits a speculative interpretation of how conflicts evolve. They start when a rebel organization is both militarily and financially viable. A dispersed population, mountainous terrain, and prior establishment in a neighboring country, all enable the rebel organization to build up an initial military force in remote areas. In such conditions governments are powerless to prevent the emergence of rebellion regardless of preventative military expenditure. The presence of primary commodity exports, an enemy neighboring government, or a rich diaspora, all might be sources of finance for a rebel organization, while if per capita income is low it reduces the cost of recruitment into the organization. Having started because it is viable the rebel organization will continue unless it becomes unviable. It is unlikely to become financially unviable unless it is dependent upon a neighboring government that settles its differences with the host government. Hence, the main threat to continued viability is if the rising military threat posed by the government reaction becomes uncontainable, or if the rebels fall out among themselves. Both forest cover and civil war in a neighbor may be important in sustaining a conflict because they make it more difficult for an increased government military effort to defeat a rebellion. Possibly, the greater difficulty of defeating rebellion in the 1980s and 1990s compared with the previous twenty years is because of the growth of international markets for supplies of armaments to rebels. A financially viable rebel organization can now purchase a wide range of military equipment whereas in an earlier period access was much more dependent upon finding political favor with one or other of the superpowers. During the 1990s many rebel organizations have become in effect entirely private enterprise activities, without political support from any major power. The evolution in the effects of social composition between the initiation and duration of a conflict is surely striking. An ethnic fractionalization score of around 50 does not make a country more prone to the initiation of civil war, but if such a war occurs it tends to lengthen its duration very substantially. This may be because the rebel organization uses ethnicity as a means of maintaining organizational cohesion. If so, it suggests that where there is no potential for ethnicity to be used as a means of rebel identity, many rebel organizations face severe problems of maintaining cohesion: hence the much shorter duration of such wars. Thus, violent conflict appears to become organized on ethnic lines, rather than prior ethnic hatreds escalating into violent conflict.

Recall that one of the results from the analysis of the initiation of conflict is that the risk is temporarily much higher in post-conflict situations. Evidently, something happens during the conflict that increases this risk. For example, one obvious hypothesis is that since conflict causes hatred, it is the accumulated hatred that makes the reversion to war more likely. Our analysis of how the peace hazard changes during conflict provides some evidence on the matter. The peace hazard is the obverse of the risk of conflict *continuation* rather than conflict *initiation*. However, it is reasonable to suppose that the risk of conflict initiation, which is latent during conflict, is monotonic in the risk of conflict continuation. Although the risk of conflict initiation is sharply increased by conflict, during the conflict the risk of conflict continuation does not monotonically increase. It is U-shaped, and is approximately the same in the sixth year of conflict as the first. This does not seem like a plausible time path for hatred, and so it is doubtful if the change in hatred is closely associated with the increased risk of conflict initiation. Rather, it appears that there must be an abrupt increase in the latent risk of subsequent conflict initiation which occurs around the onset of the initial conflict. A possible explanation for such an abrupt change in risk is that it is brought about by the emergence of the rebel organization. The switch from peace to civil war is necessarily coincident with the creation of a large, violent, non-government organization which is viable during conflict. When the war ends it may be because this organization negotiates a deal which is even more advantageous than the conflict, or because it becomes unviable. However, usually end of the conflict does not mark the end of the organization. Hence, in a typical post-conflict situation, a major difference with the pre-conflict period is that a rebel organization already exists. We would therefore expect conflict renewal to be more likely than conflict initiation in otherwise identical circumstances.

4. Conclusion

We have used a comprehensive data set on large-scale violent civil conflicts over the period 1960-99 to analyze the duration of conflict. We have estimated hazard functions, correcting for unobserved heterogeneity. We find that the duration of conflict is determined by a substantially different set of variables from those that determine the initiation of conflict. Notably, the duration of a conflict is substantially increased if the society is composed of a few large ethnic groups, if there is extensive forest cover, and if it commenced since 1980. None of these variables has any significant influence on the risk of conflict initiation. An implication is that the anticipated duration of conflict does not significantly influence the decision to rebel. In turn, this suggests that the initiation of rebellion is not closely related to the prospect of post-victory goals, as assumed in most previous theoretical models. This is consistent with our finding that neither inequality nor political repression prior to conflict is significantly related to either the initiation or the duration of conflict. Instead, our results on the duration of conflict are consistent with our previous analysis of its initiation. There we argued that the critical factors were those that determined whether rebellion was financially and militarily viable as a continuing enterprise during conflict. In this framework, since observed rebellions start out as viable enterprises, they persist unless financial or military circumstances change. Thus, for example, forest cover probably makes it harder for a given government military response to crush a rebellion. Similarly, the emergence of global unpoliced markets in armaments post-1980 may have advantaged rebels.

There is a notable dissonance between the focus of this analysis: the mundane material circumstances that make rebellion viable as an enterprise, and the focus of global political discourse on conflict resolution and prevention: the supposed ultimate goals and aspirations of rebels. An implication of our analysis is that to shorten wars, as to prevent them, policy needs to be refocused on reducing the viability of rebellion as an enterprise.

References

Barrett, D. B., ed. 1982. *World Christian Encyclopedia.* Oxford: Oxford University Press.

Barro, R. J. ed. 1997. *Determinants of Economic Growth.* Cambridge, Massachusetts and London: MIT Press.

Collier, P. 2000. 'Rebellion as a Quasi-Criminal Activity.' *Journal of Conflict Resolution* 44: 839-53.

Collier, P. and A. Hoeffler. 1998. 'On the Economic Causes of Civil War.' *Oxford Economic Papers*, 50:563-73.

_____. 2001. 'Greed and Grievance in Civil War,' Policy Research Working Paper 2355, World Bank, Washington D.C.

_____. 2001a. 'Regional Military Spillovers,' mimeo, World Bank.

Gerrard, A.J.W. 2000. 'What is a Mountain?' Mimeo. DECRG, World Bank.

Grossman, H. I. 1991. 'A General Equilibrium Model of Insurrections.' *American Economic Review* 81:912-21.

_____. 1995. 'Insurrections.' In *Handbook of Defense Economics*, ed. K. Hartley and T. Sandler, 191-212. Vol. 1. : Elsevier Science BV.

Heckman, J. J. and B. Singer (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 47, pp. 247-283.

Jaggers, Keith and Ted Robert Gurr. 1995. 'Tracking Democracy's Third Wave with the Polity III Data'. *Journal of Peace Research* 32:469-482.

Mauro, P. 1995. 'Corruption and Growth'. *The Quarterly Journal of Economics* 110:681-712.

Murdoch, J. C., and T. Sandler. 2001. Economic Growth, Civil Wars, and Spatial Spillovers. mimeo.

Singer, D. J., and M. Small. 1994. *Correlates of War Project: International and Civil War Data, 1816-1992.* Inter-University Consortium for Political and Social Research, Ann Arbor, Michigan.

Summers, R., and A. Heston.. 1991. The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. *The Quarterly Journal of Economics* 99:327-68.

USSR, 1964. *Atlas Naradov Mira*, Department of Geodesy and Cartography of the State Geological Committee of the USSR. Moscow.

Figure 1: The hazard of peace from the raw data (no explanatory variables)

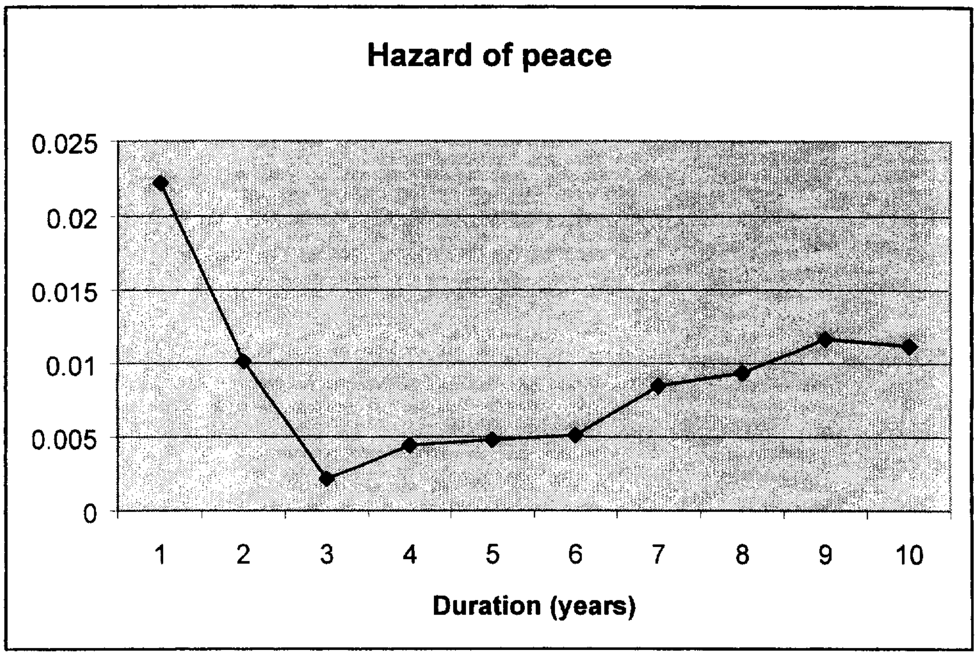

Table 1: Descriptive Statistics

	All Wars, n=52	Wars less than or equal to 48 months, n=21	Wars longer than 48 months, n=31
Per capita income, 1985 const US$	1762 (1431)	2277 (1767)	1414 (1043)
Male secondary school enrolment (%)	29.2 (23.5)	34.8 (26.0)	25.4 (21.2)
Primary commodity exports/GDP	0.13 (0.9)	0.14 (0.12)	0.13 (0.07)
Democracy	2.3 (3.4)	2.3 (3.5)	2.3 (3.3)
Ethnic fractionalization	51.7 (29.1)	43.8 (31.5)	57.0 (26.6)
Religious fractionalization	34.5 (25.1)	30.5 (24.6)	37.2 (25.5)
Mountainous terrain (%)	24.9 (23.7)	28.2 (25.6)	22.7 (22.4)
Forest coverage (%)	23.7 (19.3)	13.7 (11.3)	30.6 (20.7)
Geographic dispersion of the population	0.58 (0.15)	0.52 (0.17)	0.61 (0.13)
Neighbors at war	0.38 (0.49)	0.33 (0.48)	0.42 (0.50)
Prior peace duration (months)	230.2 (142.8)	244.0 (151.2)	220.9 (138.5)

Note: Standard deviations in parentheses.

Table 2: The Collier-Hoeffler Model, no unobserved heterogeneity[$]

	Coefficient	abs [z]
log Real GDP	-0.03	0.10
GDP Growth	-5.10	1.12
Primary commodity exports/GDP	-5.87	1.07
Primary commodity exports/GDP ^{2}	20.39	1.69*
Social Fractionalisation	0.00	0.86
Ethnic dominance	-0.06	0.16
log Peace duration	-0.02	0.12
log Population	-0.07	0.54
Geographic dispersion of the population	-2.00	1.49
d2	-1.40	2.21**
d3	-0.95	1.69*
d4	-0.33	0.68
d5	-0.20	0.52
Intercept	-1.04	0.38
Log likelihood	-86.62	
LR $\chi^2(13)$	19.97	
Prob > χ^2	0.096	
Pseudo R^2	0.10	
Number of observations	51	

Note: *z*-statistics are based on asymptotic standard errors. Significance at the 10%, 5% and 1% level is indicated by *, ** and ***, respectively.

[$] The model with controls for unobserved heterogeneity gave identical results.

Table 3: The Baseline Specification for the Hazard Function

	Model [1]: Without controls for unobserved heterogeneity		Model [2]: With controls for unobserved heterogeneity	
	Coefficient	abs [z]	Coefficient	abs [z]
Ethnolinguistic Fractionalization	-0.10	3.40***	-0.08	2.42**
Ethnolinguistic Fractionalization ^{2} / 100	0.10	3.16***	0.08	2.57***
Male secondary schooling	0.03	2.28**	0.04	3.23***
Neighbours at war	-0.79	2.21**	-0.71	1.98**
Forest coverage	-0.02	2.32**	-0.03	2.93***
Log population	-0.05	0.33	-0.26	1.82*
Religious fractionalization	-0.03	0.88	-0.01	0.33
Religious fractionalization^2 / 100	0.05	0.93	0.04	0.63
1970s	-0.36	0.80	-0.05	0.11
1980s	-1.23	2.11**	-2.26	3.86***
1990s	-2.09	1.85*	-2.51	2.23**
d2	-1.68	2.23**	-1.16	1.49
d3	-1.01	1.57	-0.13	0.19
d4	-0.18	0.57	1.72	3.00***
1990s x d2	1.93	1.20	1.59	0.97
1990s x d3	1.55	0.99	0.88	0.55
1990s x d4	-∞	--	-∞	--
Intercept	-0.69	0.35	0.53	0.25
μ_2			-3.12	5.12***
P			0.69	
Log likelihood	-79.83		-74.92	
LR $\chi^2(13)$	35.52		45.34	
Prob > χ^2	0.00		0.00	
Pseudo R^2	0.18		0.23	
Number of observations	52		52	

Note: z-statistics are based on asymptotic standard errors. Significance at the 10%, 5% and 1% level is indicated by *, ** and ***, respectively.

Table 4: The Baseline Specification with School Enrollment Interacted with Duration

	Model [1]: Without controls for unobserved heterogeneity		Model [2]: With controls for unobserved heterogeneity	
	Coefficient	abs [z]	Coefficient	abs [z]
Ethnolinguistic Fractionalization	-0.11	3.69***	-0.10	3.26***
Ethnolinguistic Fractionalization ^{2} / 100	0.11	3.37***	0.10	3.51***
Male secondary schooling	0.05	3.57***	0.05	3.77***
Neighbours at war	-1.01	2.72***	-1.00	2.66***
Forest coverage	-0.02	1.74*	-0.03	2.28**
Log population	0.07	0.48	-0.11	0.68
Religious fractionalization	-0.04	1.05	-0.01	0.31
Religious fractionalization^2 / 100	0.05	1.07	0.02	0.29
1970s	-0.45	0.98	-0.08	0.19
1980s	-1.35	2.28**	-2.32	4.04***
1990s	-2.61	2.24**	-2.70	2.36***
d2	-0.66	0.67	-0.39	0.40
d3	0.55	0.63	1.18	1.28
d4	1.02	1.56	2.97	3.84***
Male secondary schooling x d2	-0.04	1.31	-0.03	1.03
Male secondary schooling x d3	-0.07	2.03**	-0.06	1.76*
Male secondary schooling x d4	-0.04	2.25**	-0.05	2.36**
y90s x d2	2.62	1.57	2.17	1.26
y90s x d3	2.68	1.63	1.90	1.12
y90s x d4	$-\infty$	--	$-\infty$	--
Intercept	-2.87	1.25	-1.78	0.73
μ_2			-3.18	4.99***
P			0.73	

Continues…

Table 4: Continued

	Model [1]: Without controls for unobserved heterogeneity	Model [2]: With controls for unobserved heterogeneity
Log likelihood	-75.39	-71.15
LR $\chi^2(13)$	46.39	54.87
Prob > χ^2	0.00	0.00
Pseudo R^2	0.24	0.28
Number of observations	52	52

Note: *z*-statistics are based on asymptotic standard errors. Significance at the 10%, 5% and 1% level is indicated by *, ** and ***, respectively.

Appendix Table 1: Sample of 52 Wars

Country	Start of the War	End of the War
Algeria	07/62	12/62
Algeria	05/91	Ongoing
Angola	02/61	05/91
Burma/Myanmar	68	10/80
Burma/Myanmar	02/83	07/95
Burundi	04/72	12/73
Burundi	08/88	08/88
Burundi	11/91	ongoing
Chad	03/80	08/88
Columbia	04/84	ongoing
Cyprus	07/74	08/74
Dominican Rep.	04/65	09/65
El Salvador	10/79	01/92
Ethiopia	07/74	05/91
Guatemala	07/66	07/72
Guatemala	03/78	03/84
Guinea-Bissau	12/62	12/74
India	08/65	08/65
India	84	94
Indonesia	06/75	09/82
Iran	03/74	03/75
Iran	09/78	12/79
Iran	06/81	05/82
Iraq	09/61	11/63
Iraq	07/74	03/75
Iraq	01/85	12/92
Jordan	09/71	09/71
Morocco	10/75	11/89
Mozambique	10/64	11/75
Mozambique	07/76	10/92
Nicaragua	10/78	07/79
Nicaragua	03/82	04/90
Nigeria	01/66	01/70
Nigeria	12/80	08/84
Pakistan	03/71	12/71
Pakistan	01/73	07/77
Peru	03/82	12/96
Philippines	09/72	12/96
Romania	12/89	12/89
Rwanda	11/63	02/64
Rwanda	10/90	07/94
Somalia	04/82	12/92
Sri Lanka	04/71	05/71
Sri Lanka	07/83	ongoing
Sudan	07/83	ongoing
Turkey	07/91	ongoing
Uganda	05/66	06/66
Uganda	10/80	04/88
Yugoslavia	04/90	01/92
Zaire/Dem. Rep. of Congo	07/60	09/65
Zaire/Dem. Rep. of Congo	09/91	12/96
Zimbabwe	12/72	12/79

Appendix: Data Sources

War Duration

A civil war is defined as an internal conflict in which at least 1000 battle related deaths (civilian and military) occurred per year. We use mainly the data collected by Singer and Small (1984, 1994) and according to their definitions we updated their data set for 1992-99.

GDP per capita

We measure income as real PPP adjusted GDP per capita. The primary data set is the Penn World Tables 5.6 (Summers and Heston 1991). Since the data is only available from 1960-92 we used the growth rates of real PPP adjusted GDP per capita data from the World Bank's World Development Indicators 1998 in order to obtain income data for the 1990s.

Primary commodity exports/GDP

The ratio of primary commodity exports to GDP proxies the abundance of natural resources. The data on primary commodity exports as well as GDP was obtained from the World Bank. Export and GDP data are measured in current US dollars.

Population

Population measures the total population, the data source is the World Bank's World Development Indicators 1998. Again, we measure population a the beginning of each sub-period.

Social, ethnolinguistic and religious fractionalization

We proxy social fractionalization in a combined measure of ethnic and religious fractionalization. Ethnic fractionalization is measured by the ethno-linguistic fractionalization index. It measures the probability that two randomly drawn individuals from a given country do not speak the same language. Data is only available for 1960. In the economics literature this measure was first used by Mauro (1995). Using data from Barrett (1982) on religious affiliations we constructed an analogous religious fractionalization index. Following Barro (1997) we aggregated the various religious affiliations into nine categories: Catholic, Protestant, Muslim, Jew, Hindu, Buddhist, Eastern Religions (other than Buddhist), Indigenous Religions and no religious affiliation.

The fractionalization indices range from zero to 100. A value of zero indicates that the society is completely homogenous whereas a value of 100 would characterize a completely heterogeneous society.

We calculated our social fractionalization index as the product of the ethno-linguistic fractionalization and the religious fractionalization index plus the ethno-linguistic or the religious fractionalization index, whichever is the greater. By adding either index we avoid classifying a country as homogenous (a value of zero) if the country is ethnically homogenous but religiously divers, or vice versa

Ethnic dominance
Using the ethno-linguistic data from the original data source (Atlas Naradov Mira, 1964) we calculated an indicator of ethnic dominance. This variable takes the value of one if one single ethno-linguistic group makes up 45 to 90 percent of the total population and zero otherwise.

Geographic Dispersion of the Population
We constructed a dispersion index of the population on a country by country basis. Based on population data for $400km^2$ cells we generated a Gini coefficient of population dispersion for each country. A value of 0 indicates that the population is evenly distributed across the country and a value of 1 indicates that the total population is concentrated in one area. Data is available for 1990 and 1995. For years prior to 1990 we used the 1990 data. We would like to thank Uwe Deichman of the World Bank's Geographic Information System Unit for generating this data. He used the following data sources: Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); and World Resources Institute (WRI). 2000. Gridded Population of the World (GPW), Version 2. Palisades, NY: IESIN, Columbia University. Available at http://sedac.ciesin.org/plue/gpw.

Forest Coverage
We used the FAO measure of the proportion of a country's terrain which is covered in woods and forest. Source: http://www.fao.org/forestry

Mountainous terrain
The proportion of a country's terrain which is mountainous was measured by John Gerrard, a physical geographer specialized in mountainous terrain. His measure is not only based on altitude but takes into account plateaus and rugged uplands. The data is presented in Gerrard (2000).

Peace Duration
This variable measures the length of the peace period since the end of the previous civil war. For countries which never experienced a civil war we measure the peace period since the end of World War.

Neighbors at War
This binary variable indicates whether any of the neighbours was at war at the time the conflict began. We would like to thank James Murdoch and Todd Sandler for the use of their neighbourhood data set. Countries sharing land boundaries are classified as neighbours, island nations do not have any neighbours.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2655	Measuring Services Trade Liberalization and its Impact on Economic Growth: An Illustration	Aaditya Mattoo Randeep Rathindran Arvind Subramanian	August 2001	L. Tabada 36896
WPS2656	The Ability of Banks to Lend to Informationally Opaque Small Businesses	Allen N. Berger Leora F. Klapper Gregory F. Udell	August 2001	A. Yaptenco 31823
WPS2657	Middle-Income Countries: Development Challenges and Growing Global Role	Peter Fallon Vivian Hon Zia Qureshi Dilip Ratha	August 2001	D. Fischer 38656
WPS2658	How Comparable are Labor Demand Elasticities across Countries?	Pablo Fajnzylber William F. Maloney	August 2001	A. Pillay 88046
WPS2659	Firm Entry and Exit, Labor Demand, and Trade Reform: Evidence from Chile and Colombia	Pablo Fajnzylber William F. Maloney Eduardo Ribeiro	August 2001	A. Pillay 88046
WPS2660	Short and Long-Run Integration: Do Capital Controls Matter?	Graciela Kaminsky Sergio Schmukler	August 2001	E. Khine 37471
WPS2661	The Regulation of Entry	Simeon Djankov Rafael La Porta Florencio Lopez de Silanes Andrei Shleifer	August 2001	R. Vo 33722
WPS2662	Markups, Entry Regulation, and Trade: Does Country Size Matter?	Bernard Hoekman Hiau Looi Kee Marcelo Olarreaga	August 2001	L. Tabada 36896
WPS2663	Agglomeration Economies and Productivity in Indian Industry	Somik Lall Zmarak Shalizi Uwe Deichmann	August 2001	R. Yazigi 37176
WPS2664	Does Piped Water Reduce Diarrhea for Children in Rural India?	Jyotsna Jalan Martin Ravallion	August 2001	C. Cunanan 32301
WPS2665	Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree?	Martin Ravallion	August 2001	C. Cunanan 32301
WPS2666	Measuring Pro-Poor Growth	Martin Ravallion	August 2001	C. Cunanan 32301

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2667	Trade Reform and Household Welfare: The Case of Mexico	Elena Ianchovichina Alessandro Nicita Isidro Soloaga	August 2001	L. Tabada 36896
WPS2668	Comparative Life Expectancy in Africa	F. Desmond McCarthy Holger Wolf	August 2001	H. Sladovich 37698
WPS2669	The Impact of NAFTA and Options for Tax Reform in Mexico	Jorge Martinez-Vazquez Duanjie Chen	September 2001	S. Everhart 30128
WPS2670	Stock Markets, Banks, and Growth: Correlation or Causality?	Thorsten Beck Ross Levine	September 2001	A. Yaptenco 31823
WPS2671	Who Participates? The Supply of Volunteer Labor and the Distribution of Government Programs in Rural Peru	Norbert R. Schady	September 2001	T. Gomez 32127
WPS2672	Do Workfare Participants Recover Quickly from Retrenchment?	Martin Ravallion Emanuela Galasso Teodoro Lazo Ernesto Philipp	September 2001	C. Cunanan 32301
WPS2673	Pollution Havens and Foreign Direct Investment: Dirty Secret or Popular Myth?	Beata K. Smarzynska Shang-Jin Wei	September 2001	L. Tabada 36896
WPS2674	Measuring Economic Downside Risk and Severity: Growth at Risk	Yan Wang Yudong Yao	September 2001	A. Rivas 36270
WPS2675	Road Infrastructure Concession Practice in Europe	Franck Bousquet Alain Fayard	September 2001	G. Chenet-Smith 36370
WPS2676	An Alternative Unifying Measure of Welfare Gains from Risk-Sharing	Philippe Auffret	September 2001	K. Tomlinson 39763
WPS2677	Can Local Institutions Reduce Poverty? Rural Decentralization in Burkina Faso	Paula Donnelly-Roark Karim Ouedraogo Xiao Ye	September 2001	E. Hornsby 33375
WPS2678	Emerging Markets Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns?	Graciela Kaminsky Sergio Schmukler	September 2001	E. Khine 37471
WPS2679	"Deposit Insurance Around the Globe: Where Does It Work?	Aslı Demirgüç-Kunt Edward J. Kane	September 2001	K. labrie 31001
WPS2680	International Cartel Enforcement: Lessons from the 1990s	Simon J. Evenett Margaret C. Levenstein Valerie Y. Suslow	September 2001	L. Tabada 36896