ISSN 0258-6770

THE WORLD BANK ECONOMIC REVIEW

Volume 22 • 2008 • Number 1

www.wber.oxfordjournals.org

OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

THE WORLD BANK ECONOMIC REVIEW

The World Bank Economic Review is a professional journal for the dissemination of World Bank–sponsored and other research that may inform policy analysis and choice. It is directed to an international readership among economists and social scientists in government, business, international agencies, universities, and development research institutions. The Review seeks to provide the most current and best research in the field of quantitative development policy analysis, emphasizing policy relevance and operational aspects of economics, rather than primarily theoretical and methodological issues. It is intended for readers familiar with economic theory and analysis but not necessarily proficient in advanced mathematical or econometric techniques. Articles illustrate how professional research can shed light on policy choices. Consistency with World Bank policy plays no role in the selection of articles.

Articles are drawn from work conducted by World Bank staff and consultants and by outside researchers. Non-Bank contributors are encouraged to submit their work.

Before being accepted for publication, articles are reviewed by three referees—one from the World Bank and two from outside the institution. Articles must also be endorsed by two members of the Editorial Board before final acceptance.

For more information, please visit the Web sites of the Economic Review at Oxford University Press at www.wber.oxfordjournals.org and at the World Bank at www.worldbank.org/research/journals.

Instructions for authors wishing to submit articles are available online at www.wber.oxfordjournals.org. Please direct all editorial correspondence to the Editor at wber@worldbank.org.

THE WORLD BANK ECONOMIC REVIEW

Volume 22 · 2008 · Number 1

A SYMPOSIUM ON POST-CONFLICT TRANSITIONS

SUBSCRIPTIONS: A subscription to *The World Bank Economic Review* (ISSN 0258-6770) comprises 3 issues. Prices include postage; for subscribers outside the Americas, issues are sent air freight.

Annual Subscription Rate (Volume 21, 3 Issues, 2008): *Academic libraries*—Print edition and site-wide online access: US$182/£118/€182, Print edition only: US$173/£112/€173, Site-wide online access only: US$173/£112/€173; *Corporate*—Print edition and site-wide online access: US$273/£177/€273, Print edition only: US$259/£168/€259, Site-wide online access only: US$259/£168/€259; *Personal*—Print edition and individual online access: US$59/£38/€59. Please note: £ Sterling rates apply in the UK, € in Europe, and US$ elsewhere. There may be other subscription rates available; for a complete listing, please visit www.wber.oxfordjournals.org/subscriptions. *Readers with mailing addresses in non-OECD countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request by writing to the UK address below.*

Full prepayment in the correct currency is required for all orders. Orders are regarded as firm, and payments are not refundable. Subscriptions are accepted and entered on a complete volume basis. Claims cannot be considered more than four months after publication or date of order, whichever is later. All subscriptions in Canada are subject to GST. Subscriptions in the EU may be subject to European VAT. If registered, please supply details to avoid unnecessary charges. For subscriptions that include online versions, a proportion of the subscription price may be subject to UK VAT. Personal rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address.

BACK ISSUES: The current year and two previous years' issues are available from Oxford University Press. Previous volumes can be obtained from the Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. E-mail: psc@periodicals.com. Tel: (518) 537-4700. Fax: (518) 537-5899.

CONTACT INFORMATION: Journals Customer Service Department, Oxford University Press, Great Clarendon Street, OxfordOX2 6DP, UK. E-mail: jnls.cust.serv@oxfordjournals.org. Tel: +44 (0)1865 353907. Fax: +44 (0)1865 353485. *In the Americas, please contact*: Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513, USA. E-mail: jnlorders@oxfordjournals.org. Tel: (800) 852-7323 (toll-free in USA/Canada) or (919) 677-0977. Fax: (919) 677-1714. *In Japan, please contact*: Journals Customer Service Department, Oxford University Press, 4-5-10-8F Shiba, Minato-ku, Tokyo, 108-8386, Japan. E-mail: okudaoup@po.iijnet.or.jp. Tel: +81 3 5444 5858. Fax: +81 3 3454 2929.

POSTAL INFORMATION: *The World Bank Economic Review* (ISSN 0258-6770) is published three times a year, in Feb., Jun., and Oct., by Oxford University Press for the International Bank for Reconstruction and Development/THE WORLD BANK. Send address changes to *The World Bank Economic Review*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Economic Review*, The World Bank, 3, Chemin Louis Dunant, CP66 1211 Geneva 20, Switzerland.

DIGITAL OBJECT IDENTIFIERS: For information on dois and to resolve them, please visit www.doi.org.

PERMISSIONS: For information on how to request permissions to reproduce articles or information from this journal, please visit www.oxfordjournals.org/jnls/permissions.

PAPER USED: *The World Bank Economic Review* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

INDEXING AND ABSTRACTING: *The World Bank Economic Review* is indexed and/or abstracted by *CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/EconLit, PAIS International, RePEc (Research in Economic Papers)*, and *Social Services Citation Index*.

Postconflict Transitions: An Overview

Ibrahim Ahmed Elbadawi

In the two to five years immediately following end of conflicts, UN peacekeeping operations have succeeded in maintaining peace, while income and consumption growth rates have been higher than normal and recovery on key education and health indicators has been possible. Aid also has been super-effective in promoting recovery, not only by financing physical infrastructure but also by helping in the monetary reconstruction of postconflict economies. However, sustaining these short-term gains was met with two difficult challenges. First, long-term sustainability of peace and growth hinges primarily on the ability of postconflict societies to develop institutions for the delivery of public goods, which, in turn, depends on the capacity of postconflict elites to overcome an entrenched culture of political fragmentation and form stable national coalitions, beyond their immediate ethnic or regional power bases. Second, after catch-up growth runs its course, high levels of aid could lead to overvalued real currencies, at a time when growth requires a competitive exchange rate and economic diversification. Successful peace-building would, therefore, require that these political and economic imperatives of postconflict transitions be accounted for in the design of UN peacekeeping operations as well as the aid regime.

Transitioning from civil war to sustainable peace is complex. Success starts with basic transitions: from warfare to peacekeeping to ensure the security of a country and its communities; from large armies to public agencies of law and order to ensure the security of individuals and their property rights; from military- to civilian-dominated state institutions. These transitions are necessary to effectively manage aid and establish the policies and institutions needed for sustained and widely shared growth.

The five articles of this symposium deal with postconflict transition. The articles address UN peacekeeping operations and the prospects for sustainable peace, the role of policy credibility in preventing war and promoting peace, country performance in the aftermath of civil war, and the consequences of aid for macrostability and growth. All of the articles analyze panel data of global samples of conflict and nonconflict countries covering at least 30 years, drawing on extensive new evidence on postconflict economies.

Ibrahim Elbadawi is a lead economist in the Development Economics Research Group of the World Bank; his email address is ielbadawi@worldbank.org.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 1–7 doi:10.1093/wber/lhn002
Advance Access Publication February 20, 2008

I. Peacebuilding and Political Credibility

Two articles deal with peacebuilding and the related restoration of political credibility.

Peacebuilding and UN Peacekeeping Operations

The article by Nicholas Sambanis sets the stage by examining the effectiveness of UN peacekeeping operations in maintaining "sustainable" or "participatory" peace (that is, the capacity of a sovereign state to resolve by means other than war the natural conflicts to which all societies are prone). Peace involves not just the end to war but the absence of significant residual violence, undivided sovereignty, and a minimum level of political openness.

According to Sambanis's model, the probability of peacebuilding success depends on the degree of hostility and the extent of local and international competencies for peacebuilding.[1] Maximizing the peacebuilding space requires greater commitment by the international community.

Holding other control variables at their sample means, Sambanis estimates that going from facilitative (monitoring and reporting) to transformational (multidimensional, enforcement, and transitional administration) peacekeeping increases the probability of peacebuilding success by 36 percent.[2] He also investigates the long-run impact of UN missions on the duration of peace, defined as the absence of war (or "negative peace"). Using a survival probability model, he finds that UN intervention is still highly significant and reduces the risk of peace failure by 50 percent. But the strongest influence on reduced proneness to a new war outbreak, consistent with much of the literature on civil war onset, is associated with local capacity variables. Fast-growing, rapidly diversifying, and high-income economies are far more likely than others to experience longer peace duration, even when hostilities remain high.

Political Credibility and Peacebuilding

Philip Keefer suggests lack of political credibility as an additional, if not alternative, factor influencing the risk of civil war.[3] He argues that because weakly credible leaders cannot persuade the majority of their citizens to believe their promises, they have no incentive to make broad-based programmatic policies. Instead, they are likely to undertake policies that benefit the few groups that believe their promises. This tendency leads to the underprovision of public

1. Doyle and Sambanis (2000) first developed this insight.

2. See Doyle and Sambanis (2006) for extensive discussions of mandates of UN peacekeeping operations.

3. The recent empirical literature has been dominated by the debate over whether the hazard of the onset of civil war is better explained by "greed" or "grievance" factors. The dominant view, espoused by Fearon and Laitin (2003) and Collier and Hoeffler (2004a), is that greed factors are more important (see Bodea and Elbadawi 2007 for a critique of this view).

goods, the overprovision of private goods, and high levels of corruption. In such a distorted environment, he argues, citizens are less likely to resist potential insurgency or other efforts to unseat incumbent elites. Lack of political credibility also raises the risk of conflicts by weakening resistance to an incipient counterinsurgency. Moreover, weakly credible leaders are incapable of mounting effective counterinsurgency efforts, because they are inept at making credible commitments to counterinsurgents.

Keefer measures political credibility using continuous years of competitive elections for democracies and the presence of institutionalized or programmatic political parties for autocracies. Embedding these two variables in a standard model of civil war onset, he obtains broad support for his hypothesis across several samples.

Given the connection between civil war and the accumulated effects of low or negative growth, political credibility and the provision of public goods—and hence growth—should have strong implications for peacebuilding. Keefer's contribution is important because it goes to the deep institutional issue of what determines the provision of public goods, growth, and the risk of conflict as well as the prospects for sustained peace when war ends.

II. Performance and Policies in Postconflict Economies

Civil war is considered the most devastating type of conflict, because it has much more destructive effects on economic activities, social capital, and institutions than multistate wars do. The collapse of income and economic and political insecurity associated with civil wars causes agents to divert their assets and factors of production toward subsistence sectors and to move their capital out of the country (Collier 1999). The associated loss of national income reduces revenues at the very time that public investment is needed to rebuild and protect the country.

The last three articles deal with the quantitative economic performance of postconflict countries. The first examines basic indicators of economic performance, health and education, political development, demographic trends, and conflict and security issues before and after civil wars. The other two explore two neglected macroeconomic aspects of postconflict policies: the role of aid with respect to "monetary" reconstruction and the equilibrium path of the real exchange rate (RER).

The Aftermath of Civil War

Siyan Chen, Norman V. Loayza, and Marta Reynal-Querol analyze the cost of war and the prospects for recovery in the aftermath of war in cases in which resolution of conflict led to at least 10 years of uninterrupted peace. The event framework is convenient for assessing how postconflict countries perform both absolutely and relative to two control groups, one drawn from a global sample of comparable nonconflict countries, the other from the

geographic region of the conflict country. The most striking result pertains to the tremendous postwar surge in per capita income, which rises about 2.4 percent above the prewar level. The strong recovery in income is linked to the high potential for catch-up growth following the destruction of war. The nonmonotonic trend of per capita growth suggests that aid flows are time sensitive (Collier and Hoeffler 2004b). The growth rebound is also associated with a rapid decline in inflation and the realignment of fiscal policy away from military expenditure.

Interestingly, in some areas of social development (such as infant mortality and primary-school enrollment), postconflict countries have made progress despite war. The authors suggest that this progress may reflect the ability of these countries to partake in the global public goods associated with innovations in medicine and service delivery in education and health. Despite such progress, the cost of war is substantial, as reflected in the absolute and relative decline in income in conflict countries and their slower progress in improving political and economic governance or social indicators tied to combatants (such as adult male mortality and secondary-school enrollment).

Aid and "Monetary" Reconstruction

Christopher Adam, Paul Collier, and Victor Davies discuss the financing implications for the government of the decline in the demand for money associated with reduced income and asset substitution away from domestic money, which, they argue, is likely to worsen the trade-off between seigniorage and inflation. Because the government needs to finance higher military expenditure during conflict and borrowing is not likely to be an option, however, the discount rate—and the inflation rate the government is willing to tolerate—increases. As a result of the slow recovery in the demand for money in the aftermath of conflict and the high level of financing, inflation is not likely to decline even after war ends. The key insight of this article is that without aid, conflict and postconflict countries are likely to experience explosive inflation; aid can help reestablish the preconflict equilibrium level of inflation. In this sense, aid can finance postconflict "monetary" reconstruction akin to its better known role with regard to physical reconstruction in the aftermath of war.

The empirical results suggest that postconflict aid stimulates the demand for money—directly, by substituting for seigniorage, and indirectly, by restoring income growth and supporting a modest portfolio shift in favor of domestic money. Simulations of the model suggest that plausible levels of postwar aid sustained over a 10-year period would restore real money balances to about 2.8 percentage points of gross domestic product (GDP) (an increase of about 50 percent over the end-of-conflict level) and cut inflation by almost a half, to just over 10 percent a year.

Aid and Real Exchange Rate Competitiveness

During civil war, the same factors that led to the decline in income and the demand for money disproportionately affect the tradables sector. The extent of recovery in this sector is thus likely to have a significant effect on postconflict growth. Aid can directly contribute to the growth of the tradables sector as well as to nontradables productive and service sectors, such as infrastructure and financial services.

Elbadawi, Kaltani, and Schmidt-Hebbel provide evidence that despite receiving substantial aid flows, postconflict countries experienced only moderate RER overvaluation during the postconflict cycle. Although aid is robustly associated with RER in the long run, it is not a major factor in explaining the disequilibrium behavior of the RER.[4] Using the RER misalignment index in a growth model encompassing aid effectiveness, they corroborate evidence that aid promotes growth but with diminishing returns. They also find support for the views that RER overvaluation has direct negative level effects on growth and that financial development is positively associated with growth. Accounting for these level effects and other traditional growth controls, they find that RER overvaluation further reduces growth through its interaction with aid. In contrast, the interaction effect between RER overvaluation and financial development is positively associated with growth. Simulations of the effect of a one standard deviation increase in RER overvaluation suggest that the loss in per capita growth for postconflict countries that are highly dependent on aid and have weak financial sector could be as high as half a percent.

III. Design of Peacekeeping Operations and Timing and Delivery of Aid

UN peacekeeping operations—the dominant instrument for peacebuilding in the post-Soviet era—have succeeded in maintaining peace in the two to five years after the end of civil wars. Postconflict countries have enjoyed higher than normal growth rates in income and consumption. Development aid has been particularly effective in promoting recovery, not only by financing physical infrastructure but also by helping in the monetary reconstruction of postconflict economies.

Unfortunately, these short-term gains are not always sustained in the longer run, especially after the UN mission ends. Although the support of the United Nations, the World Bank, and other external actors remains critical, the long-term sustainability of peace and growth hinges primarily on the ability of postconflict societies to develop institutions for the delivery of public goods.

4. This is likely to be partly because, although aid remained high in postconflict countries, it was sustained over the 10-year postconflict period. Like most aid-recipient countries, these countries may also have managed their aid in a way that protects their macroeconomic competitiveness (Elbadawi, Kaltani, and Soto 2007).

Creating such institutions is particularly challenging in these countries, however, for two main reasons. First, most peacebuilding initiatives are based on wealth- and power-sharing agreements across regions and communities within a country. In the short-run, such agreements may be necessary to redress historical grievances and hence promote peace. In the longer run, however, they may inculcate a political culture of subnational retrenchment and entitlements that impedes the ability of political elites to establish viable national coalitions. Because a fragmented political class is not likely to be able to make credible promises outside its immediate region or ethnic group, public goods are underprovided, corruption is high, growth is suboptimal, and the risks of future conflicts may be high. Second, after postconflict economies exhaust their catch-up growth potential, high levels of aid could lead to overvalued real currencies, at a time when growth requires a competitive exchange rate and economic diversification.

Given the political and economic imperatives of postconflict transitions, how UN peacekeeping operations are designed, mandated, and supported and how and when aid is delivered appear to be the keys to success.

Funding

This article and the other five in this symposium are products of the Growth and Macroeconomics Team, Development Economic Research Group of the World Bank, and are part of the project Post-conflict Transitions, focusing on development issues particularly relevant to countries recovering from civil wars. This project was funded by a generous grant from the multi-donor Knowledge for Change Program (KCP) Trust Fund.

References

Adam, Christopher, Paul Collier, and Victor Davies. 2008. "Postconflict Monetary Reconstruction" *World Bank Economic Review* 22(1). Advance Access published January 14, 2008, doi:10.1093/wber/lhm020.

Bodea, Cristina, and Ibrahim A. Elbadawi. 2007. "Riots, Coups and Civil War: Revisiting the Greed and Grievance Debate." Policy Research Working Paper 4397. World Bank, Washington, D.C.

Chen, Siyan, Norman V. Loayza, and Marta Reynal-Querol. 2008. "The Aftermath of Civil War" *World Bank Economic Review* 22(1). Advance Access published February 10, 2008, doi:10.1093/wber/lhn001.

Collier, Paul. 1999. "On the Economic Consequences of Civil War." *Oxford Economic Papers* 51 (1): 168–83.

Collier, Paul, and Anke Hoeffler. 2004a. "Greed, Grievance and Civil War." *Oxford Economic Papers* 56(4):563–95.

———. 2004b. "Aid, Policy and Growth in Post-Conflict Societies." *European Economic Review* 48(5):1125–45.

Doyle, Michael W., and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *American Political Science Review* 94(4):779–801.

———. 2006. *Making War and Building Peace: United Nations Peace Operations*. Princeton, N.J.: Princeton University Press.

Elbadawi, Ibrahim A., Linda Kaltani, and Raimundo Soto. 2007. "Real Exchange Rate Misalignment in SSA: How Serious, How Dangerous?" Development Economic Research Group, October, World Bank, Washington, D.C.

Elbadawi, Ibrahim A., Linda Kaltani, and Klaus Schmidt-Hebbel. 2008. "Foreign Aid, the Real Exchange Rate, and Growth in the Aftermath of Civil Wars." *World Bank Economic Review* 22(1). Advance Access published February 7, 2008, doi:10.1093/wber/lhm024.

Fearon, James, and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–89.

Keefer, Philip. 2008. "Insurgency and Credible Commitment in Autocracies and Democracies." *World Bank Economic Review* 22(1). Advance Access published January 5, 2008, doi:10.1093/wber/lhm019.

Sambanis, Nicholas. 2004. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6):814–58. Advance Access published January 31, 2008, doi:10.1093/wber/lhm022.

———. 2008. "Short- and Long-term Effects of United Nations Peace Operations." *World Bank Economic Review* 22(1).

Short- and Long-Term Effects of United Nations Peace Operations

Nicholas Sambanis

In an earlier study Doyle and Sambanis (2000) [Doyle, Michael W., and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *American Political Science Review* 94(4):779–801.] showed that United Nations (UN) peace operations have made positive contributions to peacebuilding in the short term, helping parties implement peace agreements. But are the effects of UN peace operations lasting? Because the UN cannot fight wars, such operations should not be used to enforce a peace. Peacekeeping operations contribute more to the quality of the peace—that is, to securing more than the mere absence of war—than to its duration, because the effects of such operations dissipate over time. For peace to be self-sustaining, countries must develop institutions and policies that generate economic growth. UN peacebuilding lacks a strategy for fostering self-sustaining economic growth that could connect increased participation with sustainable peace. The international community would benefit from an evolution that uses economic reforms to plug the gap between peacekeeping and humanitarian assistance on the one hand and development on the other. JEL codes: D74, F35, H56

In the first quantitative study of the effects of United Nations (UN) peacekeeping in post–civil war transitions Doyle and Sambanis (2000) presented evidence that UN peace missions can help shore up the foundations for successful peacebuilding. They proposed a simple model in which peacebuilding outcomes are a function of three factors: the level of hostility prevalent at the start of the peace process, local capacities for reconstruction and development, and international capacities for peacebuilding (mainly UN peace operations). Controlling for levels of war-related hostility and pre- and postwar levels of local capacities, they found that UN peace missions have had a significant impact, increasing the probability that peace is established following civil war.

Nicholas Sambanis is professor of political science at Yale University; his email address is nicholas.sambanis@yale.edu. This article draws on joint research with Michael Doyle, whom the author thanks for many discussions. Appreciation is also expressed to Jim Fearon, Alan Gerber, Don Green, Alan Krueger, Alexandros Michaelides, and Kenneth Schevefor for useful comments on methodological issues and to Ana Maria Arjona, Katie Glassmyer, Jonah Schulhofer-Wohl, Steve Shewfelt, Douglas Woodwell, and Annalisa Zinn for help with data collection.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 9–32 doi:10.1093/wber/lhm022
Advance Access Publication January 31, 2008

This article extends the earlier study in a number of ways. It conceptualizes successful peacebuilding as the achievement of self-sustaining peace, a view that is more consistent with the policy community's understanding of peace. It analyzes the effect of UN missions (using a definition of peace that includes a modest measure of democratization) and identifies the determinants of the duration of peace (defined simply as the absence of war). The analysis looks beyond the short-term effects of UN missions to consider whether countries that receive UN assistance enjoy better outcomes over the long run than those that do not. The results of extensive robustness tests of the empirical analyses are presented in order to establish the positive effects of UN missions.[1]

The article presents several new empirical results. It shows that the same factors that may lead parties to the negotiation table make implementation of peace agreements more difficult. In countries with deep hostilities after long and bloody wars, for example, factions are more likely to sign a peace treaty when they realize they cannot win a military victory. By that point, however, levels of trust are so low and local capacities so depleted that implementing a peace agreement may prove difficult without external assistance. It is precisely such settings that UN assist in peace implementation can prove invaluable. The UN can help by providing assurances through monitoring and policing and by supporting institution-building efforts by providing technical assistance.

The UN has a greater effect on ensuring "participatory peace" than on securing the mere absence of war.[2] UN missions can have lasting effects if they help keep the peace in the early stages of the peace process, when the risk of a return to war is greatest. They can also help by designing or helping to build institutions that can make the peace self-sustaining. Over the longer term the effects of UN missions diminish; self-sustaining peace relies more on a return to economic growth and development.

The article is organized as follows. Section I discusses conceptualizations of peacebuilding that inform the measure of peacebuilding success used in the analysis. Section II presents the model, the data, and empirical results on the short-term effects of UN missions. Section III discusses how the non-random assignment of UN missions can affect estimates of the effects on UN peacebuilding efforts. Section IV analyzes the longer-term effects of UN missions. Section V concludes with a discussion of policy implications.

I. What Constitutes a Peacebuilding Success?

Peace can be thought of as a continuum, ranging from no peace (war) to "negative peace" (absence of war) to social harmony (Boulding 1964). Social harmony

1. These robustness tests are too extensive to present here. They are posted on the author's Web site, which provides the data needed to replicate them (http://pantheon.yale.edu/~ns237/index/research.html#Peace).

2. For elaboration of the argument that the UN should not fight wars and relevant case evidence, see Doyle and Sambanis (2006).

is an elusive goal for most societies. What standard of peace, then, should be the goal for societies emerging from civil war?[3] The mere absence of war does not reflect what is needed for peace to be self-sustaining in troubled societies; the best standard is what Doyle and Sambanis (2006) call *participatory peace*, a state that involves an end to the war, the absence of significant residual violence, undivided sovereignty, and a minimum level of political openness (participatory peace is inconsistent with extreme authoritarianism).

Peace can also be defined more narrowly, as the mere absence of war. This article analyzes both the determinants of peace defined narrowly as the absence of war and the broader concept of participatory peace. It argues that attaining the goals set by the participatory peace standard helps set the foundations for self-sustaining peace in the long run.

Participatory peace is meaningful if it can be sustained after the peacekeepers leave. In extensive discussions in the UN Security Council, sustainable peace was proposed as the ultimate purpose of all peace operations, with *sustainability* defined as the capacity for a sovereign state to resolve by means other than war the conflicts to which all societies are prone. "Peace-building," Secretary-General Kofi Annan noted, "is an attempt, after a peace has been negotiated or imposed, to address the sources of present hostility and build local capacities for conflict resolution" (Annan 2001, p. 2). Few observers would think that peace had been successfully built in countries in which armed peacekeepers must remain to deter attacks that would derail the peace.

This article evaluates peacebuilding outcomes, recognizing that a peace that lasts without external assistance is more solid than one that requires the UN to hold the country together. Not all civil wars have had been followed by UN peace operations; among those that did the UN departed soon after the end of the war in some cases and stayed on for several years to manage a fragile peace process in others. This makes it difficult to determine just when the war really ended, as UN reconstruction and peacebuilding mandates sometimes require a UN presence for several years after the end of armed conflict. Although it is easier to evaluate all peace processes at the same point (say, two or five years after the cessation of major hostilities), doing so would be inconsistent with the concept of self-sustaining peace that is the goal of UN intervention.

One way around this problem is to evaluate peacebuilding outcomes two years after the peace process starts, redefining the end of the war to include the peacekeepers' departure. A peace treaty, military victory, or the completion of a UN operation could mark the start of the peacebuilding process.

There is no a priori clear relation between the way a civil war ends and how long the peace process lasts. Wars that end in peace treaties typically last longer

3. Civil wars are defined as large armed conflicts between the government of a sovereign state and domestic challengers able to cause significant destruction in reciprocal violence. Discussion of the concept and measures of civil war, as well as detailed coding notes for the measure used here, can be found in Sambanis (2004).

than wars that end in victories, but the peace process after such wars may be shorter if the treaty provides a roadmap for a more rapid transition to peace.

UN peace operations also differ in their duration and mandate. There is no clear relation a priori between the duration of a UN mission and the probability of peacebuilding success. How long or how short UN missions must be to succeed depends on underlying conditions. A long mission may suggest that the peacebuilding environment is difficult; it also gives the UN more opportunities to fail. The opposite can also be true: the longer the peace is kept, the more likely it is that a stable peace will be built, partly because of the contribution of a UN mission that stays the course.

In all cases examined here the effect of the "therapy" on the eventual health of the "patient" country is evaluated two years after the therapy is complete. Treaties and victories are "therapies" with obvious dates. A UN peace "therapy" ends when military forces are withdrawn.

Participatory peacebuilding success is a binary variable, coded 1 if all the components of participatory peace are satisfied two years after the end of the war (including the departure of the peacekeepers). If war resumes before then, if sovereignty is divided, if there is significant residual violence, or if the country's Polity score (level of political openness) is below a minimum threshold, the mission is coded 0 (peace failure).[4] Peacebuilding failures can occur while the UN is still there; successes cannot. This sets a high bar for success, causing several cases of successful peacekeeping (that is, cases in which there is participatory peace but the UN remains in place) to be dropped.[5]

The data used in the empirical analysis include peace processes following all civil wars that started after December 31, 1945, and ended before January 1, 2000. Each peace transition is an observation; wars that were ongoing at the end of December 1999 are excluded, unless a significant peace process had begun before then. If a peace process starts and fails immediately, it is coded as a peace failure in the first month.[6] This coding rule leaves 121 cases, 84 of which are coded as failures and 37 of which are coded as successes of participatory peace.

4. The Polity threshold (an index measuring regime characteristics and classifying countries along a democracy-autocracy range) is low (3 on a 20-point scale). This threshold excludes extreme autocracies and the most repressive regimes. This component of the peacebuilding definition excludes "peaces of the grave" ("peaces" in which all of the enemy are dead or in prison).

5. Results are stronger with the DS2000 definition of participatory peace, which is similar to that used here but does not drop cases in which the UN mission was still present when peacebuilding outcomes were measured.

6. In some cases the UN sends troops before the war ends or already has troops in place when a new war starts (an example is Angola, where a new war began in 1997, while UN peacekeepers were on the ground). In such cases if the war does not end within the two-year period, it is coded a peacebuilding failure. In the analysis of the long-run effects of UN missions, if the violence continues (or intensifies) while UN troops are deployed, it is coded a peace failure during the first month of the peace process (an example is Somalia). An argument can be made that such cases should be dropped, because cases of enforcement missions present special challenges. If all such cases are dropped, the results for UN missions become much stronger.

II. Determinants of Peacebuilding Success in the Short Term

The main hypothesis tested here is that UN peace operations contribute positively to the peace, controlling for other relevant factors. Drawing on the DS2000 model, the analysis uses different empirical measures for each of the three dimensions that should shape peacebuilding outcomes (hostility, international capacities, and local capacities). The explanatory variables are described briefly here; Doyle and Sambanis (2006) provide an extensive discussion of the rationale for including them in the model.

Explanatory Variables

Hostility is measured by the number of deaths and displacements, the number of factions, the signing of a peace treaty, the type of war (ethno-religious or not), the level of ethnic fractionalization, and the war's duration.[7] Greater hostility (ethno-religious war, high fractionalization, no treaty, many factions, long wars, and many deaths and displacements) should make peacebuilding success less likely.

Local capacities are measured with country-level indicators of socioeconomic development. The overall level of development is measured by electricity consumption per capita or by the annual rate of change in real per capita income and the level of real per capita income—or by all three measures. A higher level of development should help peacebuilding, because the country can more easily rebuild after war.

Another local capacity indicator works in the opposite direction: if the economy is heavily dependent on natural resources, peacebuilding will be more difficult, because the economy will be more susceptible to external commodity price shocks and because resource-rich economies tend to have weak or corrupt political institutions. Resource dependence can be measured using primary commodity exports as a percentage of GDP or a binary indicator of oil export–dependent countries (defined as a country in which fuel exports make up more than a third of total merchandise exports). Local capacities can be seen as measures of institutional quality and of the economic opportunity costs of returning to war: greater capacities imply higher opportunity costs and better institutions, hence a better chance of building peace.

Deficiencies in local capacities and intense postwar hostility may be offset by international capacities. The key measures of international capacities are the presence of UN peace operations, the predominant form of multilateral peace operation since 1945.[8] The mandates of UN missions are coded based on a close reading of each mission's operational guidelines, status of forces agreements

7. Summary statistics and additional controls are presented in the Supplemental Appendix to this article, available at http://wber.oxfordjournals.org/

8. Controlling for the presence of non–UN missions (which do not seem to be effective) does not affect the results (see the additional results on the author's Web site). Shimizu and Sandler (2002) show that there is a public good component to peacekeeping, as evidenced by the pattern of contributions, which tend to be borne disproportionately by large countries in both UN and non.

(where available), and UN documents that indicate how much of the mandate was actually implemented.[9] Mandates reflect the mission's strength, its technical and military capabilities, and the level of international commitment. They are classified as observer missions (in which civilian officials or military officers monitor a truce or treaty); traditional peacekeeping (in which formed military units monitor a truce or treaty); multidimensional peacekeeping (in which a peace treaty authorizes international civilian officials and military units to help build or rebuild political, economic, and social institutions); and enforcement missions, some of which included the transitional administration of the country (in the absence of consent, international military forces intervene to impose peace).

The mere presence of a UN mission might help build peace by signaling to the factions that the international community is watching or that more help is available. A binary variable is therefore used that indicates that a UN mission is present, regardless of its mandate. Alternatively, one could distinguish between facilitative missions, which provide monitoring and reporting (observer and traditional peacekeeping operations), and transformational UN missions (multidimensional, enforcement, and transitional administration), which have a more intrusive mandate and seek to transform rather than merely contain a conflict. One could also distinguish consent-based missions from those authorized under Chapter VI of the UN Charter. Coding UN involvement in these different ways allows a nuanced argument to be developed about the conditions under which the UN can help build self-sustaining peace.

Other controls for international capacities can include a measure of foreign economic assistance. It was not possible to measure the amount of economic assistance available to all countries from all sources (NGO, bilateral, multilateral). Moreover, reconstruction assistance is likely to be an endogenous variable. A control was therefore used for the amount of per capita net current transfers to the balance of payments.[10] A number of other controls were also used in robustness tests (see the Supplemental Appendix).

Results

A logistic regression of the variable participatory peace is estimated two years after the war's end (table 1). All variables have the expected relation with peacebuilding success.[11] Model A controls for the presence of transformational

9. Although coding the mandate is difficult in some cases, the results are robust to recoding ambiguous cases (see the additional results on the author's Web site).

10. This measure is obviously not perfect. The Supplemental Appendix reports results without this variable as well as results obtained using slightly different coding of it (results on the UN are not affected). For the period after 1970, data on effective development assistance as a percent of GDP are used (see the discussion on the author's Web site).

11. Clustering and bootstrapping of the standard errors are dropped in the Supplemental Appendix. This affects the results for some variables (electricity consumption in particular is less robust). UN intervention is measured in three ways, all of which yield significant results (confidence intervals with bootstrapping exclude zero).

TABLE 1. Logit Models of Participatory Peace Two Years after the End of War

Variable	Model A	Model B
Ethnic war	−1.5885*	−1.6075*
(*Wartype*)	(0.5110)	(0.4952)
Log deaths and displaced	−0.3179*	−0.3392*
(*Logcost*)	(0.1370)	(0.1391)
Number of factions	−0.6074*	−0.5686*
(*Factnum*)	(0.2291)	(0.2699)
Net transfers per capita	0.0388*	0.0275*
(*Transpop*)	(0.0118)	(0.0118)
Multidimensional peacekeeping operation and enforcement	3.1039*	
(*TransfUN*)	(1.0290)	
Any UN intervention		1.9247*
(*Unintrvn*)		(0.6118)
Signed peace treaty	1.5799*	1.6153*
(*Treaty*)	(0.6654)	(0.6643)
Electricity consumption	0.0562*	0.0422
Per capita (*Idev1*)	(0.0281)	(0.0282)
Primary commodity	−7.7346*	−7.8967*
Exports/GDP (*Isxp2*)	(2.1829)	(2.2121)
Constant	5.3226*	5.4447*
	(1.5400)	(1.5529)
Pseudo-R^2	0.34	0.32
Log-likelihood	−49.02	−49.99

*Significant at least at the 5 percent level.

Note: Reported figures are coefficients. Figures in parentheses are robust standard errors. Number of observations = 119. Participatory peace is coded as a binary variable: it takes the value 1 if, two years after the war's end or the end of a UN mission, there is no new civil war, no significant residual violence, and no divided sovereignty over the country's territory and the political system has a minimum level of political openness (3 on a scale of 1–20 in the Polity IV dataset). It takes the value 0 otherwise. See table S.2 in the Supplemental Appendix for results using slightly different versions of this variable and for more discussion of the coding.

Source: Author's analysis based on data is described in the text and in the Supplemental Appendix on the author's Web site (http://pantheon.yale.edu/~ns237/index/research.html#Peace).

UN missions only; it compares their effect to no UN intervention and all other types of intervention. Model B controls for any UN intervention, regardless of mandate. The significant positive effects persist when an ordinal measure of peacebuilding is used that distinguishes mixed outcomes from unquestionable successes (see the additional results available in the Supplemental Appendix and on the author's Web site).

The definition of *self-sustaining peace* creates a left-censoring problem in a few cases. Because UN missions are required to leave before they can be coded as peacebuilding successes, in a few cases outcomes are coded more than two years after the end of the armed conflict; in all other cases coding is done at the two-year mark. (This issue is addressed at length in the Supplemental

Appendix, which shows that it does not affect the results.) Only 10 cases are coded with a time lag. One way to get at the fact that in these cases the UN had more time to implement its mandate is to control for the duration of all peace missions. Adding this control to the model and interacting it with UN mandates makes transformational UN missions less significant, although a joint significance test with mission duration overwhelmingly rejects the null hypothesis of no effect ($P = 0.01$).[12]

Using the core model with a categorical variable denoting the type of UN mandate (facilitative or transformational), the probability of peacebuilding success changes when one changes the values of some of the explanatory variables while holding other variables at their means (for continuous variables) or medians (for binary or categorical variables). Moving from a facilitative to a transformational peacekeeping operation increases by 36 percent (the confidence interval ranges from 9 to 55 percent).[13]

These results are robust to extensive specification and robustness tests. Some of the additional controls include the size of the government military (per capita), a measure of the government's capacity to deter external intervention, including by the UN; an indicator variable for Cold War conflicts, a measure of systemic constraints to UN intervention (the UN could intervene with greater ease after the Cold War); region-specific effects;[14] and time trends in the data, captured with a variable denoting the decade during which the war started. Results are also robust to alternative econometric assumptions and estimation methods (as shown in the Supplemental Appendix).

Getting the UN mandate right is critically important. It is not sufficient to send large numbers of troops to the field; troops must be given rules of engagement and a mandate to make peace. The number of peacekeeping troops alone is not a good predictor of peacebuilding success: there is no statistically significant difference between the number of peacekeepers per square kilometer that participate in transformational and facilitative missions.[15] This result suggests that the Security Council often underfunds and underresources transformational missions, which on an average should be given more troops to deal

12. The correlation between transformational UN and interaction with mission duration is 83 percent. Controlling for the duration of all UN missions makes the coefficient of transformational UN significant (P-value is 0.001), as shown in the Supplemental Appendix.

13. This analysis can also be conducted for joint effects, changing more than one variable at once if such changes tend to go together. One could, for example, combine the shift from facilitative to transformational peacekeeping with a change in the treaty variable from 0 to 1, as most transformational peacekeeping missions require a treaty.

14. No regional differences in peacebuilding outcomes were found, although the UN is less likely to intervene in Asia and more likely to intervene in Europe or Africa. Geographic region was thus used as an instrumental variable in a two-stage least squares model that treats UN intervention as an endogenous regressor. UN intervention is still positive and significant in these regressions (see the Supplemental Appendix and the author's Web site).

15. A comparison of the means cannot reject the null hypothesis of no difference ($P = 0.94$). See the results on the author's Web site.

with more-difficult peacebuilding environments. It also suggests that if transformational peacekeeping works better it is not because a larger military force is used.

The effects of peacekeeping troops per square kilometer on the probability of participatory peace success are negative (though non-significant, as reported in the Supplemental Appendix). A large troop deployment with a weak mandate is a sure sign that the Security Council lacks commitment, which creates an impediment for effective intervention. Large numbers of troops per capita in monitoring missions (observer missions and traditional peacekeeping operations) actually reduce the chance of peacebuilding success [examples include Cyprus, Lebanon, and Rwanda, where a large troop deployment (in per capita terms) was given no authority to intervene to stop the violence]. Such deployments are inefficient and potentially counterproductive. A large troop deployment with a narrow mandate in monitoring operations indicates that the Security Council recognizes the severity of the conflict but is unwilling or unable to give troops an adequate mandate to resolve the conflict (that is, a mismatch between the problem and the treatment). Better-targeted mandates should improve the effectiveness of UN missions.

This analysis focuses on participatory peace. Other scholars favor more restrictive definitions of peace. The analysis in the Supplemental Appendix therefore unpacks the complex peacebuilding measure, analyzing each of its components. That analysis shows that the model presented here does not explain other concepts of peace as well as it explains participatory peace. Sovereign peace (which includes all the criteria for peacebuilding success except participation) is more robustly associated with income growth. By contrast, transformational UN missions are more important than income growth for postwar democratization (see table S.11 in the Supplemental Appendix). Different parts of the model are better at explaining different components of peacebuilding.

In general, the resumption of war in the short run is explained by local capacity variables: higher income reduces the risk of a new war, as do less dependence on natural resources and less fractionalization of the society (see tables S.5 and S.6 in the Supplemental Appendix). A more developed and more rapidly growing economy with less dependence on natural resources is not less likely to experience divided sovereignty after civil war, however. Ethnic wars, by contrast, are much more likely to be followed by peacebuilding failure, because of persisting claims over sovereignty. High levels of hostility are particularly damaging with respect to higher-order, positive peace; they are also more likely to lead to persistent divisions in state sovereignty. Treaties are more important for the design of participatory peace; they are generally not significant in ensuring that sovereignty will be undivided or that war or other forms of large-scale violence will recur. UN missions are not very effective in preventing resumption of full-scale war in the short run, but they are helpful in

preventing peace failures that result from persistent divisions in sovereignty, minor armed conflict, or a failure of political institutions.

III. Where Does the United Nations Intervene?

The fact that the UN does not randomly assign its missions should be reassuring to its member countries, but it complicates the evaluation of the effect of UN missions. This section examines the strategies used to account for selection effects in the UN's decision of where and when to intervene.

The first, and more important, strategy is to use knowledge of the institutional structure of the UN to assess the plausibility of arguments that it selects easy cases and to gauge the likely direction of bias in the statistical estimates (if such bias exists). The second strategy is to use estimation methods that can account for selection on observables or unobservables.

Accounting for the Logic of UN Intervention

Underlying the logit regression results is an assumption that all right-hand-side variables are exogenous. Although right-hand-side variables in the model may be correlated with other explanatory variables, they are not caused by peacebuilding outcomes. This is a reasonable assumption, given that peacebuilding outcomes are coded years after right-hand-side variables are coded.

Estimates of the coefficient for the UN variable would also be biased if explanatory variables that are correlated both the UN variable and the dependent variable (participatory peace) are omitted. Although the results are robust to many specification tests, it is not easy to establish that some variables were not omitted.

Critics might argue that some explanatory variables may have been omitted and that there is logic to the UN's decision to intervene that must be modeled before the effects of UN intervention can be assessed. It is possible, for example, that the UN chooses easy cases in order to maximize its chances of success. It is also possible that it intervenes in very difficult situations, where help is needed most. In either case situations in which the UN intervenes may look very different from cases in which it does not.

A variable that may explain both the decision to intervene and the outcome of peace processes is the degree of interest by major powers. This variable is not directly controlled for in the model, but interest by the major powers should be a function of the measures of hostility and local capacities, as well as other included controls. The fact that the UN sends a mission means that the major powers in the Security Council are interested in the case; the degree of interest should be reflected in the mandate (a more intrusive mandate implying greater interest) or in the number of troops sent (more troops implying more interest). Controlling for type of mandate or size of troops does not change the results. The effect of UN missions remains strong when mountainous terrain, political instability, and several other variables that are typically included as

regressors in models of civil war onset are included (see the Supplemental Appendix).

The analysis accounts for possible selection problems in a number of ways. First, a two-stage model is estimated, with a "choice" equation explaining whether or not the UN will intervene, and an outcome equation that is the original peacebuilding model. Such a model is identified only if it includes an instrumental variable that explains UN intervention without also explaining the peacebuilding outcome.[16]

Because such instruments are hard to come by in cross-country studies, some variables with a plausible relation with UN intervention were tested, after establishing that they were not empirically correlated with peacebuilding outcomes. An indicator variable was used for whether or not a country had been a British or French colony, because one might expect that intervention would be more frequent in countries in which permanent members of the Security Council have political, cultural, or other interests. The size of the government military was also used, because one might expect the UN to intervene less often in countries with large militaries (in order to avoid the high costs of a potential military confrontation).[17] A dummy variable for Cold War conflicts was used, because the difficulty of reaching consensus in the Security Council meant that the UN was much less likely to intervene during this period. Dummy variables were also used for geographical regions, to reflect the expectation that, because of the regional interests of the permanent Council members, some regions (Europe or Latin America) are more likely to attract the attention of the Security Council than others (Africa or Asia).

None of these instruments is perfect; various conditions could violate the exclusion restrictions for each of them. Thus although the exogeneity assumption could not be rejected using these instruments and instrumental variables estimation of the model produced substantive results that are consistent with the results of the logit model, the analysis is merely suggestive (see the additional results on the author's Web site). Some of these instruments are significantly correlated with UN intervention, but the R^2 in the first-stage regressions is low, and weak instruments can affect the size of the estimates in the second-stage regressions. Moreover, other complications, not addressed here, could invalidate the instruments used.

Because it was not possible to find good instrumental variables, the problem of possible endogeneity of UN missions has not been fully addressed, making this analysis necessarily tentative. Nevertheless, it is possible to argue that UN missions can be treated as exogenous explanatory variables in the model

16. The covariance between the instrument and the disturbance term in the outcome equation must be zero; the covariance between the instrument and the treatment variable (UN intervention) must be different from zero.

17. The absolute size of the military is more relevant than the per capita measure. In theory the size of the military may also affect peacebuilding, but there is no significant correlation between military size and peacebuilding outcomes.

because of the particular institutional organization of the UN. This is the main claim in support of the exogeneity assumption for UN intervention. The UN's complex decision-making process suggests that there is no simple utility-maximizing logic underlying its decision to send peacekeepers. Bargaining inside the UN is too complex to respond in a straightforward manner to a particular logic of intervention. The interests of the Security Council, the Secretariat, and the General Assembly are rarely aligned in such a way as to produce a unifying logic for the deployment of blue helmets. This explains why there is often a mismatch between underlying conflict conditions and the mandate and resources given to UN peacekeeping missions. In some cases (such as El Salvador and Namibia) there is a good match, and the UN responds in a unified way to the challenges it faces. In other cases (such as Rwanda) the UN mission barely hides the major powers' indifference.

Moreover, once the mandate and resources are determined, there is a high degree of agency slack in the field. A review of cases suggests that the same mandate can be interpreted either as a ceiling or a floor under different conditions, such as different force commanders or different regions within the same country (see the discussion of Cyprus in Doyle and Sambanis 2006). Leadership decisions in the field are not easily anticipated or necessarily connected to any of the important macro-level variables in the model. The actual impact of UN operations has as much to do with how the mandate is implemented as it does with what mandate is given; the agency slack available to field commanders creates an independent institutional effect of UN peacekeeping.

Given that some variables may still be omitted, is it possible to gauge the likely direction of any bias on estimates of the effects of UN intervention? To get a sense of this, one would need to consider the likely effect of the omitted variables on peacebuilding outcomes and the likely sign of the regression coefficient of the same (omitted) variables on UN intervention. Precise calculation depends on the variance–covariance matrix of these variables, which is not known. If both these relations are positive, however, the bias would result in an overestimate of the effects of UN missions; if one or both of the coefficients is negative the bias would underestimate the effects of UN missions.

Omitted variables of interest should help explain why the UN picks the kinds of cases it does. First-stage regressions estimated in the instrumental variables models reveal that the UN intervenes in more-difficult cases. Even casual observation makes this clear: an organization concerned about its success rate would not have intervened in the quagmire of the Democratic Republic of the Congo in the late 1990s: a vast underdeveloped country with deep hostility, no democratic traditions, and many incoherent and uncooperative factions does not suggest an easy transition to self-sustaining peace. As any omitted variables would have to explain why the UN intervenes in difficult cases, the omission of any such variables would likely have a negative overall effect on estimates of the UN's effects (because positive correlates of the probability of UN intervention would be negatively associated with the probability of peacebuilding

success). Consistent with this logic, if any bias in the estimates presented is corrected by the addition of more variables to the model, the effect of UN missions should increase.

Another possible source of bias in the estimates of the effects of UN missions is selection on observables. The non-random assignment of UN missions implies that values of the other covariates may be systematically different in the treatment (cases with UN missions) and control (cases without UN missions) groups. Indeed, in general the UN picks more difficult cases in terms of deaths, other hostility, and local capacity variables. In light of this, the model is reestimated in the Supplemental Appendix, adding interactions between UN intervention and other covariates and estimating the effects of UN intervention by matching on the propensity score and on covariates.[18]

Effect of Selection on the Signing of a Peace Treaty

Peace treaties are usually necessary for certain UN mandates, especially multidimensional ones. Treaties and UN missions work in concert, enhancing the chances for peace (see the results on the author's Web site). Because treaties are sometimes necessary for the UN to send troops, the effects of UN operations on participatory peace are estimated while selecting on *treaty* and *war termination* (see table S.10 in the Supplemental Appendix). Doing so uncovers an interesting new result: some variables that lead the parties to sign a peace treaty make implementation of the peace more difficult (this is the case for longer wars and high levels of deaths and displacements, for example). Consistent with rational choice models of war, a long and bloody war should resolve uncertainty over the parties' relative resolve and capabilities, pushing them to the negotiating table. But once a settlement is reached, implementation becomes more difficult in the presence of such high hostility. It is precisely in such circumstances that UN missions can help, by guiding and reassuring the parties through the peace implementation process.

IV. How Durable Is the Peace?

UN missions appear to have a positive impact in the short term. What are the longer-run effects?

In some cases the peace fails soon after the two-year cutoff point; as with many other types of intervention, one would expect the benefits of UN missions to be concentrated in the short run. In the long run new challenges, which may be unrelated to the previous war or to the peace process, develop which can undermine the peace. Nevertheless, countries that have had some

18. See the Supplemental Appendix and the author's Web site. Matching estimates of the effects of UN peace operations indicate a positive and significant effect on peacebuilding. But matching methods are better suited to analyze datasets with many observations and many variables to explain the assignment of both the treatment and the outcome. The limited applicability of this method to the peacebuilding data is discussed on the author's Web site.

UN assistance during the peace process should have a better chance of achieving self-sustaining peace, largely because of the much-needed help they receive in the critical first years of their transition from war.

Survival analysis can be used to analyze the duration of the peace over the longer term. Survival models estimate the hazard (or risk) of peace failure at time t given that failure has not occurred until then; they can account for right-censoring (the fact that the peace has not failed up to the end of analysis time does not mean that it cannot fail afterward) (for a methodological discussion, see Box-Steffensmeier and Jones 2003). The dependent variable (*peace duration*) is continuous and measured in months, counting from the start of the peace process until the peace fails or the censoring point (end of December 1999) is reached. Peace failure implies that a new civil war erupts. This understanding of peace is a departure from the complex definition of participatory peace used up to this point (see the Supplemental Appendix for a discussion of the duration of participatory peace).

The analysis begins with the same single-record single-failure dataset used earlier; it then switches to a time-series cross-section of the data. Because time-series data do not exist for all the original model variables, the model is somewhat different.

War Recurrence as Peacebuilding Failure

The absence of war is a common standard definition of "peace" in the literature. It is used here mainly because participatory peace cannot be coded as a continuous variable in the same way in which it was coded for short-term outcomes, for several reasons. First, there are no reliable data on levels of residual violence over time. Second, levels of political openness can vary over time. Thus a country's Polity score can fall below the threshold associated with peacebuilding failure in the coding of short-run outcomes, then rise above that threshold a few years later before falling back below it. Third, coding undivided sovereignty requires some sort of deadline: peace processes are often designed to resolve persisting problems of divided sovereignty, and they do so gradually. One therefore has to decide how to treat cases in which sovereignty was divided a few months into the peace process, with gradual improvement leading to a resolution of the problem a year or two into the peace process. Such considerations illustrate the usefulness of using a cutoff point that allows one to ask whether the UN mission helped resolve these issues within a given time frame. Because of the conceptual ambiguity associated with a continuous measure of participatory peace, the analysis of long-term effects of UN missions focuses only on the recurrence of war.

The single-record, single-failure dataset of 138 peace processes includes 73 failures (*peacend*), with mean peace duration of 53 months. The shortest duration is one month, the longest 634 months.

The core model is estimated from Table 1 using a Cox proportional hazards model. Given its versatility the Cox model is a better initial choice than the

more frequently used Weibull model or other parametric hazard models. It gives the hazard rate for the ith individual as $h_i(t) = h_0(t) \exp(\beta'\mathbf{x})$, where h_0 is the baseline hazard rate and $\beta'\mathbf{x}$ are the estimated coefficients and covariates. This model assumes a proportional hazard rate and uses the ordered duration times to derive estimates for the regression covariates. The analysis tests the proportional hazard assumption. Where it is not satisfied another model—such as the Weibull model, which is appropriate if the hazard rate is monotonically increasing or decreasing—is used.

Model 1 is the core model with added controls for real per capita income, the rate of growth of real income at the end of the war, and the level of ethnic fractionalization (table 2). None of these variables has a significant association with participatory peace in the short run. They are added because the longer-run analysis uses a different concept of peacebuilding and the literature links these variables to the onset of civil war.

UN intervention is significant ($P = 0.039$) and reduces the risk of peace failure by about 50 percent (the effect is robust to controlling for non–UN missions). Consistent with much of the literature on the onset of civil war, the strongest result is that local capacities are critical in determining proneness to a new outbreak of war. Local capacity variables take away some of the effect of the hostility variables (only deaths and displacements is significant, and it is not very robust). Countries with higher levels of income, fast-growing postwar economies, and less dependence on natural resource exports are far more likely to experience longer peace durations, despite the negative effects of postwar hostility. The effects of variables such as the number of factions or the nature of the war (whether, for example, it was an ethnic conflict) are non-significant (one might expect any impact they have to be limited to the immediate postwar period).[19] High fractionalization has a significant ($P = 0.007$) negative effect on peace duration.[20]

The results on local capacities are robust to small specification changes. The effects of UN intervention are less robust. One problem is that UN enforcement missions seem to have negative effects on peace duration, probably because of the very challenging circumstances in which they are deployed and the fact that they are designed to end wars rather than build long-lasting peace.

Model 2 drops enforcement missions, controlling only for consent-based UN missions, which are significant ($P = 0.044$) and positive with respect to peace

19. Treaty is dropped from the model on the basis of a likelihood ratio test. The proportional hazards assumption cannot be rejected ($P = 0.13$). Re-estimating the model assuming a Weibull distribution improves the results slightly, as does estimating the model using Cox regression with interactions with the log of time for two variables that come close to failing the proportional hazards test (see the Supplemental Appendix).

20. Models of civil war duration also reveal a significant association between greater fractionalization and longer duration (Fearon 2004), which may be non-linear (Collier, Hoeffler, and Soderbom 2004). These results are in contrast to those of models of war onset. Consistent with the nationalism literature, ethnic differences seem to be "activated" to support mobilization for violence once conflict levels are high (during and after war).

TABLE 2. Duration Models of the Hazard of War Recurrence

Variable	Model 1	Model 2	Model 3	Model 4
Ethnic war	1.14	0.996	0.95	1.12
	(0.31)	(0.27)	(0.29)	(0.31)
Dead and displaced (log)	1.14*	1.12**	1.13	1.20*
	(0.065)	(0.070)	(0.077)	(0.08)
Number of factions	1.04	1.06	1.06	1.01
	(0.099)	(0.08)	(0.085)	(0.07)
Net current transfers	0.999	0.999	0.999	0.999
	(9.18e−07)	(1.13e−06)	(1.22e−06)	(1.01e−06)
Ethnic fractionalization	3.78*	3.81*	3.94*	4.32*
	(1.88)	(1.98)	(2.11)	(2.32)
Electricity consumption	0.999			
	(0.0002)			
Real GDP growth	0.96*	0.96*	0.96*	0.96*
	(0.012)	(0.014)	(0.015)	(0.014)
Real GDP (log)		0.78	0.79**	0.74**
		(0.10)	(0.11)	(0.13)
Primary commodity exports/GDP	3.52*	3.29*	2.92**	2.38
	(1.90)	(2.00)	(1.79)	(1.42)
Any UN intervention	0.54*			
	(0.16)			
UN Chapter VI missions		0.48*	0.47*	0.41*
		(0.17)	(0.18)	(0.16)
Negotiated settlement		0.43*	0.37*	0.33*
		(0.17)	(0.14)	(0.13)
Military outcome		0.54*	0.50*	0.38*
		(0.18)	(0.16)	(0.12)
1940s peace start				0.22**
				(0.197)
1950s peace start				3.37*
				(1.80)
1960s peace start				1.30
				(0.50)
1970s peace start				0.80
				(0.33)
1980s peace start				0.57
				(0.33)
Time dependence (P)			0.62*	
			(0.055)	
Number of observations	129	131	131	131
Number of failures	69	70	70	70
Log pseudo-likelihood	−267.95	−268.91	−182.74	−260.32
Wald χ^2 (degrees of freedom)	73.73 (9 degrees of freedom)	102.88 (11 degrees of freedom)	108.93 (11 degrees of freedom)	154.32 (16 degrees of freedom)

*Significant at least at the 5 percent level.

**Significant at the 5 percent level with one-tailed test.

Note: Reported figures are hazard ratios. Numbers in parentheses are robust standard errors.

Source: Author's analysis based on data described in the text and in the Supplemental Appendix on the author's Web site (http://pantheon.yale.edu/~ns237/index/research.html#Peace).

duration (the odds ratio indicates that consent-based missions reduce the risk of peace failure by about 50 percent). This effect is not reduced when the outcome of the previous war is controlled for. Both negotiated settlements and military victories lead to longer peace durations than do truces or military stalemates. Model 2 replaces the generally non-significant electricity consumption per capita variable with the log of per capita real income, which is positively associated with longer peace durations ($P = 0.054$), consistent with similar results from the literature on the onset of civil war. The proportional hazard assumption is not satisfied.[21]

Model 3, estimated by Weibull regression, yields substantively similar results, suggesting that peace becomes more stable over time. To account for the fact that exposure to the risk of peace failure is higher in countries in which the war ends early relative to the end of the analysis period, the model includes controls for the decade during which the peace process started (model 4). Model 4 marginally satisfies the proportional hazards assumption underlying the Cox model and improves some of the results, making the deaths and displacements variable highly significant ($P = 0.008$).[22]

These results highlight the importance of including economic rehabilitation in a peacebuilding mandate. They dovetail with recent findings in the literature on civil war that demonstrate the power of these variables in influencing the risk of war onset. Civil war scholars use per capita income as a measure of state strength. They show that heavy dependence on oil results in authoritarian state structures.

The results presented here suggest that postwar authoritarianism and state weakness may increase the risk of war recurrence. Consent-based UN missions have a positive impact, but their impact is overshadowed by local capacity variables (though it may be possible that, if the UN presence helps a country return to growth, some of its positive impact is captured by the growth variable). These results suggest that in countries emerging from civil war, a UN war-prevention strategy should be to help build institutions that resist the corrupting pressure of resource-dependent economies and facilitate economic growth. The UN' impact in rebuilding institutions will be particularly important in ethnically divided societies, which are at higher risk of a return to civil war.[23]

21. The model was re-estimated by adding time interactions for those variables that failed the test. The hazard ratio for per capita income drops substantially and is significant at the $P = 0.008$ level. The results on UN consent-based missions are also stronger.

22. The R^2 test of the proportional hazards assumption is only narrowly not rejected ($P = 0.055$). The results are similar using a Weibull model (see the Supplemental Appendix).

23. The Supplemental Appendix presents more robustness tests. A different measure of peace duration is used, based on different assumptions about which cases of war resumption represent war recurrence and which represent new wars. Cases that some scholars might not characterize as civil wars are dropped. The results are generally robust to those changes, although they differ somewhat from the results obtained using the alternative peace duration variable. The logistic regression estimates are robust to all of these coding changes.

Analysis of the Hazard of Peace Failure Using Time-Varying Covariates

The peacebuilding data can be set up as a multiple-event, multiple-failure dataset with both ordered and unordered events (peace failures). The dependent variable (peace duration) is measured in months at peace until peace failure or censoring. Because there are multiple peace failures in a country and the cross-sectional unit is the peace process (*cnumb*), multiple records for each cross-section are added, making clustering on country necessary, as multiple records of country-level variables will be the same across conflicts. The unit of analysis is the peace process year, not the country year, as is typical with binary time-series cross-sectional datasets in international relations and comparative politics. (This set-up is discussed in more detail in the Supplemental Appendix.)

The model is slightly different, but it includes a measure for all three concept variables (hostility, international capacities, and local capacities). A Cox proportional hazards model is estimated with UN intervention as the main variable and lagged log of per capita GDP, lagged oil export-dependence, log of deaths and displacements, ethnic war, signed treaty, level of ethnic fractionalization, and population size as additional covariates.

The analysis concentrates on results pertaining to the presence of any UN mission (a binary variable coded 0 if there was no UN mission and 1 if there was a UN mission, with coding starting at the year of initial deployment and continuing until the peace either fails or analysis time runs out). This allows the effect of UN missions to be studied even after the UN departs by distinguishing cases that received some UN assistance at some point in their peace process from cases that received no UN assistance.[24] A different version of this variable might code a case 1 only after the UN departs; doing so would cause all observations with ongoing missions to be dropped. The difference in the results between these two variables might reveal something about the difference in the effectiveness of UN missions in peacekeeping compared with peacebuilding. Both versions are used in the analysis, lagged one year (table 3).

Lagged per capita GDP (in logs) is very significant ($P = 0.005$) and reduces the hazard of peace failure substantially. By contrast, higher ethnic fractionalization ($P = 0.01$) and more deaths and displacements ($P = 0.03$) significantly increase the risk that the peace will fail (regression 1). Dependence on oil exports, signed treaty, population size, and ethno-religious war do not have significant effects.

UN missions are weakly significant ($P = 0.056$), as in the single-record, single-failure data. They become more significant ($P = 0.025$) if cases in which

24. Coding the presence of UN missions only for years in which they were actually there and peace lasts beyond the departure of UN troops reveals a negative association between UN intervention and self-sustaining peace (by construction there would be no observations of self-sustaining peace coinciding with UN presence).

TABLE 3. Duration Models of the Risk of War Recurrence with Time-Varying Covariates

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Ethnic war	1.391	1.402	1.283	1.284	1.199	1.290	1.221	1.201	1.209
	(0.422)	(0.436)	(0.410)	(0.411)	(0.401)	(0.412)	(0.400)	(0.396)	(0.396)
Ethnic fractionalization	2.973*	3.119*	3.222*	3.109*	3.409*	3.268*	3.228*	3.066*	3.104*
	(1.293)	(1.387)	(1.420)	(1.340)	(1.570)	(1.445)	(1.449)	(1.356)	(1.384)
Dead/displaced (log)	1.130*	1.178*	1.145*	1.148*	1.127*	1.146*	1.130*	1.131*	1.129*
	(0.065)	(0.064)	(0.065)	(0.066)	(0.064)	(0.065)	(0.063)	(0.064)	(0.064)
GDP per capita (log)	0.670*	0.707*	0.707*	0.710*	0.690*	0.705*	0.704*	0.708*	0.702*
	(0.095)	(0.099)	(0.107)	(0.109)	(0.102)	(0.108)	(0.099)	(0.097)	(0.097)
Population (log)	0.909	0.928	0.912	0.909	0.947	0.910	0.934	0.935	0.936
	(0.077)	(0.087)	(0.073)	(0.073)	(0.076)	(0.073)	(0.070)	(0.070)	(0.070)
Oil exports	1.597	1.468	1.456	1.473	1.420	1.475	1.375	1.373	1.393
	(0.543)	(0.434)	(0.546)	(0.566)	(0.457)	(0.562)	(0.434)	(0.423)	(0.437)
Treaty signed	1.411	1.453	1.194	1.241	1.066	1.174	1.180	1.256	1.199
	(0.499)	(0.501)	(0.405)	(0.417)	(0.352)	(0.409)	(0.409)	(0.420)	(0.401)
UN intervention	0.483**	0.446*	0.490**			0.470**	0.349		
	(0.184)	(0.162)	(0.195)			(0.202)	(0.247)		
Chapter VI UN operations				0.418*				0.177**	0.218
				(0.180)				(0.168)	(0.211)
Multidimensional peacekeeping operation					0.449*				
Interacted with peacekeeping operation duration					(0.036)				
Traditional peacekeeping operation					1.001				
Interacted with peacekeeping operation duration					(0.003)				
Time dependence (p)			0.618*	0.620*	0.617*	0.623*	0.629*	0.635*	0.633*
			(0.059)	(0.059)	(0.060)	(0.059)	(0.059)	(0.062)	(0.062)

(Continued)

TABLE 3. Continued

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Wald χ^2 (degrees of freedom)	50.99	46.42	46.15	44.64	167.93	45.48	45.42	39.65	37.94
Log-likelihood	−246.01	−168.99	−171.47	−170.95	−172.13	−171.40	−171.65	−170.64	−171.28

*Significant at least at the 5 percent level.

**Significant at the 5 percent level with one-tailed test.

Note: Figures are hazard ratios. Numbers in parentheses are coefficient robust standard errors. Number of observations = 1,323. Number of subjects = 122. Number of failures = 64.

Source: Author's analysis based on data described in the text and in the Supplemental Appendix on the author's Web site (http://pantheon.yale.edu/~ns237/index/research.html#Peace).

the peace lasted for at least 10 years are dropped in an informal "splitting" model (see the Supplemental Appendix). This result suggests that the effects of UN missions are greater in the first few years of the peace process.[25] This interpretation is consistent with the hypothesis that the UN has become better at peacekeeping over time, as by dropping cases in which the peace lasts more than 10 years, one is more likely to select against cases that occur early in the data.[26] The significance of the UN variable also improves ($P = 0.026$) in a stratified Cox regression (regression 2, where the variable *pstrata* identifies the number of civil wars in the country).

A theoretical argument can be made that the hazard of war recurrence should decrease with time spent at peace if peace (like war) generates its own rewards and incentives for people to keep it going. Thus even though the proportional hazards assumption cannot be rejected, the Weibull model (estimated in regression 3) may be more appropriate. The substantive results remain unchanged. Duration dependence is negatively associated with the recurrence of war, confirming the intuition that the highest risk of recurrence occurs during the first few years following the end of civil war. Given that UN peace missions have their greatest effect in the short run; this suggests that their long-term effect on the durability of the peace reflects successful peace implementation during the first few years of transition, which helps set the stage for self-sustaining peace.

Trying to get at the impact of different types of UN mandates is difficult. Multidimensional mandates cannot be included in the regression, because they predict peacebuilding success perfectly. Although earlier results from the survival analysis suggest that enforcement missions increase the risk of a new war, there have been so few missions that the results would not be robust (and there may well be a selection effect, as the war is often not over when enforcement missions are sent to the field). For this reason enforcement is dropped and only consent-based operations are considered in regression 4. The results suggest that they have a significant effect ($P = 0.04$) in reducing the risk of a new war.[27]

Do longer-lasting UN interventions reduce the risk of peace failure? This question is investigated by creating an interaction between the type of mandate and the duration of UN missions. Regression 5 indicates that longer-lasting multidimensional missions do have such an effect. This result is consistent with the fact that rebuilding institutions and enhancing local capacities takes time, suggesting that a quick exit is not likely to be a good strategy on an average. By contrast, longer-lasting traditional peacekeeping does not make a difference.

25. This procedure also addresses the problem of uneven exposures to failure risk for peace processes that started earlier than others.

26. The author thanks an anonymous reviewer for making this point.

27. This result is robust to adding year dummies and controls for the geographical region.

This result is consistent with the fact that such missions do not transform a conflict and may well be merely monitoring a political stalemate.[28]

Several time lags of the UN intervention variable are used, as well as a version of this variable that takes the value of 1 only after the UN departs. The results of a Weibull regression (regression 6) become weakly significant after the first lag but non-significant after the second lag. It may not matter how a sustainable peace is achieved, but achieving it, whether by domestic or international means, does make a difference in the long run. If one controls for UN intervention only after the UN has left (regression 7), there is no effect (although results are close to statistical significance with other methods[29]), because all mandates, including enforcement, are lumped together in this variable. If one looks separately at consent-based peace operations, more-lasting effects ($P = 0.068$) are found, even after these missions depart (regression 8), although the effects become non-significant after the second lag (regression 9). Overall, these results suggest that the UN's positive contribution with respect to the avoidance of war recurrence is concentrated in the short term, primarily to the period in which the UN forces remain on the ground.

V. Conclusion

UN missions have a robust positive effect on peacebuilding outcomes, particularly participatory peace, but the effects occur mainly in the short run and are stronger when peacekeepers remain. This finding is reassuring, because the longer a peace lasts, the more stable it becomes. Interventions that shore up the peace in the immediate postwar period can thus have a lasting effect, as the best predictor of participation and peace tomorrow is participation and peace today.

Economic factors drive the long-term prospects of peace in the average post–civil war country. This finding is consistent with the prevailing opinion in the quantitative literature on the causes of civil war (see, for example, Fearon and Laitin 2003; Collier and Hoeffler 2004). Long-term peace can be facilitated by the presence of a UN mission with a mandate to monitor and police cease fires and rebuild institutions. Simply sending large numbers of troops will not solve the problem; they need to have a well-defined mandate. Economic assistance alone is also unlikely to do the trick, although more evidence is necessary to fully assess this claim. While economic factors may be more significant than other factors in preventing the resumption of war, more than mere financial assistance is needed to rebuild political institutions and implement complex peace agreements in the early years of postwar transitions.

28. In light of the fact that there have been too few missions of different mandates, this analysis is not pushed farther. Interested readers can read extensive case studies in Doyle and Sambanis (2006).

29. Results from random effects probit model are presented in the Supplemental Appendix.

The value added of UN missions lies precisely in its ability to play a vital role in transitions from civil war.

Countries that have had a UN mission tend to recover in a more sustainable manner from civil wars than countries that have not had a UN mission, but the effect wanes over time. Success in the early postwar years may account for the return to long-run growth, so that some of the impact of UN missions may well be reflected in the economic variables. Over time economic development is critically important; UN missions must be designed so that they can make a significant difference with respect to this critical determinant of peace. As good as UN peacebuilding is in expanding political participation, it must also help to jump-start self-sustaining economic growth. Economic reconstruction is usually treated as follow-on to peace rather than the vital partner that it is. It provides the visible benefits of peace that help mollify hostilities, the civilian jobs that absorb demobilized soldiers, and the tax revenues that strengthen state capacity. Economic reconstruction also funds the national army, which sustains public order. It pays for soldiers who demine rather than mine battlefields.

In addition to being a significant determinant of sustainable peace, growth is also a determinant of sustainable democracy (Przeworski and others 2000; Collier and Hoeffler 2004). Efforts are under way to fill the policy gap between peacekeeping (which focuses on humanitarian assistance) and development assistance (which emphasizes structural transformation). One example of such efforts is the creation, in 2006, of the UN Peacebuilding Commission. This commission should coordinate UN peacekeeping interventions with development assistance provided by international financial institutions, such as the World Bank.

Much remains to be learned about recovery from civil war: no recipes are available for the right mix or sequencing of security and development strategies following such conflicts. The international community would benefit from an evolution that uses economic reform to plug the gap between peacekeeping and humanitarian assistance on the one hand and development on the other.

Supplementary Material

Supplementary Material is available at The World Bank Economic Review Online.

Funding

Financial support is from the World Bank's Project on Political Institutions, Development, and Domestic Civil Peace.

References

Annan, Kofi. 2001. "No Exit without Strategy: Security Council Decision-Making and the Closure or Transaction of United Nations Peacekeeping Operations." Report presented to the Security Council April 20. Document S/2001/394. United Nations, New York.

Boulding, Kenneth. 1964. "Toward a Theory of Peace." InRoger Fisher, ed., *International Conflict and Behavioral Science*, 70–87. New York: Basic Books.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2003. *Timing and Political Change: Event History Modeling in Political Science*. Ann Arbor: University of Michigan Press.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–95.

Collier, Paul, Anke Hoeffler, and Mans Soderbom. 2004. "On the Duration of Civil War." *Journal of Peace Research* 41(3):253–74.

Doyle, Michael W., and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *American Political Science Review* 94(4):779–801.

———. 2006. *Making War and Building Peace: United Nations Peace Operations*. Princeton, N.J.: Princeton University Press.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer than Others?" *Journal of Peace Research* 41(3):275–301.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review* 97(1):75–90.

Przeworski, Adam, Michael Alvarez Fernando Limongi, and Jose Cheibub. 2000. *Democracy and Development*. Cambridge: Cambridge University Press.

Sambanis, Nicholas. 2004. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6):814–59.

Shimizu, Hirofumi, and Todd Sandler. 2002. "Peacekeeping and Burden-Sharing, 1994–2000." *Journal of Peace Research* 39(6):651–68.

Insurgency and Credible Commitment in Autocracies and Democracies

Philip Keefer

The inability of political actors to make credible promises to broad segments of society—a previously unexplored determinant of civil war—causes both elected and unelected governments to pursue public policies that leave citizens worse off and more prone to revolt. Noncredible political actors are also less able to build counterinsurgency capacity. Popular dissatisfaction with rulers reduces the costs to counterinsurgents of overthrowing regimes, discouraging rulers from building counterinsurgency capacity in the first place; lack of credibility prevents rulers from writing contracts with counterinsurgents that maximize counterinsurgency effort. Empirical tests across numerous subsamples using various measures of political credibility support the conclusion that broad political credibility ranks at least as high as social fractionalization and natural resource rents as a cause of conflict. JEL Codes: D73, D74

The literature on the determinants of civil war has explored three reasons why normal politics give way to armed conflict: resource endowments raise the stakes of violent conflict and allow its financing; governments are unable to organize a response to rebellion, because of, among other reasons, country characteristics (such as mountains) that facilitate rebel activity; and social conditions, particularly ethnic fragmentation, increase popular support for insurgency or at least acquiescence to it. This article argues for a fourth possible explanation: in conflict-prone countries, "normal politics" are undermined by the inability of politicians to make credible promises to large segments of the citizenry.

The lack of political credibility increases the risk of conflict in two ways. First, in both democratic and nondemocratic countries weakly credible leaders have incentives to make policies only in the interests of the few groups that believe their promises. In this setting public goods are underprovided, and

Philip Keefer is a Lead Research Economist in the Development Research Group at the World Bank; his email address is pkeefer@worldbank.org. The author gratefully acknowledges the assistance of Cristina Abodea, as well as comments from participants in seminars at the International Peace Research Institute in Oslo, Stanford University, the University of Lund, and the University of Basel.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 33–61 doi:10.1093/wber/lhm019
Advance Access Publication January 5, 2008

private good provision and corruption swell. In this distorted policy environment citizens are less likely to resist efforts, including efforts by potential insurgents, to replace the incumbent.

Second, noncredible leaders are less able than credible leaders to build counterinsurgency capacity. On the one hand, capable counterinsurgents always have the choice of overthrowing their employer. Like insurgents, their costs of doing so are lower when the government is unpopular. Noncredible governments must therefore pay more to dissuade counterinsurgents from undertaking coups. On the other hand, optimal compensation contracts for counterinsurgents entail future payments contingent on success. However, leaders who are noncredible to citizens broadly are less able to make the credible commitments to counterinsurgents that are essential to these contracts. Reduced counterinsurgency capacity and citizen support make noncredible governments more vulnerable to insurgency.

Sections I and II examine how political credibility reduces popular support for insurgencies in democracies and nondemocracies and increases the government's ability to mount a counterinsurgency capacity. The rest of the article tests the proposition that conflict is more likely where politicians are less credible, using three different proxies for political credibility. Two reflect political party characteristics: whether parties are programmatic or institutionalized. The third is continuous years of competitive elections. All significantly affect the risk of conflict. The evidence suggests that credibility differences distinguish conflict from nonconflict countries at least as definitively as determinants of conflict ranging from oil to ethnic fractionalization.

I. Political Credibility and Popular Support for Insurgency

The literature points to a significant role for credibility in starting and ending civil wars. Garfinkel and Skaperdas (2000) argue that political competitors resist peaceful redistribution because they cannot credibly refrain from using the benefits of redistribution to arm themselves in the future. Fearon (2004) explains that some types of interstate conflict are difficult to end because the prospect of significant and exogenous variation in the relative strength of contending states makes self-enforcing peace agreements difficult to create. In fact, as Walter (1997) reports, only 20 percent of civil wars have ended by agreement, and most of them relied on third-party enforcement.

This article focuses on a different aspect of credibility, one that is common in discussions of political competition more generally but less common in the literature on conflict: whether leaders can make credible promises to large numbers of citizens. Lack of credibility, by inducing governments to pursue policies that give privileged access to government resources to some while doing little for the vast majority of citizens, reduces citizen support for incumbents and encourages insurgency. While the conclusion is the same in both

democracies and nondemocracies, the underlying logic differs somewhat. The two cases are therefore treated separately.

Credibility and Popular Support for Insurgencies in Democracies

In democracies, if challengers cannot make credible promises, citizens have no reason to believe that they will enact different policies or perform better than poorly performing incumbents. Knowing this, incumbents ignore citizen welfare and throw themselves into rent-seeking. Ferejohn (1986) notes, however, that in real-world democracies, even where credibility appears to be absent, citizens are nevertheless able to hold politicians accountable. He attributes this to voter coordination on a performance threshold: if incumbents meet this threshold, they are re-elected; if they do not, they are thrown out of office. This type of postelectoral accountability mechanism exerts a relatively weak influence on government performance. Persson and Tabellini (2000) extend this logic to government decisions regarding the provision of private and public goods and rent-seeking.

Keefer and Vlaicu (forthcoming) relax the assumption that credibility is an exogenous characteristic of countries and allow politicians either to make promises with the assistance of patrons (who in turn appeal to their clients) or, at some cost, to make promises directly to some fraction of society. When the costs of appealing directly to some fraction of society are sufficiently high relative to the costs of employing patrons, credible promises can be made to only a few. This condition yields predictions consistent with actual policy choices in less credible states: policymaking is targeted to those few who believe political promises, leading to high levels of targeted or private spending, lower public goods provision, and greater rent-seeking by politicians. As welfare falls or inequity in the provision of public services increases, the threat of insurgency rises correspondingly.

For several reasons, politicians are less likely to be credible in younger than older democracies. Keefer (2007) finds broad evidence that the policy choices of young democracies exhibit precisely this policy pattern. This evidence is robust to numerous other explanations, ranging from per capita income to ethnic fragmentation to the formal political institutions that young democracies adopt. The number of years of competitive elections is therefore used as one of the proxies for political credibility in the analysis below.

Political parties are key to credibility in democracies. For promises to be credible candidates must bear a loss if they renege on them. That loss could be their personal reputation. However, if they come from a programmatic political party (one with a reputation for pursuing particular policies), the reputational costs of reneging affect not only the candidate but also all party members. Any single candidate who reneges on the policy promises of the party undermines the party's reputation, so that in future elections the party's label (now no longer credible) will contribute nothing to a candidate's electoral support. Party members therefore have strong incentives to police one another.

Programmatic parties are also better able to implement their promises, because they can enforce a higher degree of party discipline: party members lose the electoral benefits of the party label when they are expelled. Keefer (2005) presents evidence that where programmatic parties are absent, public goods provision is lower and private good provision and corruption greater. Two proxies for the extent to which parties are programmatic or the extent to which party members can discipline party leaders are used in the empirical analysis presented below.

Credibility and Popular Support for Insurgencies in Nondemocracies

Credibility is also a significant determinant of regime support in nondemocracies, where most civil wars occur (56 out of 71 since 1975, the period analyzed below). For example, clientelist strategies—the targeting of benefits to particular supporters at the expense of citizens at large—are a staple of decisionmaking in poorly performing nondemocracies (see, for example, Lewis 1998). In addition, like democracies, autocracies vary in their ability to make credible promises to large numbers of citizens. In many circumstances, as Haber, Razo, and Maurer (2003) argue in the case of the Mexican autocrat Porfirio Díaz, autocrats rely on personal relationships or family ties to make credible commitments. Their ability to make credible commitments only to a narrow slice of society correspondingly reduces the quality of public policies that benefit the vast majority, increasing incentives to mount insurgencies (such as the Mexican Revolution of 1910).

Autocrats can also use political parties to broaden their credibility. As in democracies, one strategy is to construct a ruling party with a strong ideological identity, one reflected in autocrat policies. Communist parties are prototypical examples of such a strategy, in which insurgency is deterred to the extent that insurgents cannot make similarly credible policy promises.

Another strategy is to create ruling party institutions that raise the costs to the autocrat of reneging on promises to the supporters of the ruling party. Ruling party institutionalization can take many forms. For example, autocrats can share power within the ruling party, creating intraparty political checks and balances that reduce the scope for arbitrary treatment of party members by leaders. Following the Glorious Revolution of 1688, for example, the British Crown acceded to institutional arrangements (a strengthened Parliament) that tied its hands; Parliament represented a tiny fraction of the British population but a considerably larger fraction of its moneyed citizens (North and Weingast 1989). Investment opportunities for the members of Parliament and employment opportunities for the rest of society opened up, reducing the risk of revolt.

Gehlbach and Keefer (2006) analyze a second type of ruling party institutionalization, the decision of leaders to allow more supporters into the ruling party, where they can exchange information and coordinate more freely (to overthrow the leader, for example) than nonparty members (in response to

leader expropriation, for example). When the ruling party is larger, welfare is higher, both because party members are more numerous and richer and because they employ more nonparty members. Insurgency is correspondingly less likely.

China illustrates the key role of ruling party institutionalization. It moved from a system in which the top leadership's capacity for arbitrary decisionmaking (under Mao) was nearly unfettered to one in which a number of intraparty institutional arrangements place checks on arbitrary leader behavior. These include rules governing intraleadership decisionmaking (checks and balances at the top of the party) and expensive and elaborate personnel evaluation schemes that may have facilitated coordination among party members. Such institutional arrangements increased the credibility of leadership promises to millions of Communist Party cadres, who responded to economic liberalization with massive investments, convinced by the new intraparty institutional arrangements that Beijing would not expropriate the proceeds of those investments (Keefer 2006).

Despite their usefulness, institutionalized ruling parties are uncommon in nondemocracies. Institutionalization require leaders to surrender authority and, therefore, rents. The autocratic leader who decides to share power with a junta of nine other autocrats could see his rent share drop 90 percent. In countries with large natural resource deposits the rents from faster growth, reduced insurgency threats, and enhanced counterinsurgency capacity may be lower than the rents lost from sharing. The empirical analysis below uses the two political party variables to examine whether there is an association between insurgency and the institutionalization of ruling parties in autocracies.

II. Political Credibility and the Construction of Counterinsurgency Capacity

Lack of credibility increases the probability of insurgency by raising the costs of building counterinsurgency capacity, in two ways. First, the low popularity of noncredible rulers not only reduces the likelihood of popular resistance to insurgency, but it also reduces the likelihood of resistance to coups by counterinsurgents. Second, while rulers would prefer to pay counterinsurgents contingent on the success of their efforts, they cannot do so if they are not credible.

With respect to the first issue, citizens, who are less happy under noncredible leaders are also less likely to oppose coups by counterinsurgents against noncredible leaders. In response, rulers can raise the share of rents they offer to counterinsurgents, reducing counterinsurgency incentives to overthrow the ruler. This is an expensive solution: the rent share necessary to make it succeed may far exceed the effort counterinsurgents must invest to put down rebellion. In fact, military spending is about 2 percentage points of GDP

higher in nondemocracies that lack a programmatic ruling party than in those that have one.[1] These costs are a substantial disincentive to forming a large and capable counterinsurgency force, encouraging insurgency.

Rulers can also organize counterinsurgency forces in a way that makes it more difficult for the forces to launch coups. They can increase the coordination costs of counterinsurgents by creating separate and competing counterinsurgency units, rotating unit leaders frequently, or putting family members in command of the units. Haber (2006) makes a similar argument in the context of dictators who would like to prevent their supporters from overthrowing them, but would still like to make credible promises to them. Unfortunately, this strategy weakens counterinsurgency capacity, because uncoordinated security forces are less able to fend off insurgencies, and uncles and brothers may be incompetent.

Credible rulers are also better able to make credible offers of future compensation to counterinsurgents. The link between the credibility of ruler promises to citizens broadly and to counterinsurgents specifically is straightforward. If credible rulers fail to compensate counterinsurgency forces as agreed, public security declines, as security forces react to nonpayment by quitting, striking, or using their arms against citizens. At the same time, the tenure in office of broadly credible politicians is more likely to depend on their ability to provide the public good of law and order. They therefore lose credibility with citizens and shorten their tenure in office when they fail to honor commitments to security forces, increasing the costs to them of reneging on agreements with security forces.

The inability to make credible promises of future compensation undermines counterinsurgency capacity because leaders cannot easily observe counterinsurgency effort. They would therefore prefer to offer success-based contracts to counterinsurgents, with some payments contingent on defeating the insurgency. If counterinsurgents believe rulers will renege on compensation promises should they succeed, however, they will demand payment upfront, weakening their incentive to exert effort against insurgents. Broadly noncredible rulers can build their counterinsurgency capacity around individuals who believe their promises (for example, family members). These players may not be the most capable, however, making the counterinsurgency less effective than it otherwise would be.

In addition, the current costs of fighting an insurgency may exceed the current ruler's resources, including those from borrowing, so that a successful counterinsurgency effort can be mounted only if promises of future compensation are possible. Assume that governments cannot borrow and that with a total effort by counterinsurgents of E (expressed in units of rents), an insurgency can be defeated, extending the expected duration of the government to

1. It is possible that these countries confront higher risks of international military conflict; of course, that risk may itself be endogenous to the political dynamics inside the country.

T years. The expected value of rents that the government receives from each additional year of tenure is given by r; total rents are given by $R = rT$. If ruler promises are credible, counterinsurgents will accept future rewards for current effort and the government can expend up to R to defeat the insurgency, prevailing if $R > E$. If these assurances are not credible, the government can defeat the insurgency only if $r > E$.[2]

III. Country Examples

Examples of specific conflicts support the argument that the lack of political credibility makes conflict more likely. The insurgencies reviewed here all occurred when governments could not make broadly credible promises and pursued policies consistent with their limited credibility. In cases in which civil wars ended with peace accords, as opposed to total victory, a key attribute of the accords was the enforceability of policies that broadly credible politicians would, in any case, have had incentives to pursue (such as broadly redistributive policies). The cases here also underline the close connection between the lack of broad political credibility and ethnically driven conflict.

Guatemala

The redistributive demands of rebels during the Guatemalan civil war were precisely those that credible governments in poor democracies with significant income inequality would have been driven to provide naturally, even in the absence of rebellion. These policies did not emerge, however, because pre-civil war governments could not make broadly credible promises. In 1995, the year before the final peace accord was signed, taxes in Guatemala amounted to only 8 percent of GDP, the lowest in Latin America. From 1975 to 1995, according to party data in the Database of Political Institutions (Beck and others 2001), Guatemala had no broadly credible left-wing party. Rebellion driven by redistributive demands was therefore more likely.

The Guatemalan case also illustrates the importance of the reputational effects of continuous competitive elections. Negotiations to end more than 30 years of conflict began when the first democratic government of Vinicio Cerezo took office in early 1986. They resumed in 1991, under the (elected) government of Jorge Serrano; continued under his elected successor, Ramiro de Leon; and concluded under the presidency of Alvaro Arzu in 1996 with the signing of the last 7 of the 12 accords that constituted the agreement to end the conflict (Rodas-Martini 2007).

One explanation for the extended process was insurgent doubt over whether government commitments reached in the peace accords were credible. Insurgents would have been more reluctant to trust elected governments in new

2. Rulers also compensate counterinsurgents extracontractually, by providing them opportunities for looting or drug-dealing, for example. These strategies effectively reduce R, however.

democracies, because they were unlikely to be broadly credible. This in turn meant that they had weaker political incentives to engage in broad-based redistribution. Agreement was reached when it was more credible to guerrillas that governments saw it in their own interests to increase taxes and spending. This in turn was more likely in a democracy that had seen 10 years of continuous competitive elections and several changes of president than it had been earlier, when democratic institutions were newer or nonexistent.[3]

Uganda

Uganda also suffered from a decades-long internal conflict, one with a strong ethnic character. Most observers blame ethnic conflict on the governing practices of the British during the colonial period (Ndikumana and Nannyonjo 2007). The British strongly favored the Baganda, in the south, at the expense of other groups, especially those in the north. When the Baganda became restive, the British began to arm northern groups, establishing a rough balance between the economic power of the Baganda in the south and the military power of the Acholi and Langi in the north. These practices meant that, postindependence, leaders from the Acholi or Langi could not easily make credible promises to the Baganda and instead focused on their own ethnic groups at the expense of others. The lack of cross-ethnic political credibility, combined with the imbalance of military forces, made conflict more likely.

The first postindependence regime of Milton Obote, a Langi, was rooted in the northern ethnic communities of the Acholis and Langis. Obote immediately began to discriminate against the Buganda region, to increase northern domination of the military, and to cement his regime with repression (Ndikumana and Nannyonjo 2007). In 1971 Idi Amin, from the Kakwa community, overthrew Obote and targeted not only the Baganda but also the Acholi and Langi, as well as foreign-owned businesses, whose assets he redistributed among his closest supporters. Yet another group overthrew Amin in 1979; it was replaced in 1980 by Obote, who pursued repressive policies that further widened the gulf between the Langi and other ethnic groups, including the Acholi. An Acholi replaced Obote in 1985, but in 1986 Yoweri Museveni, with support rooted in the Buganda and some western communities, defeated Obote's army. Immediately, though, conflict resumed in the north between Acholi-dominated groups and the Langi and between the Acholi groups and the Museveni government (which has included the president's brother as defense minister).

The usual lesson drawn from the Uganda experience is that ethnic conflict emerges when armed ethnic groups cannot make credible promises not to attack one another. The argument here suggests two further lessons. First,

3. By 2004 the Guatemalan government had not honored the letter of the peace accords (calling for an increase in the tax ratio to 12 percent of GDP). However, after oscillating between about 6.5 and 8.0 percent of GDP from 1990 to 1995, the gross tax burden rose 35 percent from 1995 to 2004, rising continuously from 8.0 percent to 10.8 percent of GDP.

ethnic conflict is also more likely when ethnic groups simply cannot make credible promises to make policy decisions that are broadly beneficial. Second, conflict is more likely when political leaders cannot make broadly credible promises even within their own ethnic groups. In Uganda political incentives were to provide targeted payoffs to narrow segments of political leaders' ethnic constituencies, paid for by predation on other constituencies. Leaders could still retain the support of their ethnic constituency since even the members of the constituency who did not benefit—the majority—could see that their own leaders were their only defense against predation by other groups. Unfortunately, the net result of such a political equilibrium is enhanced risk of conflict.

Lebanon

The fragility of interethnic agreements and the particularistic nature of intraethnic political promises also characterize conflict in Lebanon (Makdisi and Sadaka 2005). Until 1975 Lebanon had been governed by a coalition of Maronite Christians, Shia Muslims, and Sunni Muslims. This equilibrium was disrupted by the emergence of Palestinian organizations and disaffected, nonsectarian Muslims. They undermined the balance of power among the sectarian groups that had allowed the coalition arrangements to be self-enforcing, precipitating violence.

Why was the equilibrium sufficiently vulnerable to shocks as to precipitate conflict? One answer appears to be the large mass of disaffected Muslims. Their dissatisfaction was a natural consequence of sectarian leaders who relied for support not on the promises of broad policies and public goods that would improve the welfare of all members of their respective groups but rather on promises of particularistic policy benefits to some members of their groups. Appointments to most, if not all, public administration positions were subject to sectarian considerations, with party leaders having a large say in which members of their sectarian group would receive jobs. Parliamentary seats were also distributed to communities according to sectarian formulas. Not surprisingly, there was little political support for the public good of a well-performing public administration, which was instead characterized by corruption and laxity (Makdisi and Sadaka 2005).

Macedonia

In Macedonia, as in Lebanon, politics was organized around ethnic appeals and political parties that depended on particularistic policies and charismatic leaders (Lund 2005). Why did the lack of political credibility not precipitate civil war? Lund suggests two reasons, both absent in Uganda and Lebanon. First, multiple parties competed for support within each major ethnic group. This increased political pressures on the parties not to concentrate the benefits of clientelism on narrow groups within their ethnic communities, which reduces popular dissatisfaction with incumbent governments. Second, a broad

international presence provided large rents that were contingent on continued peace. "The key to obtaining political support was patronage, the key to patronage was international assistance, the key to assistance was international recognition, and the key to recognition was responsiveness to international norms" (Lund 2005, p. 237). Such assistance, which was relatively minor in Lebanon and was not provided to Uganda, reduced the relative payoffs for any ethnic group of capturing the Macedonian state through conflict.

IV. Specification and Data

More systematic evidence also points to the importance of credibility. The proposition examined here is that the lack of broad political credibility increases the risk of insurgency, whether because it reduces popular resistance to insurgency or because it makes the development of effective counterinsurgency capacity more difficult. More continuous years of competitive elections (in democracies) and institutionalized or programmatic political parties (in both democracies and autocracies) endow rulers with credibility and reduce the risk of insurgency.

To test these propositions, the empirical work presented below follows the literature (for example, Fearon and Laitin 2003), estimating the following logistic specification:

$$(\text{Onset of Conflict})_{it} = f(\text{political variables, income/capita, ln(population), ln(mountainous), oil, ethnic fractionalization, religious fractionalization, and the noncontiguity of state territory})_{i,t-1} + \varepsilon_{it}.$$

The conflict data come from the correlates of war data base (Sambanis 2004) and, like the nonpolitical control variables, are used in Fearon and Laitin (2003). A country is coded as experiencing a civil war when three conditions are met: the conflict caused more than 1,000 deaths, it threatened the sovereignty of the state, and rebel opposition to the state was militarily organized. The control variables are taken from Fearon and Laitin because they went to great lengths to supplement the usual sources of data, which are often missing in poor, conflict-torn countries. Population and geographical variables capture the ease of mounting an insurgency; oil is a dummy variable (0–1) constructed by Fearon and Laitin indicating whether a country is a significant oil exporter; and fractionalization variables proxy for social polarization.[4] To capture geographic features of countries that affect the cost of insurgency, Fearon and Laitin control for how mountainous countries are (the underlying index ranges from 0 to 94.3, where Bahrain is among the countries at 0 and Bhutan is 94.3)

4. Fractionalization variables also enter in other ways in the literature to better capture polarization, as defined axiomatically by Esteban and Ray (1994): quadratically, in Keefer and Knack (2002), or transformed according to a formula that links fractionalization more precisely to their formalization of polarization, in Reynal-Querol (2002). The focus here, however, is on the political variables; the linear specification was therefore retained.

and include a dummy variable indicating whether territory is contiguous (24 countries are noncontiguous).

Fearon and Laitin constructed the per capita income variable based on Penn World Tables 5, 6 (covering the years 1950–92) data on real per capita income (chain method), measured in 1985 U.S. dollars. They extrapolated from 1950–99 using per capita income growth rates from *World Development Indicators 2001*. For countries for which they had at least 11 observations, they added additional observations based on regressions of income on year and energy consumption. For the period 1975–99, their data set includes income data on many more poor countries than the *World Development Indicators* does.[5] This allows 13 more conflicts (out of a total of 73) to be analyzed than is possible using *World Development Indicators* data.

The three political measures used in the analysis to capture various notions of credibility are new to the conflict literature. The first, following Keefer (2007), is the number of years of continuous competitive elections. Consistent with the assumption that these are associated with broad political credibility, the lower is this variable, the lower is public goods provision (such as secondary school enrollment, bureaucratic quality, and public information flows) and the greater are corruption and private good provision (government jobs and public investment). The elections variable comes from the Database of Political Institutions (DPI) (Beck and others 2001) for the period 1975–2004. This variable is zero in countries in which leaders were not elected in competitive elections (elections in which multiple candidates compete in legislative and executive elections and none receives more than 75 percent of the vote, as judged by the Legislative and Executive Index of Electoral Competitiveness from the DPI) and one otherwise.

Two political party variables are also used, both derived from the DPI. One is a dichotomous variable indicating whether the main governing party is programmatic. The DPI indicates whether parties can be judged as right, left, or center. If they can be, this is an indication that they can make more programmatic appeals to citizens. The DPI assessment of this variable is based on sparse information. However, it is highly correlated with assessments in detailed regional studies of whether parties are programmatic (Kitschelt and others 1999 in Eastern Europe; Jones 2005 in Latin America).

The other party variable is the number of years by which the age of the political party of the country's leader exceeds the years that the leader has been in power. When this number is low or negative, it is more likely that the party is constructed around the personality of the leader. In this case, it is only the reputation of the leader that is hurt when party members fail to pursue the leader's program. When, however, the number is positive and high, the party is more

5. For 1976 Fearon and Laitin (2003) have data on 32 countries that are missing in the *World Development Indicators 2007* (World Bank 2007). The average income of these countries is about 40 percent less than the average income of the countries for which data from both sources are available.

likely to have an existence independent of the leader, and party members have greater incentives to develop and preserve the party's programmatic reputation. The age variable is also likely to be associated with the extent to which rulers confront intraparty checks and balances, which is smaller when the party is younger than the ruler's tenure.

These political variables offer more nuances to the description of pre–civil war politics than is typically the case in the conflict literature. The literature generally characterizes countries using the degree of democratization, as measured broadly by the 20-point Polity democracy-autocracy measure (Jaggers and Gurr 1995). Although Polity covers more years than any other political dataset, it does not include data on political parties, an essential element of the analysis here. In addition, the Polity democracy-autocracy measure is more difficult to interpret than the DPI variables. The Polity democracy measure is a composite of other Polity variables, each assigned a weight and then summed to yield the final democracy score. By construction, similar democracy measures over time can be associated with significantly different configurations of the component indicators, as long as they add up to the same final number. Similarly, quite different models for constraining executives or electing leaders could drive similar values of the component variables that make up the democracy measure.[6]

Political credibility is significantly related to many of the control variables that are common in the conflict literature. Political strategies to make clientelist rather than broad programmatic promises, for example, are more likely in poor countries; ethnic appeals are more likely when politicians are not broadly credible.[7] Inclusion of such controls as income and social cleavage therefore biases the estimations presented below against finding a significant role for political variables. This issue is discussed more extensively below.

V. Does Political Credibility Reduce the Risk of Conflict?

The proposition that political credibility reduces the probability of civil war is tested in four samples of countries: all countries, all poor countries, all nondemocracies, and all poor nondemocracies. All of the political variables are significant in each of the three specifications in the sample of all countries

6. At the same time, some values of the Polity measure are, by construction, endogenous to insurgency. For example, Hegre and others (2001) and other researchers find that countries with intermediate values of the Polity measure are most vulnerable to civil war. Their operationalization of the Polity measure is common in the literature. However, countries can fall into the intermediate category, for several reasons. One is the violent overthrow of the regime, which leads Polity evaluators to downgrade the country with respect to the institutionalization of political competition (PARREG), one of the elements of the Polity democracy index.

7. This is the case, for example, to the extent that members of the same ethnic group believe that candidates' individual welfare is tied to the welfare of their ethnic group or that candidates have had more repeated interactions with members of their ethnic group than with others.

TABLE 1. Credibility Effects in the Whole Sample (Dependent Variable: Conflict Onset, 0–1)

Variable	(1)	(2)	(3)
Age of largest party minus leader years in office	0.987 (0.04)		
Years of continuous competitive elections		0.959 (0.02)	
Largest government party programmatic? (1–0)			0.507 (0.05)
Ln (population)	1.186 (0.06)	1.145 (0.12)	1.175 (0.07)
Oil exporter dummy variable (0–1)	1.797 (0.04)	1.466 (0.15)	1.387 (0.31)
Ethnic fractionalization	3.036 (0.05)	2.890 (0.07)	3.486 (0.03)
Religious fractionalization	4.174 (0.05)	4.074 (0.04)	3.301 (0.06)
N (country years)	2,603	2,822	2,847
Log-likelihood	−272.233	−298.801	−310.844
Pseudo *R*-squared	0.059	0.057	0.054

Note: Logistic estimation with clustered standard errors. All time-varying independent variables are lagged one year. Odds ratios are reported with *p*-values based on robust, clustered, standard errors in parentheses: coefficient values below one indicate that the odds of conflict fall with increases in the corresponding variable. The *p*-values for years and party age are calculated assuming that the interaction term is zero. All variables are from the year before the crisis. Other independent variables (noncontiguous state territory, mountainous terrain) and the constant are not reported.

Source: Political variables, Database of Political Institutions (Beck and others 2001); conflict variables, Correlates of War (Sambanis 2004); control variables, (Fearon and Laitin 2003).

(table 1). Their effects are also economically meaningful. To show this more clearly, the table reports the odds ratios, constant across parameter values, rather than coefficient values. Odds ratios less than one imply that an increase in the variable reduces the risk of insurgency; odds ratios greater than one imply that increases raise the risk of insurgency.[8]

The presence of a programmatic governing party reduces the odds of conflict by half. Every additional year by which the age of the governing party exceeds the years a ruler has been in office reduces the odds of conflict by about 2 percent a year. The regressions in the all-country sample omit per capita income. In the presence of per capita income none of the political variables is significant, but this is almost surely because of the effect of income on political incentives to invest in the credibility of broadly programmatic promises, as discussed at length in the next section.

Among other determinants of insurgency, oil is statistically significant in only one regression; religious and ethnic fractionalization are significant determinants of insurgency in all cases. An increase in religious and ethnic fractionalization from the minimum possible value (zero) to the maximum (one)

8. The odds ratio reveals how the odds of conflict change after a one-unit increase in the independent variable. If the odds before the change in a variable are 3:1 and the odds after are 4:1, the odds ratio is 1.33. The marginal effects of coefficient changes estimated from logistic specifications depends on parameter values.

TABLE 2. Credibility Effects in Poor Countries (Dependent Variable: Conflict Onset, 0–1)

Variable	(1)	(2)	(3)
Age of largest party minus leader years in office	0.982 (0.04)		
Years of continuous competitive elections		0.968 (0.10)	
Largest government party programmatic? (1–0)			0.572 (0.11)
Income per capita	0.597 (0.14)	0.631 (0.16)	0.653 (0.20)
Ln (population)	1.178 (0.07)	1.135 (0.21)	1.155 (0.10)
Oil exporter dummy variable (0–1)	1.558 (0.21)	1.266 (0.55)	1.152 (0.72)
Ethnic fractionalization	0.687 (0.54)	0.967 (0.96)	0.926 (0.89)
Religious fractionalization	2.741 (0.20)	2.33 (0.29)	2.961 (0.14)
N (country years)	1,272	1,353	1,377
Log-likelihood	−185.416	−203.590	−211.887
Pseudo *R*-squared	0.047	0.034	0.043

Note: Logistic estimation with clustered standard errors. All time-varying independent variables are lagged one year. Odds ratios are reported with *p*-values based on robust, clustered, standard errors in parentheses: coefficient values below one indicate that the odds of conflict fall with increases in the corresponding variable. The *p*-values for years and party age were calculated assuming that the interaction term is zero. All variables are for the year before the crisis. Other independent variables (noncontiguous state territory, mountainous terrain) and the constant are not reported. Poor countries in year *t* are those with incomes below the world median that year.

Source: Political variables, Database of Political Institutions (Beck and others 2001); conflict variables, Correlates of War (Sambanis 2004); control variables, (Fearon and Laitin 2003).

increases the odds of conflict by 300–400 percent. Consistent with the argument that fractionalization is associated with conflict precisely when politicians are not credible and that large natural resource rents create a disincentive for politicians to invest in broad credibility, the omission of any of these variables substantially boosts the significance of the two party variables. Population size is also significant: countries with larger populations have a greater likelihood of experiencing conflict in this specification.[9]

The vast majority of civil wars occur in countries that are poorer than the world median in the year they occur. In the period 1975–2000, 57 conflicts broke out in countries with per capita income at or less than the world median. Of the remaining 15 countries in which conflict occurred, 9 had no income data. Of these, at least six (Afghanistan, Azerbaijan, Georgia, Liberia, Moldova, and Yemen) were near or below the median the year before war broke out. Examining poorer countries separately is therefore a sensible estimation strategy, unless one has strong prior reasons to believe that the dynamics of civil war are on average the same in poor and rich countries. Given that citizen income affects political incentives to invest in credible,

9. The population result is common. Collier and Hoeffler (1998) attribute it to the increased desire for secession in larger countries. Fearon and Laitin (2003) point to the greater difficulties that larger populations create for control by the state. The argument in this article suggests a third explanation: the costs to politicians of making credible appeals to citizens rises with population size.

TABLE 3. Credibility Effects in Nondemocracies (Dependent Variable: Conflict Onset, 0–1)

	All nondemocracies		Poor nondemocracies	
Variable	(1)	(2)	(3)	(4)
Age of largest party minus leader years in office)	0.990 (0.35)		0.980 (0.09)	
Largest government party programmatic? (1–0)		0.469 (0.05)		0.407 (0.04)
Income per capita			0.597 (0.20)	0.680 (0.32)
Ln (population)	1.277 (0.06)	1.250 (0.04)	1.298 (0.09)	1.224 (0.15)
Oil exporter dummy variable (0–1)	1.314 (0.41)	1.080 (0.82)	1.049 (0.94)	0.883 (0.83)
Ethnic fractionalization	1.519 (0.51)	1.520 (0.52)	0.536 (0.38)	0.656 (0.50)
Religious fractionalization	3.623 (0.16)	2.820 (0.20)	2.399 (0.41)	2.093 (0.45)
N (country years)	1,553	1,678	1,019	1,058
Log-likelihood	−202.658	−215.868	−149.379	−156.230
Pseudo *R*-squared	0.038	0.040	0.041	0.046

Note: Logistic estimation with clustered standard errors. All time-varying independent variables are lagged one year. Odds ratios are reported with *p*-values based on robust, clustered, standard errors in parentheses: coefficient values below one indicate that the odds of conflict fall with increases in the corresponding variable. The *p*-values for years and party age were calculated assuming that the interaction term is zero. All variables are for the year before the crisis. Other independent variables (noncontiguous state territory, mountainous terrain) and the constant are not reported. Poor countries in year t are those with incomes below the world median that year.

Source: Political variables, Database of Political Institutions (Beck and others 2001); conflict variables, Correlates of War (Sambanis 2004); control variables, (Fearon and Laitin 2003).

programmatic promises to citizens, there is ample reason to suspect that the opposite is true.

Unlike the specifications in table 1, the poor-country specifications include per capita income (table 2); the political results strengthen if income is excluded, however. Despite the presence of per capita income, the relevant political variables are highly significant and the odds ratios are little changed from the all-country results. Equally important, however, is that the significance of other commonly investigated correlates of rebellion changes substantially in poorer countries. Though income is significant in the sample of all countries (not reported), it is insignificant in the sample of poor countries. It appears that income effects are driven by the complete absence of conflicts among the richest 30 percent of countries. The effects of oil exports and ethnic and religious fractionalization all fall significantly compared with the results for all countries. Even if income is excluded, ethnic fractionalization and oil exports remain insignificant. The insignificance of ethnic and religious fractionalization is not driven by their simultaneous entry into the specifications; although they are correlated at 0.32, neither becomes significant when the other is dropped.

The hypotheses related to the institutionalization of political parties are relevant to both democracies and nondemocracies; nondemocracies are therefore

also examined separately. Countries are identified using two DPI variables, the legislative and executive indices of electoral competition. Countries in which each of these indices takes a maximum score of seven (multiple parties compete in the elections and none receives more than 75 percent of the vote) are classified as democracies; the remainder are classified as nondemocracies.

Two sets of regressions are presented in table 3: those looking only at nondemocracies, using the specifications of table 1, and those looking only at poor nondemocracies, using the specifications of table 2 (the years of continuing competitive elections are omitted, because they are always zero in the nondemocratic subsample). The presence of a programmatic governing party has an even stronger effect on outcomes, whether compared with all countries or only poor countries (the presence of a programmatic party reduces the odds of conflict more than in tables 1 and 2). The party age variable also has a stronger effect on poor nondemocracies than it does on poor countries only, although the party age variable is insignificant in the sample of all nondemocracies.

Other common determinants of conflict are uniformly insignificant across all four nondemocracy regressions. Per capita income is insignificant in the sample of poor nondemocracies. The magnitudes of the effects of oil exports and religious and ethnic fractionalization drop substantially; they are insignificant in both subsamples. Only population, among the alternative explanations of insurgency, continues to matter.

VI. Robustness: Income, Oil, Social Polarization, and Political Credibility

The results in tables 1–3 are striking because they are almost entirely robust to controls for variables that theory predicts should influence or be influenced by political credibility: per capita income, ethnic and religious fractionalization, and oil. Oil is a typical control variable in the conflict literature, because, as Collier and Hoeffler (1998) and Elbadawi and Sambanis (2002) argue, access to natural resource rents is both a goal of rebel movements and a way for rebel movements to finance themselves (Fearon 2005 disputes the importance of natural resources in civil war). However, oil or other natural resource rents also make rulers more reluctant to invest in credibility. For example, when rents are high, the benefits to rulers of adopting institutions that force them to share rents (for example, larger, more-institutionalized ruling parties) are smaller.

Indeed, there is a strong negative correlation between countries coded by Fearon and Laitin (2003) as oil exporters and the variables used here to capture the extent to which politicians can make broadly credible promises to voters. Moore (2004) provides an illustration of the discovery of oil leading to new and more circumscribed institutional arrangements. Before the discovery of oil, Kuwaiti merchants shared political control with the sheiks of the dominant al-Sabah family: "Manpower and financial power gave Kuwaiti

merchants an early sense of equality with the ruling al-Sabahs.... Commerce was not viewed as subordinate to politics. Indeed, politics needed commerce" (p. 31). However, with the discovery of oil in 1938, "the elected municipality board, which had served as an *asil* [elite] merchant enclave since 1932, was replaced with an appointed board of shaikhs. As royal family members took control of government ministries, merchant committees within those bodies, designed to provide policy input, were disbanded" (p. 42). Nevertheless, despite the relation between political variables and rents, the estimates of the political variables in tables 1–3 are not sensitive to the presence of the oil control.

Using various measures, researchers (Reynal-Querol 2002; Soysa 2002) have also found that polarization is a determinant of civil war (Fearon and Laitin 2003 argue that this relation is not robust). The argument is that sharp differences in policy preferences across social groupings create social cleavages that increase insurgent pressures. However, as the earlier discussion of several civil wars concludes, social polarization is endogenous to political decisionmaking. In particular, noncredible politicians seek to expand the set of citizens to whom they can make credible commitments. In socially fragmented countries the low-cost way for them to do this may be by making ethnic or religious appeals. Politicians build a reputation for policy stances favoring particular ethnic groups, but in broadening the reach of their credible commitments, they exacerbate political polarization.

Van de Walle (2003) documents this in the case of emerging African democracies, where he finds that the overwhelming organizational principle of new political parties is ethnic or linguistic. Under these conditions, social polarization leads to conflict, because noncredible politicians use social identity to build support. Nevertheless, the presence of controls for the social polarization variables does not affect the credibility results reported earlier.

Per capita income is the most troublesome of the three controls, because none of the results in table 1 is significant when a control for income is added. However, the arguments linking income to clientelist political strategies are well established in the literature. Poorer voters are more susceptible than richer voters to targeted transfers (Dixit and Londregan 1996). The political advantages of relying on clientelist strategies that target the few voters who believe politician promises are therefore greater in poorer countries, reducing political incentives to invest in credibility but also increasing insurgency incentives and giving rise to an association between insurgency and income that runs through political credibility.[10]

10. The fact that the lack of political credibility might be both a cause and consequence of low incomes suggests the possibility of a credibility-based poverty trap. Keefer and Vlaicu (forthcoming) anticipate such a possibility, proving that where the costs to politicians of clientelist electoral strategies are sufficiently low, they will delay the investments needed to make their promises broadly credible.

While it is reasonable to exclude per capita income from the table 1 regressions, given the links with the credibility variables, there is some concern that to the extent that there are additional reasons for an association between income and civil war that do not operate through credibility, the exclusion of income biases the results. The fact that income is insignificant in tables 2 and 3 provides some reason to believe that this bias is small. At the same time, the three reasons in the literature for including income seem less persuasive than the arguments that link income to political credibility. Moreover, two of these reasons have a credibility interpretation that reinforces the argument here for excluding income.

Some authors (including Collier and Hoeffler 1998, Elbadawi and Sambanis 2002, and Fearon and Laitin 2003) argue that income captures state capacity to combat insurgencies. However, state capacity is a public good that broadly credible politicians have a greater incentive to provide. For example, law and order and strong administrative capacity, from which counterinsurgency capacity springs, benefit all citizens. The earlier discussion predicts that such public goods will be underprovided in low-credibility states. In fact, Keefer (2007) presents evidence that both the rule of law and bureaucratic quality are lower in low-credibility democracies.

Moreover, it is not clear whether income is a binding constraint on the development of state capacity in any case. Military spending, for example, seems to bear no systematic negative relation with per capita income; if anything, military spending as a fraction of GDP varies inversely with per capita income. Moreover, countries confronting threats show remarkable agility in increasing military spending independent of income. Despite roughly similar incomes per capita, India, Pakistan, and Sri Lanka exhibit vast differences in military spending: Pakistan spends twice as much as India and Sri Lanka 50 percent more than India, differences that seem more correlated with the fraction of their territory the countries view as vulnerable to armed dispute than with differences in per capita income.[11]

Hegre and others (2001) posit a second channel through which income might drive conflict: as incomes rise, class conflicts moderate and opposing segments of society are more amenable to peaceful resolutions of their disagreements. One indicator of class-based mobilization is the existence of political parties advocating the interests of different classes. The argument here predicts that it is precisely the absence of such parties that makes conflict more likely. The tests below examine this thesis directly, showing that direct measures of programmatic—possibly class-based—parties reduce the probability of insurgency.

Collier and Hoeffler (1998) and Elbadawi and Sambanis (2002) interpret income as the opportunity cost of rebel labor: insurgencies can form at lower

11. India, with per capita purchasing power parity–adjusted income of $2,312 in 1999, is richer than Pakistan ($1,818) but poorer than Sri Lanka ($3,229).

cost in poorer countries. However, the same argument could be made about the ability of governments to recruit soldiers into its counterinsurgency forces, suggesting no clear bias from omitting income from a conflict regression.

VII. Other Robustness Issues

The results in tables 1–3 raise two broad concerns. One is that the specifications do not control for reasonable alternative explanations for conflict; the other, related concern is that the results are the consequence of endogeneity bias, which reflects either reverse causality or omitted-variable bias. These concerns are endemic in estimating the determinants of conflict.[12] Valid instruments are rare and entirely absent for investigations of the political drivers of conflict.

One response to endogeneity bias is to conduct case studies, which may reveal causal mechanisms using information that is not available with aggregate data. The earlier thumbnail sketches of several conflicts illustrate the role of political credibility in conflict. The estimations reported in this section are based on two additional strategies: estimation across different subsamples and estimation with additional control variables. Each captures at least a substantial subset of potential causes of endogeneity bias. The all-country results of table 1 are robust to all alternative estimation strategies; other results are somewhat more fragile, although the smaller sizes of the samples make this unsurprising. In both tables 4 and 5 the estimation of the effects of party age relative to executive tenure in the case of all nondemocracies is omitted: this variable is not significant in the original regression in table 3 or in any of the specifications in tables 4 and 5.

Estimates across Subsamples

Conflicts are rare: the onset of conflict occurs in less than 2 percent of the country-year observations investigated here. As a consequence, subsamples can vary significantly in the number of conflicts they contain, affecting the robustness of results simply because the number of conflicts in the subsample is small. Nevertheless, one might reasonably ask whether the results reported in tables 1–3 are robust to four additional subsamples.

First, the fall of communism might have had a substantial effect on civil war; the end of the Cold War ended proxy wars contested and financed by the two superpowers. Nevertheless, results reported earlier are robust when one uses only observations from the 1990s (table 4, row 1). In only one case is a significant result from the earlier tables insignificant in the 1990s subsample. The magnitude of the effect of years of continuous competitive elections is

12. Conflict almost always shortens the duration of democracies and the lives of parties. The estimations in tables 1–3 avoid this source of endogeneity, however, because all of the explanatory variables are measured in the year before conflict occurs.

TABLE 4. Political Credibility and Conflict Risk across Subsamples (Dependent Variable: Conflict Onset, 0–1)

	All countries			Poor countries			Nondemocracies	Poor nondemocracies	
Subsample	Party age (years in office)	Election years	Programmatic party	Party age (years in office)	Election years	Programmatic party	Programmatic party	Party age (years in office)	Programmatic party
1990s only	0.977 (0.002)	0.79 (0.00)	0.370 (0.06)	0.981 (0.06)	0.916 (0.35)	0.716 (0.54)	0.289 (0.03)	0.980 (0.10)	0.27 (0.02)
Excluding country years with two or more conflicts	0.985 (0.06)	0.964 (0.04)	0.363 (0.005)	0.98 (0.06)	0.993 (0.82)	0.406 (0.03)	0.46 (0.06)	0.978 (0.12)	.303 (0.02)
Cross-section, average 1975–99 (ordinary least squares, not odds ratios)	−0.005 (0.04) $N = 122$, N (conflicts) = 42	−0.011 (0.001) $N = 124$, N (conflicts) = 44	−0.68 (0.004) $N = 125$, N (conflicts) = 45	−0.005 (0.37) $N = 60$, N (conflicts) = 21	0.014 (0.77) $N = 60$, N (conflicts) = 21	−0.74 (0.031) $N = 61$, N (conflicts) = 21	−1.59 (0.001) $N = 35$, N (conflicts) = 13	−0.008 (0.45) $N = 20$, N (conflicts) = 5	−1.97 (0.002) $N = 11$, N (conflicts) = 5
Cross-section, conflicts 1990–99, political variables 1975–90 (ordinary least squares)	−0.002 (0.07) $N = 104$, N (conflicts) = 19	−0.009 (0.001) $N = 124$, N (conflicts) = 26	−0.41 (0.005) $N = 125$, N (conflicts) = 27	−0.003 (0.54) $N = 56$, N (conflicts) = 11	−0.03 (0.12) $N = 60$, N (conflicts) = 14	−0.54 (0.007) $N = 61$, N (conflicts) = 14	−0.60 (0.03) $N = 35$, N (conflicts) = 11	−0.001 (0.80) $N = 18$, N (conflicts) = 3	−1.77 (0.06) $N = 11$, N (conflicts) = 5

Note: Estimated coefficients of respective political variables using specifications from tables 1–3, with change noted at far left. See notes to tables 1–3. N (conflict) indicates number of countries in which conflicts occurred over the period.

Source: Political variables, Database of Political Institutions (Beck and others 2001); conflict variables, Correlates of War (Sambanis 2004); control variables, (Fearon and Laitin 2003).

TABLE 5. Political Credibility and Conflict Risk with Additional Controls (Dependent Variable: Conflict Onset, 0-1)

	All countries			Poor countries			Nondemocracies	Poor nondemocracies	
Control variable	Party age (years in office)	Election years	Programmatic party	Party age (years in office)	Election years	Programmatic party	Programmatic party	Party age (years in office)	Programmatic party
Controlling for year effects	0.986 (0.03)	0.958 (0.02)	0.484 (0.04)	0.978 (0.02)	0.965 (0.08)	0.493 (0.08)	0.493 (0.08)	0.976 (0.04)	0.380 (0.04)
Substituting per capita income from *World Development Indicators*, filling missing values through extrapolation	n.a	n.a	n.a	0.98 (0.02)	0.96 (0.02)	0.55 (0.10)	n.a	0.980 (0.10)	039 (0.09)
Controlling for per capita income growth	0.986 (0.03)	0.954 (0.020)	0.553 (0.10)	0.98 (0.02)	0.967 (0.08)	0.673 (0.24)	0.628 (0.26)	0.983 (0.14)	0.528 (0.17)
Controlling for continent dummies	0.988 (0.07)	0.951 (0.02)	0.502 (0.04)	0.977 (0.02)	0.939 (0.01)	0.574 (0.15)	0.546 (0.11)	0.986 (0.26)	0.478 (0.06)
Controlling for Köppen-Geiger climate zones	0.988 (0.062)	0.96 (0.04)	0.60 (0.11)	0.98 (0.01)	0.96 (0.006)	0.674 (0.238)	0.630 (0.22)	0.980 (0.09)	0.53 (0.17)

n.a Not applicable.

Note: Estimated coefficients of respective political variables using tables 1–3 specifications, with change noted at far left. See notes to tables 1–3. Five Köppen-Geiger variables report the fraction of country covered by each of five climate zones.

Source: Political variables, Database of Political Institutions (Beck and others 2001); conflict variables, Correlates of War (Sambanis 2004); control variables, (Fearon and Laitin 2003); Köppen-Geiger climate variables: Sachs (2003).

TABLE 6. Political Credibility and Political Institutions (Dependent Variable: Conflict Onset, 0–1)

	Table 1: All countries			Table 2: Poor countries		
Formal institutional variables	Party age (years in office)	Election years	Programmatic party	Party age (years in office)	Election years	Programmatic party
Political variables from column heading	0.97 (0.00)	0.95 (0.01)	0.31 (0.01)	0.972 (0.00)	0962 (0.00	0.46 (0.03)
Presidential = 0, semipresidential = 1, parliamentary = 2	1.003 (0.00)	1.002 (0.02)	1.002 (0.04)	1.002 (0.00)	1.003 (0.00)	1.003 (0.00)
District magnitude	1.00 (0.93)	1.00 (0.94)	1.00 (0.86)	1.00 (0.15)	1.00 (0.43)	1.00 (0.43)
Plurality = 1; Proportional representation = 0	1.00 (0.94)	1.00 (0.91)	1.00 (0.86)	1.00 (0.19)	1.00 (0.19)	1.00 (0.25)
Number of observations	1,847	1,972	1,972	795	847	847
R-squared	0.12	0.12	1972, 0.12	0.09	0.10	0.11

Note: Each column reports the results of specifications corresponding to tables 1 and 2, augmented with the formal political institutions. Other table 1–2 control variables are not reported. Institutional variables reduce sample sizes compared with tables 1 and 2, but samples are close to tables 1 and 2 and results are robust if district magnitude and voting rule variables are excluded.

Source: See text.

much greater in the poor country, 1990s-only regressions of table 4 than in the poor-country regressions of table 2. However, their impact on conflict is no longer statistically significant.

For all of the regressions in tables 1–3, clustered standard errors are reported; this reduces the risk that multiple conflict countries drive the significance of the results reported in these tables. This may not eliminate all sources of bias resulting from including country-year observations in countries that experienced two or more conflicts, however. To be sure that this is the case, regressions in tables 1–3 are re-estimated excluding all country-year observations beginning with the onset of a second conflict (table 4, row 2). The earlier results remain largely robust. One of the significant results in table 2 and one in table 3 are insignificant in table 4; however, the insignificant result in table 2 (the effect of programmatic parties) is highly significant in the second row of the poor-country sample in table 4.

Most of the explanatory power of the political variables comes from their cross-country variation.[13] It might be argued, therefore, that the statistical significance of the results in tables 1–3 is exaggerated by the large sample sizes of panel data. The clustered standard errors used in tables 1–3 substantially reduce this downward bias in standard errors. An additional solution to this problem is to examine only the purely cross-section variation in the conflict and control variables (table 4, row 3). The left-hand side variable is therefore the total number of conflicts a country experienced between 1975 and 1999. The control variables are averaged over the period, and coefficients are estimated using simple ordinary least squares. Once again, all of the results from table 1, and five of eight of the significant results in tables 1–3, remain significant. Moreover, one of the insignificant results (in table 2) is significant in table 4.

The row 3 results, using whole-period averages of both the political and conflict variables, could be driven by reverse causality. To determine if this is the case, the estimation is repeated, this time averaging the political variables over the period 1975–90 (row 4). The effects of early political variables are then estimated on the number of conflicts over the period 1990–99 (the other control variables are 1975–99 averages, as in row 3). While the coefficient estimates drop somewhat compared with row 3, the pattern of statistically significant results is the same. This suggests that reverse causality is not driving row 3 results, indicating that the significance of the results in tables 1 is not the spurious result of using the many observations contained in the panel data.

13. More important, it is not clear how tests should be specified over time. Among other things, it is unclear over what time period changes in the political variables should begin to exert an effect on the risk of conflict. Yearly fixed-effects estimations are inappropriate, for example, because early increases in party age or years of continuous elections almost surely do not have the same effects as later increases, while the effect of programmatic parties is unlikely to emerge the year after they are recognized.

Adding New Control Variables

It is also possible that the political variables in table 1–3 are sensitive to the use of particular versions of some control variables (for example, income or social polarization) or some alternative explanations of conflict are omitted. Numerous alternative specifications indicate that this is unlikely to be the case. The first row of table 5 includes controls for year effects: as in the 1990s regressions in table 4, global shocks such as the end of the Cold War could have influenced both political institutions within countries and the risk of conflict. However, all of the significant results and one of the insignificant results in tables 1–3 are significant when year effects are taken into account (row 1).

The per capita income variable taken from Fearon and Laitin (2003) is not adjusted for purchasing power parity. The political variables could therefore be significant because they spuriously capture mismeasured income effects. In fact, this is not the case. *World Development Indicators* does not provide income data of any kind for many countries in the mid-1970s, including a number of conflict countries. To adjust for this, missing values of the available purchasing power parity–adjusted per capita income data were filled in using simple linear extrapolation (this procedure never accounted for more than four years of a country's income data). The results are robust to the use of purchasing power parity–adjusted per capita income (table 5, row 2).

Results are similarly robust when alternative measures of ethnic and religious fractionalization compiled by Alesina and others (2002) are substituted into the regressions of tables 1–3 (results not reported). Specifically, all of the political coefficients in tables 1 and 2 are significant (including one that was previously insignificant). However, only the last specification in table 3 (which looks at the effects of programmatic parties in poor nondemocracies) is significant (results not reported). It is not surprising that the nondemocratic regressions are less robust, because the political variables are less noisy, more-accurate descriptions of political credibility in democracies than in nondemocracies.

Economic growth is a proximate cause of conflict and therefore captures political sources of conflict (such as credibility) that affect both growth and conflict. Despite this, five of the eight significant results in tables 1–3 remain significant when growth is added; as with the fractionalization variables, the table 1 and 2 results are particularly robust (table 5, row 3).

Continent controls (for Africa, East Asia, and South Asia) and climate zone variables (five Köppen-Geiger climate zones, such as whether countries are primarily desert, tropical, dry, or wet) are separately added to the table 1–3 regressions to capture other unobserved influences on conflict. Their inclusion is not precisely a robustness test of the regressions in tables 1–3, however, because theory predicts that their influence should operate in part through the political variables under examination here. For example, political incentives to build broadly credible institutions (for example, programmatic political parties or institutionalized ruling parties in autocracies) are directly related to the growth

payoff that leaders can expect if they encourage investment with such institutions. Unfavorable locations (for example, poor climactic conditions unfavorable to agriculture or continents characterized by adverse neighborhood effects, such as generally thin population densities) have more limited growth potential. The lower the growth payoff to credible institutions, the lower is the increment to regime security that leaders can achieve by adopting these institutions. Under these circumstances one is less likely to see the political arrangements that are the focus of attention here. In fact, continent and climate zone dummy variables are significant determinants of each of the three political variables used here, in the all-country, poor-country, and nondemocracy samples.

Despite the significant, theoretically predicted relation between continent and climate zone dummy variables and the political variables, nearly all of the regressions in tables 1 and 2 are robust to the inclusion of the dummy variables (rows 4 and 5, table 5). Most of the significant nondemocracy results from table 3, which are more vulnerable in any case to the noisiness of the political variables, are insignificant. However, even in these cases, the estimated impact of the political variables remains large. For example, though not statistically significant, the presence of programmatic parties still reduces the odds of conflict in nondemocracies by about 46 percent.

Reynal-Querol (2002) argues that democracies with more inclusive institutions (proportional representation as opposed to majoritarian or presidential systems) confront a lower risk of ethnic civil war. It is possible that the political effects measured here are significant only because they spuriously capture the effect of these formal institutions. However, results are robust to controls for three formal institutional variables: whether countries are presidential, semipresidential, or parliamentary; whether they are majoritarian, as reflected in their district magnitudes (low indicates more majoritarian); and whether they use plurality or majoritarian electoral rules. All of these are available in the DPI.

All of the results in tables 1 and 2 (the only samples containing democracies) are actually stronger when formal institutional variables are taken into account (table 6). At the same time, the estimated effects of these formal institutions on the probability of civil war differ from those found in previous research. Neither district magnitudes nor electoral proportionality are significant determinants of insurgency. Contrary to expectations, the odds of a conflict occurring are significantly higher in parliamentary systems. This result contrasts with predictions in the insurgency literature and is related to an ongoing debate in the literature regarding the stability of presidential and parliamentary systems. One study (Cheibub 2006) also finds that, contrary to received wisdom, presidential systems are not less stable once one controls for whether democracy is preceded by military government.

It might be the case that the political variables simply capture the fact that countries that have been sovereign for a longer period have had more time to develop key political institutions and to work out underlying conflictual issues. *The World Factbook* of the Central Intelligence Agency (2007) has data on

countries' first year of sovereignty. The years since the first year of sovereignty, particularly in log form, are positively correlated with the age of the largest government party and with the years of continuous competitive elections. However, whether in log or linear form, the variable *years of independence* has no effect on the results reported in tables 1–3. In all cases the magnitude and significance of the estimated political effects hardly change. Only the log of years since independence is ever close to significant and then only in the poorer country samples. It turns out that more years of independence is weakly associated with greater odds of experiencing a conflict.

Despite the robustness of the results to a variety of control variables, it is still possible that omitted variables drive the negative correlation of conflict risk and political credibility. One argument made here is that the ability to make credible promises, as captured by the political variables used in the empirical analysis, undermines counterinsurgency capacity and encourages insurgency. It is conceivable, however, that some governments simply enjoy a better counterinsurgency endowment or, more generally, are endowed with unobserved social conditions that lower the risk of conflict. Given the reduced risk of conflict, politicians find it more cost-effective to invest in the ability to make credible promises into the future. To the extent that this is true, results reported here would be the product of reverse causality.

While the empirical analysis does not completely foreclose the possibility of reverse causality, this is not the most plausible explanation for the results here. The row 4 results in table 4—which show that in the cross-section estimations, the 1975–90 values of the political variables are significant predictors of conflict over 1990–99—suggest that any such reverse causality would need to be driven by time-invariant omitted variables. Furthermore, these omitted factors would have to be largely uncorrelated with the control variables examined here, including years of independence, geography, ethnic and religious fractionalization, and per capita income, and strongly related to the political variables that are the object of the analysis.

It is possible that such time-invariant omitted effects exist. However, it is also difficult to identify social conditions that reduce the risk of conflict without also directly affecting political incentives to invest in the ability to make credible commitments. Unobserved ethnic tensions, for example, may make conflict more likely, but they do so in large part by raising the costs to politicians of making credible promises across ethnic groups. The endowment of an effective counterinsurgency force may deter insurgencies from forming, but it is difficult to see how leaders can sustain effectiveness if they are unable to make credible commitments to the force. Indeed, it is more likely that the omitted variables that would have the greatest impact on the analysis here are precisely those that influence political incentives to make credible commitments. The influence of omitted factors such as these, however, is entirely consistent with the argument proposed here.

VIII. Conclusion

The credibility of political promises plays a crucial role in "normal" politics. The evidence presented here shows that it also heavily influences the transition from normal politics to civil war. In both democratic and nondemocratic countries, where political actors are unable to make credible promises to a sufficient fraction of citizens, they pursue public policies that benefit a few at the expense of most citizens, increasing citizen tolerance for insurgent movements. Moreover, the willingness and ability of governments to build effective counterinsurgency capacity declines. As a consequence, political credibility—proxied by the age of the ruling party, whether the ruling party is programmatic, or by the years of continuing competitive elections—significantly reduces the probability of civil war.

These results are relevant to postconflict recovery strategies. Countries that have experienced conflicts are at particular risk of subsequent conflicts; the evidence here suggests that this is related to the effect of past conflicts on the credibility of a country's political competitors. Efforts to build the broad credibility of political actors is therefore likely to be key to successful postconflict recovery. Doing so implies, among other things, ensuring that citizens give governments (rather than donors) credit for improvements in welfare; controlling access to rents to reduce incentives for renewed conflict in noncredible settings; and supporting the informal institutions of politics (political parties and information transmission to voters).

These efforts contrast with other alternative reform priorities that are often voiced. For example, some argue for delivering quick wins in the form of enhanced service delivery; doing so contributes to sustainable recovery, however, only if it builds political reputations for broad public goods provision. Others argue for fine-tuning institutional arrangements that grant contending groups a veto over policymaking, making intergroup agreements credible. These arrangements need to be supplemented by changes that allow the leaders of contending groups to make credible promises to citizens more broadly.

Funding

This research was part of the ongoing World Bank Research Project on Post-Conflict Transitions, made possible through the Knowledge for Change Program with support from the Norwegian Ministry of Foreign Affairs.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2002. "Fractionalization." *Journal of Economic Growth* 8(2):155–94.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1):165–76.

Cheibub, Jose. 2006. *Presidentialism, Parliamentarism and Democracy*. New York: Cambridge University Press.

Collier, Paul, and Anke Hoeffler. 1998. "On Economic Causes of Civil War." *Oxford Economic Papers* 50(4):563–73.

Dixit, Avinash, and John Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *Journal of Politics* 58(4):1132–55.

Elbadawi, Ibrahim, and Nicholas Sambanis. 2002. "How Much War Will We See? Explaining the Prevalence of Civil War." *Journal of Conflict Resolution* 46(3):307–34.

Esteban, Joan-Mara, and Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica* 62(4):819–51.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer than Others?" *Journal of Peace Research* 49:438–507.

———. 2005. "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution* 41(3):275–301.

Fearon, James D, and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50(1):5–26.

Garfinkel, Michelle R, and Stergios Skaperdas. 2000. "Conflict without Misperceptions or Incomplete Information: How the Future Matters." *Journal of Conflict Resolution* 44(6):793–807.

Gehlbach, Scott, and Philip Keefer. 2006. "Ruling Party Institutionalization and Autocratic Success." World Bank, Development Research Group, Washington, D.C.

Haber, Stephen. 2006. "Authoritarian Government." InBarry Weingast, and Donald Wittman, eds., *The Oxford Handbook of Political Economy*. New York: Oxford University Press.

Haber, Stephen, Armando Razo, and Noel Maurer. 2003. *The Politics of Property Rights: Political Instability, Credible Commitments, and Economic Growth in Mexico, 1876–1929*. New York: Cambridge University Press.

Hegre, Havard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change and Civil War, 1816–1992." *American Political Science Review* 95(1):33–48.

Jagger, Keith, and Ted Robert Gurr. 1995. "Transitions to Democracy: Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32(4):469–82.

Jones, Mark P. 2005. "The Role of Parties and Party Systems in the Policymaking Process." *Paper prepared for the Workshop on State Reform, Public Policies, and Policymaking Processes*. Inter-American Development Bank, Washington, D.C., February 28–March 2.

Keefer, Philip. 2005. "Programmatic Parties: Where Do They Come from and Do They Matter?" Paper presented at the 2005 meetings of the International Society for New Institutional Economics, Barcelona.

———. 2006. "Governance and Economic Growth in China and India." World Bank, Development Research Group, Washington, D.C.

———. 2007. "Clientelism, Credibility and the Policy Choices of Young Democracies." *American Journal of Political Science*. 51(4):804–21.

Keefer, Philip, and Stephen Knack. 2002. "Polarization, Property Rights and the Links between Inequality and Growth." *Public Choice* 111(1–2):127–54.

Keefer, Philip, and Razvan Vlaicu. (forthcoming). "Clientelism, Credibility and Democracy." *Journal of Law, Economics and Organization*.

Kitschelt, Herbert, Zdenka Mansfeldova, Radoslaw Markowski, and Gabor Toka. 1999. *Post-Communist Party Systems: Competition, Representation, and Inter-Party Collaboration.* Cambridge: Cambridge University Press.

Lewis, Peter, ed. 1998. *Africa: Dilemmas of Development and Change.* Boulder, CO: Westview Press.

Lund, Michael S. 2005. "Greed and Grievance Averted: How Macedonia Averted Civil War, 1990–2001." In P. Collier, and N. Sambanis, eds., *Understanding Civil War: Evidence and Analysis.* Washington, D.C.: World Bank.

Makdisi, Samir, and Richard Sadaka. 2005. "The Lebanese Civil War, 1975–90." In P. Collier, and N. Sambanis, eds., *Understanding Civil War: Evidence and Analysis.* Washington, D.C.: World Bank.

Moore, Pete W. 2004. *Doing Business in the Middle East: Politics and Economic Crisis in Jordan and Kuwait.* New York: Cambridge University Press.

Ndikumana, Lonce, and Justine Nannyonjo. 2007. "From Failed State to Good Performer? The Case of Uganda." InJames K. Boyce, and Madalene O'Donnell, eds., *Peace and the Public Purse: Economic Policies for Postwar Statebuilding*, 12–54. Boulder, CO: Lynne Rienner Publishers.

North, Douglass C., and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49(4):803–32.

Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Public Policy.* Cambridge, MA: MIT Press.

Reynal-Querol, Marta. 2002. "Ethnicity, Political Systems and Civil Wars." *Journal of Conflict Resolution* 46(1):29–54.

Rodas-Martini, Pablo. 2007. "Building Fiscal Provisions into Peace Agreements: Cautionary Tales from Guatemala." In James K. Boyce, and Madalene O'Donnell, eds., *Peace and the Public Purse: Economic Policies for Postwar Statebuilding*, 85–118. Boulder, CO: Lynne Rienner Publishers.

Sachs, Jeffrey D. 2003. *Institutions Don't Rule: Direct Effects of Geography on per Capita Income.* NBER Working Paper 9490. Cambridge, MA: National Bureau of Economic Research.

Sambanis, Nicholas. 2004. "What is Civil War? Conceptual and Empirical Complexities." *Journal of Conflict Resolution* 48(6):814–58.

Soysa, Indra de. 2002. "Paradise Is a Bazaar? Greed, Creed and Governance in Civil War, 1989–1999." *Journal of Peace Research* 39(4):395–416.

van de Walle, Nicolas. 2003. "Presidentialism and Clientelism in Africa's Emerging Party Systems." *Journal of Modern African Studies* 41(2):297–321.

Walter, Barbara F. 1997. "The Critical Barrier to Civil War Settlement." *International Organization* 51(3):335–64.

The Aftermath of Civil War

Siyan Chen, Norman V. Loayza, and Marta Reynal-Querol

Using an event-study methodology, the article analyzes the aftermath of civil war in a cross-section of countries. It focuses on cases where the end of conflict marks the beginning of relatively lasting peace. The analysis considers 41 countries involved in internal wars over the period 1960–2003. To provide a comprehensive evaluation of the aftermath of war, a range of social areas is considered: basic indicators of economic performance, health and education, political development, demographic trends, and conflict and security issues. For each indicator the post- and pre-war situations are compared and their dynamic trends during the post-conflict period are examined. The analysis is conducted in both absolute terms and relative to control groups of countries that are similar except for conflict. The findings indicate that even though war has devastating effects and its aftermath can be immensely difficult, when the end of war marks the beginning of lasting peace, recovery and improvement are achieved. JEL code: O11

War has devastating consequences, including death, displacement of people, and destruction of public infrastructure and physical and social capital. Two comprehensive analyzes of post-conflict situations conclude that the economic and social costs of civil wars are not only deep but also persistent, lasting for years after the end of conflict (World Bank 2003; Fosu and Collier 2005). However, when the end of war represents the beginning of lasting peace, there are good reasons to believe that recovery is possible, albeit gradual. This is what neoclassical models of economic growth and convergence would predict and what the evidence of recovery in Europe, the Republic of Korea, and Vietnam, among other examples, would seem to indicate.

Siyan Chen is a Master's student in computational finance at Carnegie Mellon University; her email address is siyanc@andrew.cmu.edu. Norman V. Loayza (corresponding author) is a lead economist at the World Bank; his email address is nloayza@worldbank.org. Marta Reynal-Querol is Ramon y Cajal Research Fellow at the Universitat Pompeu Fabra and research affiliate at the Center for Economic Policy Research (CEPR) and the Munich Society for the Promotion of Economic Research (CESifo); her email address is marta.reynal@upf.edu. The authors are grateful to Ian Bannon, Paul Collier, Ibrahim Elbadawi, Jim de Melo, and Nicholas Staines; participants at the post-conflict conferences at Oxford University, held in November 2005, and the World Bank, held in April 2007; and three anonymous referees for helpful comments.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 63–85 doi:10.1093/wber/lhn001
Advance Access Publication February 10, 2008

Most of the scarce literature on the consequences of civil war focuses on the costs during conflict; few studies analyze the costs of civil war after peace agreements are signed. Working with a cross-section of countries with well-defined pre- and post-war periods, this study uses an event-study methodology to evaluate the economic, social, and political effects in the aftermath of civil wars. Although mainly descriptive, the article provides motivation and evidence on various hypotheses concerning the consequences of internal wars. It is hoped that this study will induce more specific and more analytical research on these issues.

The article is organized as follows. Section I reviews the literature on the costs of civil wars. Section II describes the data, their sources, and the methodology used. Section III presents the results on pre-war–post-war changes and identifies trends after the war. Section IV offers some concluding remarks.

I. Review of the Literature

Wars kill people, destroy infrastructure, weaken institutions, and erode social trust. The destruction of infrastructure and institutions leaves the population living in conditions that increase the risk of disease, crime, political instability, and further conflict. The World Bank (2003) reviews the literature on the costs of civil war. The collection of articles in Fosu and Collier (2005) provides a critical analysis of post-conflict situations, with a particular focus on Sub-Saharan Africa. These articles examine the conditions and policies that support or undermine sustainable peace after conflict. Collier (1999) finds that economic growth is 2.2 percent points lower during civil wars than it is during peace. Using World Health Organization data on 23 major diseases, Ghobarah, Huth, and Russett (2003) find that civil war significantly increases the incidence of death and disability resulting from contagious diseases. Soares (2006) estimates the welfare cost of violence in a sample of countries applying a willingness-to-pay approach to account for the health consequences of war. He estimates that the civil conflict in Colombia, which reduced life expectancy at birth by 2.2 years, cost the country 9.7 percent of GDP.

Other studies focus on the effects of civil war on neighboring countries. Murdoch and Sandler (2002, 2004) show that civil wars reduce growth over an entire region of neighboring countries. Montalvo and Reynal-Querol (2007) explore the influence of refugees from civil wars on the incidence of malaria in refugee-receiving countries. They show that for every 1,000 refugees, 2,000–2,700 cases of malaria occur in the refugee-receiving country.

The empirical literature on the aftermath of civil and international war is scarcer. It seems to indicate that countries do recover in the post-conflict period, to at least their pre-war situations.

Organski and Kugler (1977, 1980) analyze the economic effects of both world wars on a sample of mainly European countries. They find that the

effects of war dissipate in the long run (15–20 years) for both losers and winners, with both types of countries usually returning to pre-war growth trends.

Barro and Sala-i-Martin (1995) explain post-war recoveries in Germany and Japan. They claim that whenever a war impairs one factor of production more than others, the rate of return of the other factors increases, creating the forces of convergence that spur rapid growth.

In a cross-country empirical analysis Przeworski and others (2000) find that post-war economic recovery is rapid. Their results indicate that the average annual rate of growth during the 5 years following a war is 5.98 percent. They also find that wars cause more damage under dictatorships than under democracies but that recoveries are more rapid under dictatorships than under democracies.

Collier and Hoeffler (2004) provide a systematic empirical analysis of aid and policy reform in the post-conflict growth process, based on a comprehensive data set of large civil wars covering 17 societies during their first decade of post-conflict economic recovery. They find that absorptive capacity during the first three post-conflict years is no greater than normal but roughly doubles over the next 7 years. They also find that growth is more sensitive to policy in post-conflict societies.

Miguel and Roland (2005) analyze the impact of the U.S. bombing of Vietnam on that country's subsequent economic development. Comparing heavily bombed districts with the rest of the country, they find that U.S. bombing did not have a lasting negative impact on poverty rates, consumption levels, infrastructure, literacy, or population density, as measured in 2002. They conclude that local recovery from the damage of war can be achieved if "certain conditions" are met.

Several studies use the methodology adopted here. Davis and Weinstein (2002) consider the Allied bombing of Japanese cities in World War II as a shock to the relative size of the cities. They find evidence of an extremely powerful recovery in the wake of the destruction, with the relative sizes of most cities returning to their pre-war levels within about 15 years.

Abadie and Gardeazabal (2003) analyze the impact of terrorism on firms in the Basque country. They find that firms with a significant presence there performed better than other firms did when the truce became credible and worse when the ceasefire ended.

Chen and Siems (2004) assess the effects of terrorism on global capital markets. They examine the U.S. capital market's response to 14 terrorist attacks since 1915 and the response of global capital markets to Iraq's invasion of Kuwait in 1990 and the September 11, 2001, terrorist attacks in the United States. They find that terrorist attacks and military invasions have great potential to affect international capital markets over a short period of time and that U.S. capital markets recover sooner than other global capital markets.

II. Data and Methodology

An event-study methodology is used to analyze the aftermath of war in a cross-section of countries. Calendar time is transformed into "event time" in order to aggregate a collection of experiences that share a particular event in common and extract meaningful statistics from them.[1]

A host of social areas is examined, measured by basic indicators of economic performance, health and education, political development, demographic trends, and conflict and security issues. Each indicator is compared before and after the war, and dynamic trends during the post-conflict period are identified.[2] The aim is to understand the nature of the recovery from war in order to document the costs of war and the extent of any peace dividend.

In exploring the patterns of behavior of various economic, social, and political variables in post-war countries, the study focuses on internal (or civil) conflicts. Information on these conflicts comes from the armed conflict data set of the International Peace Research Institute (PRIO) in Oslo. Internal and internationalized internal armed conflicts are classified as internal wars.[3] The analysis focuses on major conflicts, limiting the sample to countries with more than 1,000 battle-related deaths a year.

The study examines the following variables (defined in appendix table A-1):

(1) Economic performance, including the level and growth rate of GDP per capita, the share of domestic investment in GDP, the share of government expenditure in GDP, the share of military expenditure in government expenditure, and the inflation rate.
(2) Health and education, represented by the rates of infant mortality, adult female mortality, adult male mortality, and primary- and secondary-school enrollment.
(3) Political development, including indices of democracy and autocracy, civil liberties and political rights, and law and order.
(4) Demographic development, measured by the old-age dependency ratio, the youth dependency ratio, and the female–male ratio.
(5) Other forms of conflict, specifically the incidence of terrorist attacks.

The occurrence of a war is the event that anchors the data. The last year before the start of a war is defined as event year −1, the next-to-last year as event

1. For other presentations of this methodology, see Bruno and Easterly (1998) and Wacziarg and Welch (2003).

2. As explained later in the text, the comparative analysis takes into account unobserved country-specific effects. Moreover, it considers the experience of conflict countries both on their own and with respect to two control groups of countries.

3. According to PRIO's definitions internal armed conflict occurs between the government of a state and internal opposition groups without intervention from other states; internationalized internal armed conflict occurs when such conflict involves intervention by other states.

year −2, and so on; the first year after the end of a war is defined as event year 1, the second year as event year 2, and so on. Given the nature of the comparative exercise and the availability of data, the war years are excluded from the analysis.

The definition of the war event is crucial to the empirical evaluation. It is defined so that the pre- and post-conflict periods can be characterized as relatively free of war. At least 10 years of peace after the war are needed to ensure that the aftermath following the true resolution of an armed conflict is analyzed. This means that in cases of prolonged conflicts with temporary ceasefire periods, the "war event" includes the war, a (short) interwar peace, and the resumption of war. Where a country undergoes two wars with more than 10 years of peace in between, the wars are treated as independent events.[4]

The pre-conflict period is defined as the 7 years before the war and the post-conflict period as the 7 years after the war. One problem in applying the event-study methodology is that the sample changes across event years. Ideally, one should have a constant sample made up of the same countries for all event years. Unfortunately, for each variable, data on several countries are available only for a subset of the years under consideration. Data on per capita GDP growth, for example, may be available for the first 3 years after the war but not thereafter. Moreover, because the sample period is 1960–2003 and the analysis looks at 7 years before and after the war, conflicts had to begin no earlier than 1967 and end no later than 1996 to be included. If, in addition, the requirement of a perfectly constant sample were imposed, the sample could end up being too small.

Given these tradeoffs, the criteria for inclusion in the sample are set in the following way. For the comparison of pre- and post-war periods a country is included in the sample for a particular variable if it has at least 7 years of observations on the variable during the 7-year period before the war and the 7-year period after the war. (To be considered in the sample, the country would still have to meet the criterion of being war-free 10 years before and 10 years after the war.) For the analysis of the aftermath of conflicts the data availability restriction is imposed only on event years after the war (that is, a country need not have full pre-war data). It is likely that a country meets the requirement for one variable but fails to do so for another.

The empirical analysis studies the typical patterns of countries that experienced civil war, examining first the average difference between the pre- and post-war periods and then the average rates of change in the years after the war. The analysis considers the experience of conflict countries both on their own and with respect to two control groups of countries: the full sample of non-conflict developing countries and the subset located in the geographic

4. A concern arises when some countries experience external war during the period before or after the civil conflict. In such cases the periods around the war event cannot be characterized as peaceful. To eliminate this contamination, these countries are excluded from the samples for all variables.

region of the conflict country. As some of the variables under consideration may follow world trends (the wave of democratization in the case of political development variables, the discovery of new vaccines in the case of health indicators), comparison with the full sample of developing countries is necessary to separate these trends from the real costs of war and the merits of pacification. The comparison with respect to countries in the region is relevant because it can capture some of these trends while matching more closely the level of development of corresponding conflict countries. The main disadvantage of the regional control group is that its geographic proximity to conflict countries may make countries in the group susceptible to the effects of war.

The potential disadvantages of both control groups are reduced by the way in which the comparisons are made. For each indicator variable the control value is measured as the median for the control group in the calendar year corresponding to the event year. The difference between the conflict-country value and the control value in a given event year is then calculated for each variable. A series of differences is generated for each of the control groups.

Seventeen countries are examined in the comparison of the per capita GDP growth rate before and after the war; these countries, together with another seven on which pre-war information is not available, are used to evaluate the post-war period only. (Summary information on the country samples is presented in table S.1 in the supplemental appendix to this article, available at http://wber.oxfordjournals.org/.) Because of the lack of data in the pre-war period, three variables—military expenditure, law and order, and terrorist attacks—are examined only in the years after the war.[5]

The sample includes 41 countries (17 in Asia, 15 in Africa, 6 in Latin America, and 3 in Europe) for the period 1960–2003. Six of these countries (Cambodia, Iraq, Liberia, Myanmar, Sri Lanka, and Sudan) were entangled in two internal conflicts.

III. Results

This section reports results on two complementary exercises. First, it estimates and compares the central tendency of each variable before and after the war, both by itself and with respect to the two control groups. It also analyzes the extent to which the duration of the war affects the difference between pre- and post-war growth in per capita GDP. Second, it estimates the average rate of change of each variable during the post-conflict period both by itself and with respect to the two control groups. The analysis also examines whether the duration of the war affects the level and the rate of change of per capita GDP in the post-war period.

5. *World Development Indicators* (World Bank, 2005), for example, began collecting military expenditure data (as a percentage of central government expenditures) in 1990; the International Country Risk Guide (PRS Group, 2005) began providing ratings on law and order only after 1984.

The following regression equations are used. For the pre-war–post-war comparison:

$$y_{i,t} = \alpha_1 + \alpha_2^* \text{Post}_{i,t} + \mu_i + \varepsilon_{i,t} \tag{1}$$

$$y_{i,t} = \alpha_1 + \alpha_2^* \text{Post}_{i,t} + \alpha_3^* \text{Dur}_i^* \text{Post}_{i,t} + \mu_i + \varepsilon_{i,t} \tag{2}$$

where subscripts i and t represent country and event year; y is the variable under consideration; Post is a dummy variable that takes the value of 1 for post-war years and 0 for pre-war years; α_2, the main parameter of interest, represents the average difference in the variable y during the pre- and post-war periods; α_3 represents the effect of each additional war year on the difference between the pre- and post-war periods; μ is a country-specific effect (modeled as a country dummy); Dur is the duration of the war in number of years; and ε is the regression residual.

For the post-war exercise, the following regression equations are used:

$$y_{i,t} = \beta_1 + \beta_2^* \text{Year}_t + \mu_i + \varepsilon_{i,t} \tag{3}$$

$$y_{i,t} = \beta_1 + \beta_2^* \text{Year}_t + \beta_3^* \text{Dur}_i^* \text{Year}_t + \mu_i + \varepsilon_{i,t} \tag{4}$$

where Year indicates the event year after the war (1–7); β_2, the main coefficient of interest, represents the average change in the variable of interest from year to year in the post-war period; and β_3 represents the effect of each additional year of war on the post-war average change.

The dependent variable, y, is measured by itself and in terms of its deviation from the median of each control group. It can take three values:

$$y_{i,t} = \begin{pmatrix} y_{i,t} \\ y_{i,t} - \bar{y}_{i,t} \\ y_{i,t} - \tilde{y}_{i,t} \end{pmatrix}$$

where $\bar{y}_{i,\,t}$ represents the median of the non-conflict developing-country control group associated with country i in year t, and $\tilde{y}_{i,\,t}$ denotes the median of the non-conflict regional-country control group for the same country and year. Given the large number of variables under consideration, tables 1 and 3 report only the estimated α_2 and β_2 coefficients and associated standard errors for the three versions of each dependent variable.

Two sets of figures are provided as complements to the tables. In figure 1 are plotted the medians in each event year (7 years before the war and 7 years after) for the conflict countries and the control groups. In figure 2 are plotted the medians in each event year after the war (the sample for post-war analysis

TABLE 1. Pre-war–Post-war Changes in Economic, Social, and Political Indicators in Conflict Countries

Dependent Variable	Absolute Pre-war–Post-war Difference	Pre-war–Post-war Difference relative to Developing-Country Control Group	Pre-war–Post-war Difference relative to Regional-Country Control Group	Number of Observations/ Number of Countries
Economic				
GDP per capita	−0.150** (0.046)	−0.252** (0.044)	−0.283** (0.043)	249/18
GDP per capita growth rate	2.381* (1.299)	3.395** (1.277)	4.609** (1.537)	235/17
Investment Share	0.166 (0.477)	1.387** (0.432)	1.014* (0.589)	192/14
Government expenditure	0.957** (0.341)	0.137 (0.357)	−0.782 (0.512)	165/12
Inflation	13.048** (5.404)	10.643** (5.319)	11.673** (5.350)	176/13
Health and education				
Infant mortality	−24.311** (1.259)	−3.314** (0.924)	2.095 (1.565)	280/20
Adult female Mortality	−32.202** (5.332)	−5.129 (4.956)	6.219 (4.663)	292/21
Adult male Mortality	−32.966** (4.932)	22.962** (4.670)	10.356** (4.665)	292/21
Primary-school enrollment	13.207** (1.679)	5.814** (1.515)	3.613** (1.799)	292/21
Secondary-school enrollment	16.680** (1.235)	−5.853** (1.343)	−1.907 (1.368)	276/20
Political				
Polity2	4.437** (0.531)	−3.128** (0.469)	−1.067** (0.474)	227/17
Civil liberties and political rights	−0.853** (0.130)	0.342** (0.127)	0.073 (0.153)	165/12
Demographic				
Old-age dependency ratio	0.510** (0.067)	0.372** (0.067)	0.111 (0.071)	333/24
Youth dependency ratio	−5.321** (0.745)	0.540 (0.725)	2.940** (0.712)	333/24
Female–male ratio	0.741** (0.216)	1.025** (0.221)	1.606** (0.280)	333/24
Conflict				
Terrorist attacks	0.634 (1.421)	0.634 (1.421)	0.677 (1.410)	265/19

Note: Numbers in parentheses are standard errors.

*Significant at the 10 percent level; **significant at the 5 percent level.

Source: Authors' calculations based on data described in the text.

in figure 2 is larger than that for the pre-war–post-war comparisons in figure 1).

Pre- and Post-war Comparisons

Visual examination of the trends before and after war can be illustrative and motivate more precise statistical analyzes. Three types of behavior are

TABLE 2. Effect of Duration of War on Pre-war–Post-war Changes in the Level and Growth of per Capita GDP in Conflict Countries

Explanatory Variable	Absolute	Relative to Developing-Country Control Group	Relative to Regional-country Control Group	Number of Observations/ Number of Countries
Dependent variable: GDP per capita				
Constant	7.071** (0.032)	−0.394** (0.030)	−0.705** (0.029)	249/18
Post-war dummy variable	−0.144 (0.091)	−0.189** (0.087)	−0.136* (0.082)	249/18
Interaction term (post-war dummy variable* years of war)	−0.001 (0.007)	−0.008 (0.007)	−0.017** (0.006)	249/18
Dependent variable: GDP per capita growth rate				
Constant	2.868** (1.123)	2.010* (1.099)	2.153* (1.159)	235/17
Post-war dummy variable	2.499 (2.676)	2.502 (2.634)	4.858 (2.981)	235/17
Interaction term (post-war dummy variable* years of war)	−0.015 (0.202)	0.118 (0.200)	−0.033 (0.217)	235/17

Note: Numbers in parentheses are standard errors.

*Significant at the 10 percent level; **significant at the 5 percent level.

Source: Authors' calculations based on data described in the text.

evident (figure 1). Some variables (per capita GDP level and growth, investment share, inflation rate, Polity 2, civil and political rights, female–male ratio, and incidence of terrorism) exhibit different patterns (including different levels) before and after the war. Other variables (mortality rates, educational enrolment rates, and dependency ratios) show a change in level that seems to correspond to the continuation of a (declining or increasing) trend established before the war. Two other variables (investment rate and government expenditures) display no discernible level change.

Statistical analysis can reveal whether average or typical patterns are representative of the sample or if cross-country heterogeneity prevents any summary conclusion. For this purpose fixed-effects regressions are used to estimate and compare the means per period. Estimation through country fixed-effects allows to control for inherent country characteristics that are unrelated to the transition from war to peace.

MACROECONOMIC INDICATORS. The average level of per capita GDP is significantly lower after the war than before it, particularly in relation to the control groups (table 1). This is undoubtedly a direct reflection of the cost of war. In contrast, the average growth rate of per capita GDP in conflict countries appears to be significantly higher after than before it (by about 2.4 percent points). The increase is even more pronounced when compared with the

TABLE 3. Post-war Trends in Economic, Social and Political Indicators in Conflict Countries

Dependent Variable	Absolute Trend	Trend Relative to Developing-Country Control Group	Trend Relative to Regional-Country Control Group	Number of Observations/ Number of Countries
Economic				
GDP per capita	0.036** (0.006)	0.026** (0.006)	0.027** (0.006)	167/24
GDP per capita growth rate	0.458 (0.492)	0.349 (0.487)	0.372 (0.493)	166/24
Investment share	0.141 (0.111)	0.017 (0.100)	0.275* (0.163)	129/19
Government expenditure	−0.259** (0.092)	−0.136 (0.094)	−0.185 (0.121)	139//20
Military expenditure	−1.355** (0.415)	−0.923** (0.417)	−0.627 (0.503)	26/5
Inflation	−6.931** (2.753)	−6.391** (2.732)	−4.905* (2.770)	156/23
Health and education				
Infant mortality	−1.155** (0.197)	−0.317 (0.201)	0.151 (0.197)	195/28
Adult female mortality	−2.459** (0.538)	−0.335 (0.571)	−0.666 (0.592)	181/26
Adult male mortality	−2.038** (0.555)	−0.616 (0.525)	−0.227 (0.558)	181/26
Primary-school enrollment	2.064** (0.478)	1.592** (0.478)	1.684** (0.509)	189/27
Secondary-school enrollment	0.820** (0.213)	−1.695** (0.242)	−1.206** (0.256)	187/27
Political				
Polity2	0.059 (0.092)	−0.242** (0.114)	0.054 (0.113)	181/26
Civil liberties and political rights	−0.058** (0.023)	−0.017 (0.024)	−0.035 (0.025)	202/29
Law and order	0.176** (0.042)	0.151** (0.034)	0.111** (0.038)	104/15
Demographic				
Old-age dependency ratio	0.030** (0.015)	0.022 (0.014)	−0.016 (0.014)	202/29
Youth dependency ratio	−0.559** (0.120)	−0.124 (0.127)	0.073 (0.142)	202/29
Female–male ratio	−0.066** (0.029)	−0.041 (0.029)	−0.122** (0.039)	202/29
Conflict				
Terrorist attacks	−1.047** (0.398)	−1.047** (0.398)	−0.995** (0.396)	202/29

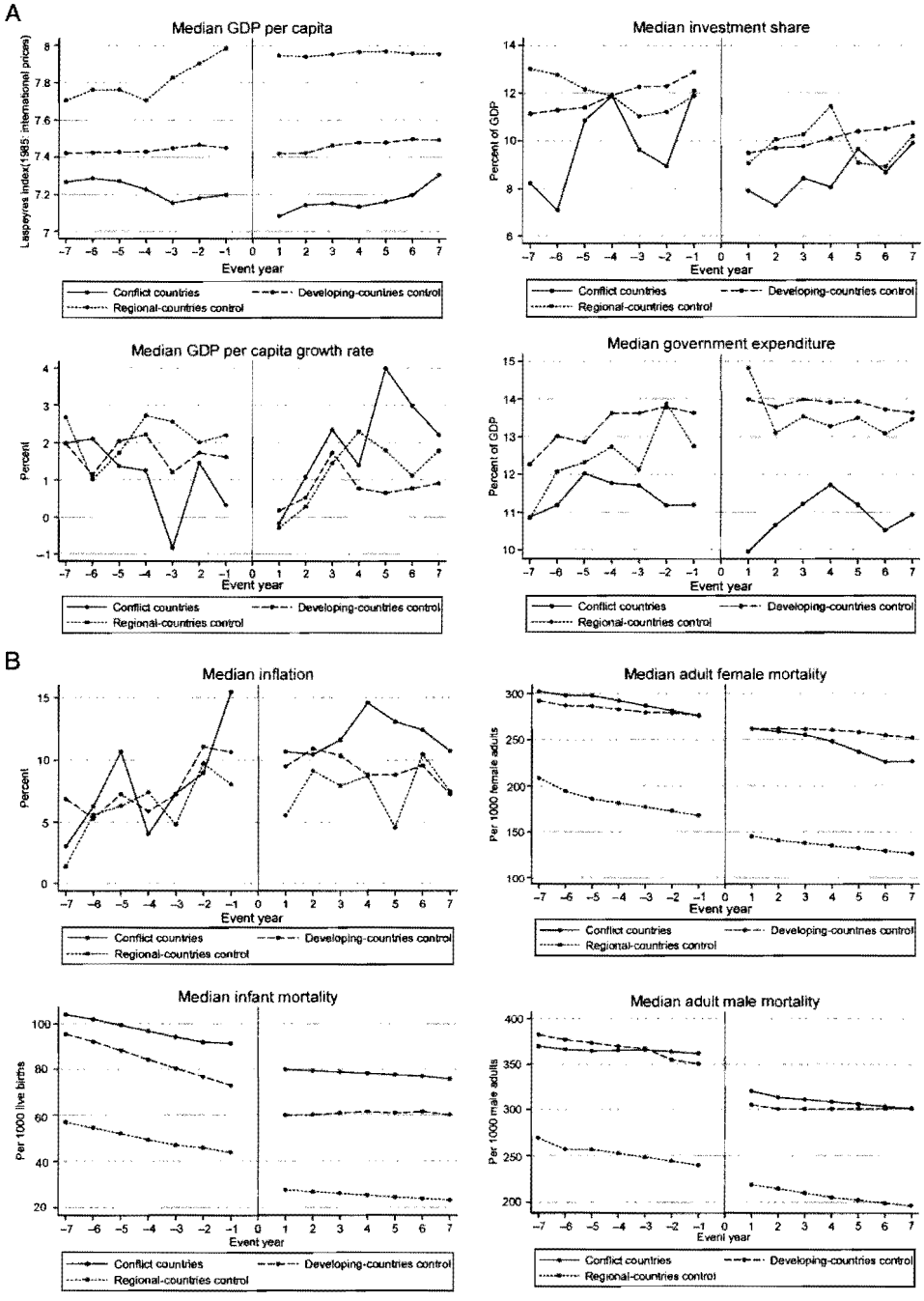
Note: Numbers in parentheses are standard errors.

*Significant at the 10 percent level; **significant at the 5 percent level.

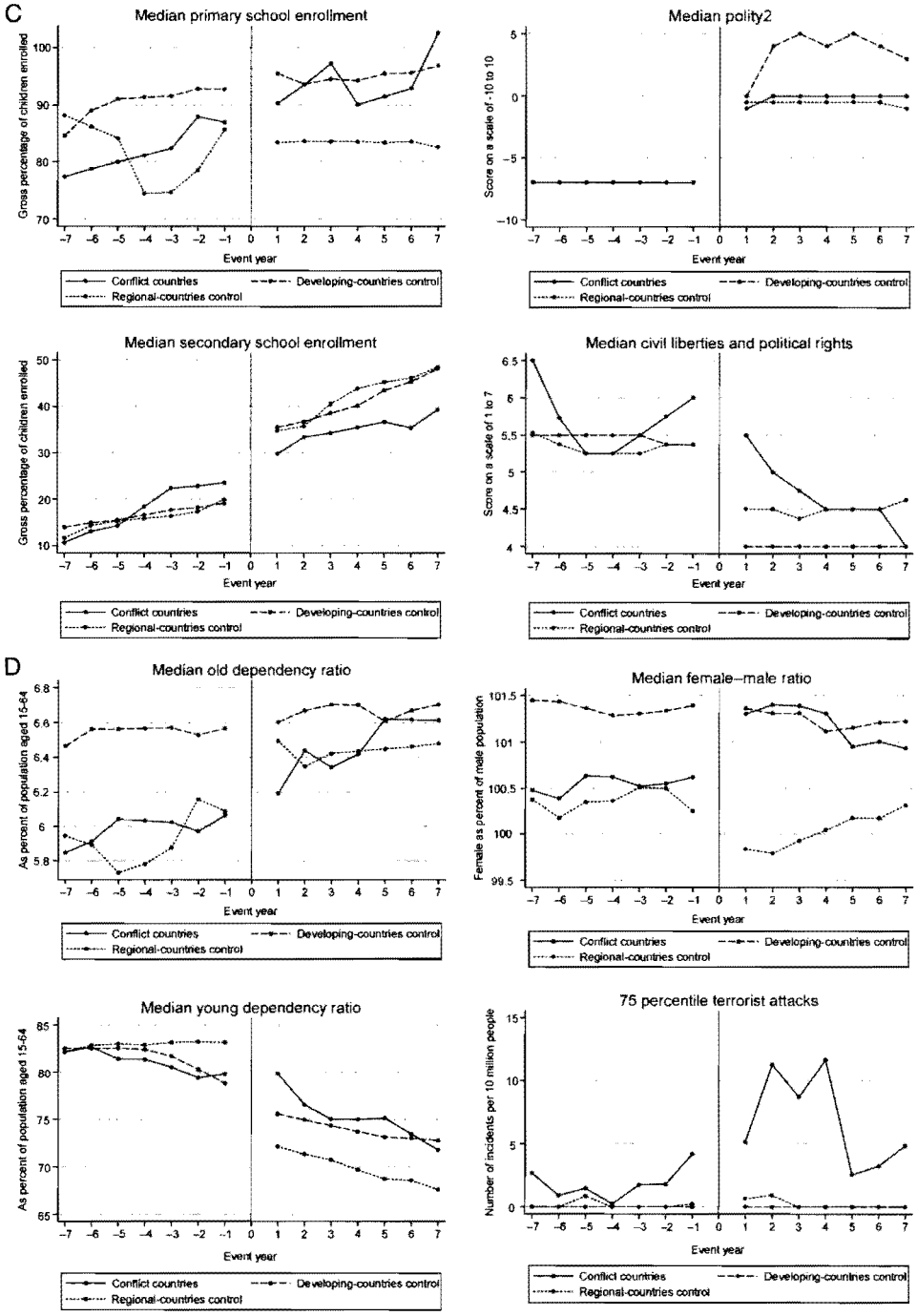
Source: Authors' calculations based on data described in the text.

change experienced by the control groups. These two results are in line with those reported by Barro and Sala-i-Martin (1995) and Przeworski and others (2000): after the destruction from war, recovery is achieved through above average growth. The increase in growth is supported by an increase in the

FIGURE 1. Economic and Social Indicators in Conflict and Non-conflict Countries



Continued



Source: Authors' analysis based on data described in the text.

investment rate. The contribution from capital accumulation, however, seems to be weak and significant only compared with the control groups, suggesting that the increase in growth also reflects a recovery in capacity utilization and, possibly, improved factor productivity.

Government expenditure (as a percentage of GDP) increases about 1 percentage point between the pre- and post-war periods and may contribute to higher growth. The change is not significantly different from that experienced in the control groups however, perhaps because the expansion in government expenditures has less potential to increase growth in countries that are not suffering the consequences of war.

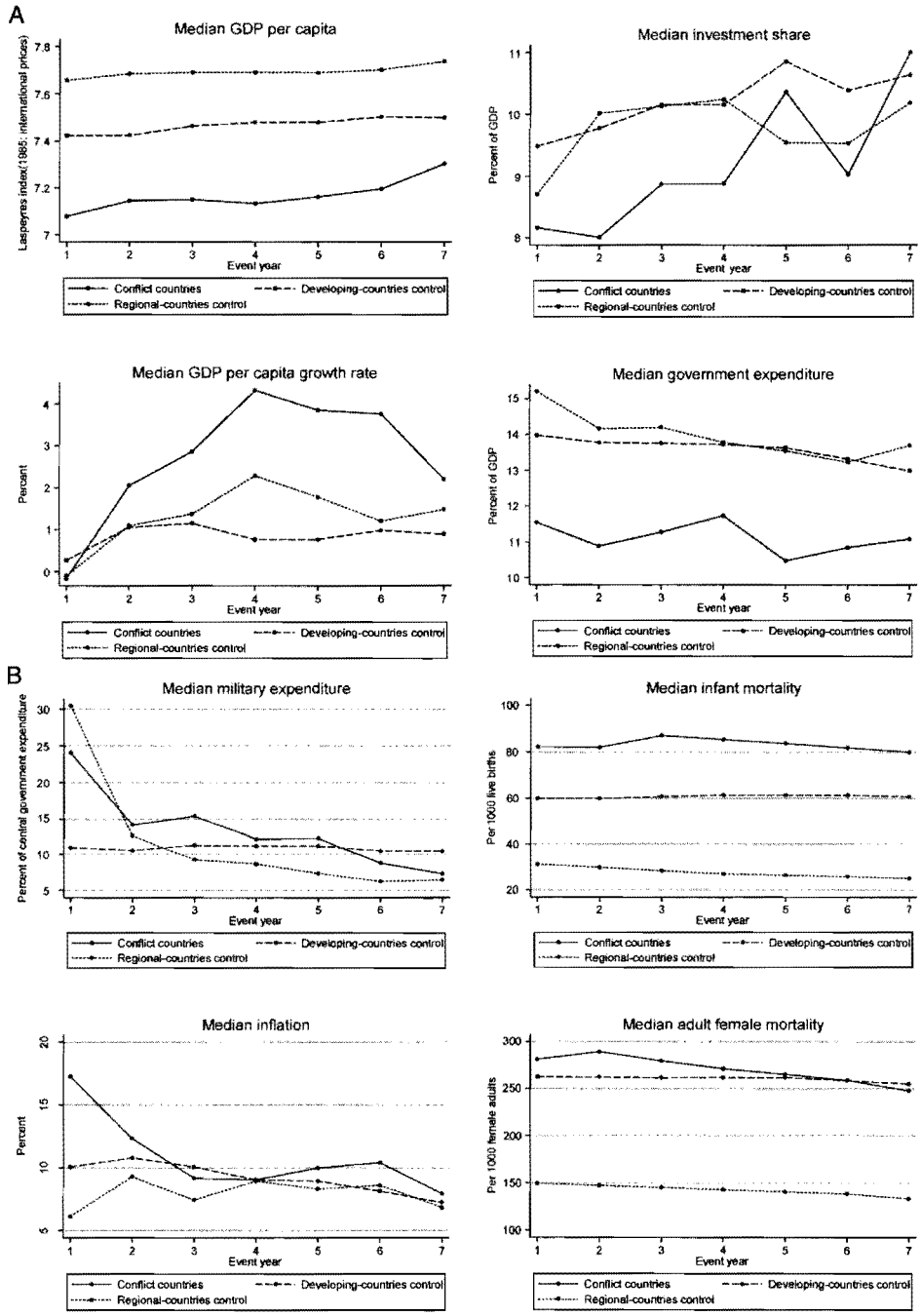
The inflation rate is significantly higher after the war, both absolutely and compared with the control groups. For the few cases for which reliable inflation data during the war are available, inflation increases sharply during the war, as government revenue sources dwindle and then decreases at the onset of peace. (The rate of inflation in the aftermath of war is discussed in the next section.)

HEALTH AND EDUCATION INDICATORS. Considered by themselves (that is, without reference to the control groups), conflict countries display marked improvement in health and education in the post-war period compared with the pre-war period. Compared with the control groups the improvements are less clear-cut. In the case of primary-school enrolment, conflict countries improve not only with respect to their pre-war level but also with respect to the control group. For the other indicators, improvement is the same as or lower than in at least one of the control groups. In the case of infant and adult female mortality, improvement in conflict countries is not significantly different from that of either control group. For adult male mortality and secondary-school enrolment—two variables related to direct combatants—improvement in conflict countries is significantly below that of the control groups. That these health and education indicators improve in absolute terms signals the important influence of world trends even for conflict-ridden countries. That the improvements fall below international standards reflects the unquestionable cost of war.

POLITICAL INDICATORS. Polity2, which measures the prevalence of democracy and the absence of autocracy, is higher after the war than before. Freedom House's Gastil measure of civil liberties and political rights (in which a smaller number represents an improvement) also indicates improvement. For both variables, however, improvement falls short of that achieved by the control groups, indicating that the cost of the war is reflected in the failure of conflict countries to achieve international standards.

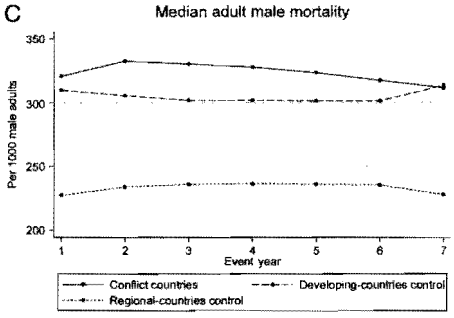
DEMOGRAPHIC INDICATORS. The old-age dependency ratio (the ratio of the number of people over 65 to the number of people in the ages 15–64) increases following war, changing in a manner similar to the demographic transition.

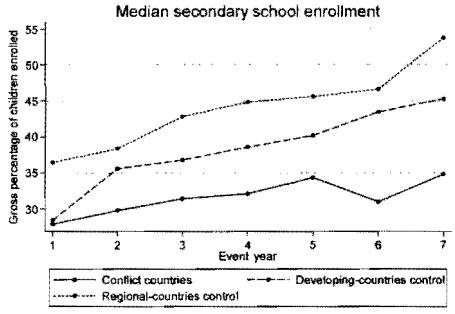
FIGURE 2. Aftermath of Conflicts

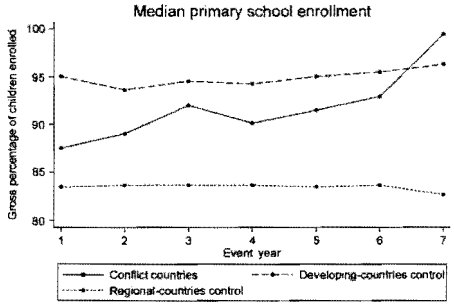


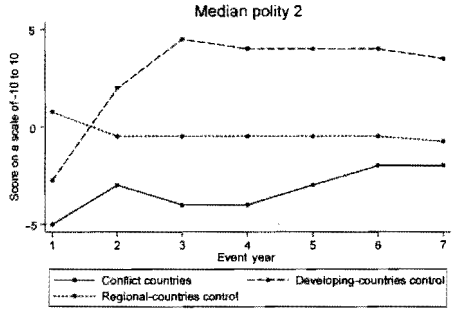
Continued

C

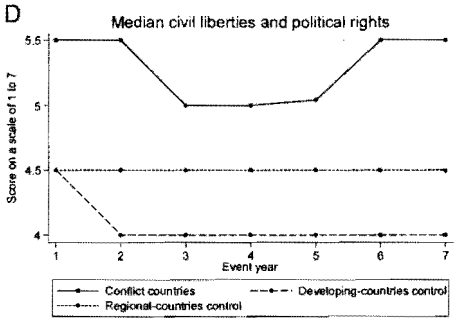


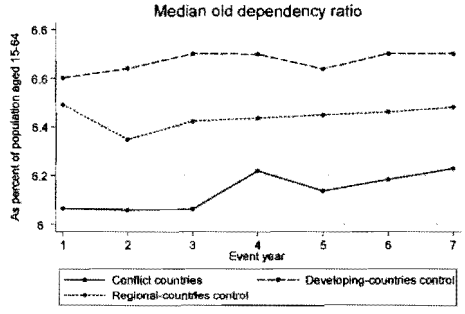


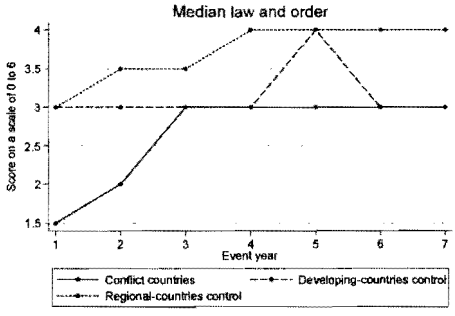


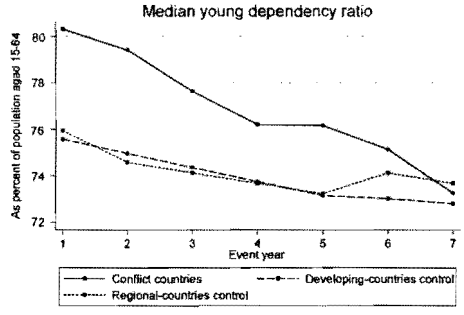


D

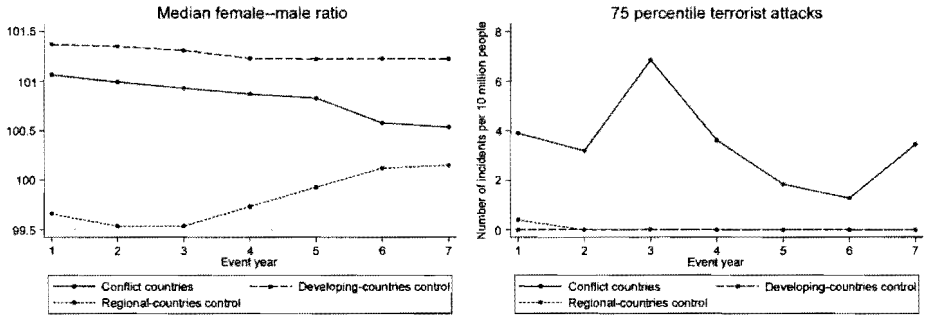








Continued



Source: Authors' analysis based on data described in the text.

The increase is more pronounced than that of the developing-country control group but not significantly different from that of countries in the same region. The youth dependency ratio (the ratio of the number of people under age 14 to the number of people in the ages 15–64) declines in absolute terms, which may also be consistent with a demographic transition. However, the youth dependency ratio increases in relative terms relative to other countries in the same region, suggesting a higher death toll among working-age adults during war.

The female–male ratio also changes in a statistically significant manner, with the ratio of women to men larger after the war than before it. This increase is even more pronounced and statistically significant when compared with the experience of either control group. The imbalance created in conflict countries in this regard is likely generated by the fact that most war fatalities are men.

CONFLICT INDICATORS. During the post-war period, the level of terrorist attacks rises over the pre-war level, but the change is not statistically significant in absolute or relative terms (largely because of the large variation across countries). (Figure 1 shows not the median for the variable, which is always 0, but the 75th percentile.)

Table 2 examines the effect of the duration of the war on the change in GDP per capita, in levels and growth rates, between the pre- and post-war periods. Regarding the level of GDP per capita, the coefficients of interest become statistically significant when the variable is expressed as deviation from the regional-country control group. The results indicate that the loss of GDP per capita as result of a major war is significant even if it is brief and that this loss increases gradually with the war?s duration. Regarding the growth rate of GDP per capita, the duration of the war does not seem to have an impact on its change between the pre- and post-war periods.

The Aftermath of War

The medians for each indicator for the conflict countries and the two control groups for each of the 7 years after the war reveal several interesting findings (figure 2). Most striking is the pattern of recovery in all dimensions after the war: in most cases the indicators show a dynamic pattern that is consistent with gradual social improvement. In the other cases improvement appears to occur early in the aftermath of internal wars. There are no clear or significant signs of worsening conditions after the onset of peace. Although recovery does not always represent progress relative to the control groups, it is nonetheless remarkable.

The average trend (or slope) of each indicator is estimated for the sample of conflict countries (table 3). A fixed-effects estimator is used to allow for different intercepts per country. (To save space, the table presents only the slope coefficients for each variable of interest, specified in absolute terms and as deviations from the control groups.)

MACROECONOMIC GROWTH INDICATORS. Per capita GDP in conflict countries has a significantly positive trend that is larger than that of any of the control groups. This gradual improvement reflects the higher levels of GDP growth in conflict countries after war. Per capita GDP growth shows no significant linear trend; its pattern appears to follow an inverted *U*, with the strongest results achieved in the fourth or fifth year after the onset of peace. The investment rate shows a positive slope, but it is statistically significant only when compared with the regional-country control group. The average investment rate in conflict countries is initially lower than in the regional-country control group, but the two rates converge within a decade of the end of the war. Conflict countries thus appear to raise their per capita GDP to the average levels in their region partly through higher investment rates in the years after the war.

Public finances also changes in the aftermath of civil wars. Government expenditure (as a percentage of GDP) reveals a declining trend that is statistically significant in absolute terms but not relative to the control groups. Military expenditure (as a percentage of government expenditure) reveals a clear and significant declining trend in the aftermath of war, both absolutely and with respect to the developing-country control group. It does not hold when conflict countries are compared with the regional-control groups, possibly because the threat of a civil war becoming an international war may induce countries in the region to increase their military expenditure during the war and decrease it afterward. In brief, after peace is achieved conflict countries gradually reduce their government expenditures and sharply de-emphasize the importance of military expenditure in the use of fiscal resources.[6]

6. Notable is the contrast between sustainable peace and insecure post-conflict. Collier and Hoeffler (2006) investigate the effects of post-conflict military spending on the risk of resumed hostilities. They find that high military spending significantly increases the risk of renewed conflict.

HEALTH AND EDUCATION INDICATORS. Health and education indicators display a statistically significant improving trend over time in absolute terms. Regarding relative improvement, the average recovery rate for primary-school enrolment is larger in conflict countries than in either of the control groups. In contrast, for secondary-school enrolment, conflict countries underperform compared with both control groups. The average rate of improvement in infant and adult female and male mortality rates is not different from that of at least one of the control groups.

POLITICAL INDICATORS. There are some signs of absolute improvement in the political indicators, as measured by the democracy index of Polity2 (positive slope) and the Gastil civil liberties index (negative slope). Only the Gastil civil liberties index is statistically significant, however. Conflict countries do not perform better than the control groups however, especially countries in the same region. In contrast, the International Country Risk Guide Index on law and order shows a marked and significant rate of progress in conflict countries, both in absolute terms and in comparison with the control groups. This result suggests that in the aftermath of civil war, when political rights are slow to advance, police and judicial systems improve at an accelerated rate.

DEMOGRAPHIC INDICATORS. The demographic transition continues in conflict countries in the aftermath of war. In absolute terms, the old-age dependency ratio rises and the youth dependency ratio falls. Relative to both control groups, there is no discernible trend in either ratio, indicating that the pattern of demographic transition in post-conflict countries is the same as in otherwise similar countries. In contrast, the female–male ratio, which rises during the war, exhibits a statistically significant declining trend in its aftermath, both in absolute terms and relative to the regional-country control group. This trend reflects the gradual recovery of the male population from its losses during the war.

CONFLICT INDICATOR. The incidence of terrorist attacks decreases significantly in the aftermath of civil wars, in both absolute terms and in relation to both control groups.[7] When more complex, nonlinear behavior is allowed (not shown in the table), terrorist attacks seem to follow a quadratic trend, with some increase early in the aftermath of war and a marked decline subsequently. The end of the civil war appears to eventually lead to pacification of other types of internal strife.

Table 4 examines whether the duration of the war has an impact on the speed of post-conflict recovery. As previously, this is studied only for the case

7. The terrorism data come from the ITERATE project (Mickolus and others 2004). These data largely cover incidents of terrorism with a transnational component. They may therefore imperfectly reflect the domestic nature of terrorist attacks that characterizes post-conflict situations.

TABLE 4. Effect of Duration of War on the Post-war Trends in the Level and Growth of per Capita GDP in Conflict Countries

Variable	Absolute	Relative to Developing-Country Control Group	Relative to Regional-country Control group	Number of Observations/ Countries
Dependent variable: GDP per capita				
Constant	6.991** (0.025)	−0.439** (0.024)	−0.716** (0.027)	167/24
Post-war trend	0.056** (0.011)	0.042** (0.011)	0.047** (0.012)	167/24
Interaction term (post-war trend * years of war)	−0.002** (0.001)	−0.002** (0.001)	−0.002** (0.001)	167/24
Dependent variable: growth of GDP per capita				
Constant	1.730 (2.428)	1.186 (2.388)	1.121 (2.461)	166/24
Post-war trend	1.552* (0.912)	1.574* (0.894)	1.292 (0.936)	166/24
Interaction term (post-war trend * years of war)	−0.125* (0.067)	−0.141** (0.066)	−0.106 (0.068)	166/24

Note: Numbers in parentheses are standard errors.
*Significant at the 10 percent level; **significant at the 5 percent level.
Source: Authors' calculations based on data described in the text.

of GDP per capita. Regarding its level, the results indicate that GDP per capita has a positive trend in the aftermath of conflict whose slope is diminished with the duration of the war. This is significantly so for the comparisons in absolute and relative terms. Regarding economic growth, the result is qualitatively similar: the growth rate of GDP per capita has a positive trend which declines as the duration of the war is larger. This is true in absolute terms and in relation to the developing-country control group (the pattern of signs is the same in the comparison to the regional-country control group but the statistical significance is weaker). In brief, the cost of war is here manifested in the negative effect which its duration has on the level and growth of GDP per capita.

IV. CONCLUSIONS

War has devastating effects, and its aftermath can be immensely difficult. However, when the end of war marks the beginning of lasting peace, recovery and improvement are feasible.

The cost of war is reflected in the substantial drop in per capita income suffered by conflict countries during war and in their failure to make as rapid progress in key areas of political development (such as civil liberties and democratic rule) and some aspects of health and educational achievement closely related to combatants (such as adult male mortality and secondary-school enrolment) as countries that did not experience civil war. In other, more basic areas of social development (such as infant mortality and primary-school

enrolment), conflict countries have been able to participate in international progress, despite the war. This is arguably a testament to the beneficial impact of medical innovations, educational programs, and the international campaigns to promote them.

The problems associated with war do not start when the fighting begins. They would have been present before and may have precipitated or generated the civil conflict. It therefore stands to reason that the resolution of war, when it promotes enduring peace, may signal the start of the solution to these problems. The behavior of economic growth provides evidence of this notion: economic growth is low (or negative) before the war. After the war, growth is strongly positive, with an average rate that is 2.4 percent points higher than that before the war.

The aftermath of war is a period of recovery. Virtually all aspects of economic, social, and political development experience gradual improvement in absolute terms. Recovery occurs swiftly in macroeconomic areas: output per capita increases, capital investment rises, and inflation decreases at rates that are high enough that conflict countries gradually converge with otherwise comparable non-conflict countries.

Some social and political indicators also display this pattern of relative improvement (and thus convergence). Measures of primary-school enrolment, demographic imbalances, the rule of law, and the incidence of terrorist attacks improve more rapidly than in other developing countries. This progress is accompanied by a continuous reallocation of public resources away from military expenditure. In contrast, indicators directly related to victims, combatants, and political processes (such as mortality, secondary-school enrolment, and democratic rights) improve no more rapidly (and often less rapidly) than otherwise similar non-conflict countries.

Several lessons can be learned from the behavior of social and political variables in post-conflict situations. Democratic rights, for example, are slow to advance and may require the foundation of long-run institutions to be consolidated. In contrast, the perception of law and order can be improved rapidly by a variety of strong government regimes. Even then, pacification after civil war does not occur overnight: terrorist attacks can be pervasive in the years immediately following the cessation of hostilities, although this trend tends to subside over time, giving way to a true resolution of the war.

These conclusions are based on the responses of typical conflict countries. This article has also attempted to account for the heterogeneity across conflict countries—both in the change between pre- and post-war situations and in the rate of recovery in the aftermath of war—by assessing the effects of the duration of armed civil conflict. The results indicate that the decline in per capita GDP during the war is greater the longer the war persists; the rate of increase and even acceleration of per capita GDP in the aftermath of war declines significantly with the length of conflict.

The article's shortcomings suggest a rich agenda for future research. That agenda should include a deeper analysis of the heterogeneity in the recovery patterns of conflict countries, examination of the causal mechanisms underlying these patterns, and an evaluation of policies for successful post-conflict recovery, including demobilization of former combatants, external intervention and aid, domestic redistributive programs, and institutional reform.

Supplementary Material

Supplementary Material is available at The World Bank Economic Review Online.

Funding

M.R.-Q. acknowledges financial support from the Spanish Ministerio de Educación (grant SEJ2006-10974/ECON).

Appendix

Table A-1. Definitions and Sources of Variables

Variable	Definition	Source
Internal/external wars	Conflicts resulting in more than 1,000 battle-related deaths a year during every year in the period	International Peace Research Institute (PRIO), Oslo Freedom House, 2005
GDP per capita	Real GDP per capita	Authors' calculation based on data from Penn World Tables 5.6 and *World Development Indicators* (World Bank 2005)
GDP per capita growth rate	Real GDP per capita growth rate (percent)	Authors' calculation with data from Penn World Tables 5.6 and *World Development Indicators* (World Bank), 2005
Investment share	Investment share of real GDP per capita (percent of GDP)	Penn World Tables 6.1
Government expenditure	General government expenditure on final consumption (percent of GDP)	*World Development Indicators* (World Bank 2005)
Inflation	GDP deflator (annual percent)	*World Development Indicators* (World Bank 2005)
Military expenditure	Percent of central government expenditure	*World Development Indicators* (World Bank 2005)
Infant mortality	Number of infant deaths per 1,000 live births	*World Development Indicators* (World Bank 2005)
Female mortality	Number of adult female deaths per 1,000 female adults	*World Development Indicators* (World Bank 2005)

(Continued)

TABLE A-1. Continued

Variable	Definition	Source
Male mortality	Number of adult male deaths per 1,000 male adults	*World Development Indicators* (World Bank 2005)
Primary school enrollment	Gross percentage of children enrolled in primary school	*World Development Indicators* (World Bank 2005); Barro and Lee (1994) data set
Secondary school enrollment	Gross percentage of children enrolled in secondary school	*World Development Indicators* (World Bank 2005); Barro and Lee (1994) data set
Polity2	Combined Polity score (computed by subtracting autocracy score from democracy score); additive 21-point scale (−10 to 10), on which 10 represents the highest degree of democracy and −10 the lowest	Polity IV (2005)
Civil liberties and political rights	Sum of political rights and civil liberties divided by two; countries whose combined average ratings are 1.0–3.0 are designated "free," countries with rankings of 3.0–5.5 are designated "partly free," countries with ranking of 5.5–7.0 designated "not free"	Freedom House (2005)
Law and order	Measured on 0–6 scale, on which 6 represents the highest quality of law and order and 0 the lowest	International Country Risk Guide (PRS Group 2005); monthly data for June are selected to represent the whole year
Old-age dependency ratio	Population over 65 divided by population 15–64, stated as percent	Authors' calculation from *World Development Indicators* (World Bank)
Youth dependency ratio	Population under 14 divided by population 15–64, stated as percent	Authors' calculation from *World Development Indicators* (World Bank)
Female–male ratio	Female population divided by male population, stated as percent	Authors' calculation from *World Development Indicators* (World Bank)
Terrorism	Number of terrorism incidents per 10 million people. A terrorism incident is ascribed to the country where it occurs or, in the case of a hijacking, where it starts.	ITERATE (Mickolus and others 2004)

Source: Authors' compilation from sources indicated.

REFERENCES

Abadie, A., and J. Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93(1):113–32.

Barro, Robert, and Jong-Wha Lee. 1994. *Data Set for a Panel of 138 Countries.* Cambridge, MA: Harvard University, Center for International Development.

Barro, R., and X. Sala-i-Martin. 1995. *Economic Growth.* New York: McGraw Hill.

Bruno, M., and W. Easterly. 1998. "Inflation Crises and Long-Run Growth." *Journal of Monetary Economics* 41(1):3–26.

Center for International Development and Conflict Management. 2005. *Polity IV Data Set.* College Park, MD: Center for International Development and Conflict Management.

Chen, A., and T. Siems. 2004. "The Effects of Terrorism on Global Capital Markets." *European Journal of Political Economy* 20(2):349–66.

Collier, P. 1999. "On the Economic Consequences of Civil War." *Oxford Economic Papers* 51(1):168–83.

Collier, P., and A. Hoeffler. 2004. "Aid, Policy and Growth in Post-Conflict Countries." *European Economic Review* 48(5):1125–45.

———. 2006. "Military Expenditure in Post-Conflict Societies." *Economics of Governance* 7(1):89–107.

Davis, D., and D. Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review* 92(5):1269–89.

House, Freedom. 2005. *Freedom in the World Country Ratings.* Washington, D.C.: Freedom House.

Fosu, A.K., and P. Collier, eds. 2005. *Post-Conflict Economies in Africa.* New York: Palgrave Macmillan.

Ghobarah, H., P. Huth, and B. Russett. 2003. "Civil Wars Kill and Maim People Long after the Shooting Stops." *American Political Science Review* 97(2):189–202.

Mickolus, E., T. Sandler, J. Murdock, and P. Flemming. 2004. *International Terrorism: Attributes of Terrorist Events, 1968–2003 (ITERATE).* Dunn Loring, Va.: Vinyard Software.

Miguel, Edward, and Gerard Roland. 2005. "The Long-Run Impact of Bombing Vietnam." Berkeley: Press release, University of California Press.

Montalvo, J.G., and M. Reynal-Querol. 2007. "Fighting against Malaria: Prevent Wars While Waiting for the 'Miraculous' Vaccine." *Review of Economics and Statistics* 89(1):165–77.

Murdoch, J., and T. Sandler. 2002. "Civil Wars and Economic Growth: A Regional Comparison." *Defense and Peace Economics* 13(6):451–64.

———. 2004. "Civil Wars and Economic Growth: Spatial Dispersion." *American Journal of Political Science* 48(1):138–51.

Organski, A.F.K., and Jacek Kugler. 1977. "The Costs of Major Wars: The Phoenix Factor." *American Political Science Review* 71(14):1347–66.

———. 1980. *The War Ledger.* Chicago, IL: University of Chicago Press.

Przeworski, A., M. Alvarez, J. Cheibub, and F. Limongi. 2000. *Democracy and Development: Political Institutions and Wellbeing in the World 1950–1990.* Cambridge, United Kingdom: Cambridge University Press.

Group PRS (Political Risk Services Group). 2005. *International Country Risk Guide.* East Syracuse, N.J.: PRS Group (Political Risk Services Group).

Soares, R. 2006. "The Welfare Cost of Violence across Countries." *Journal of Health Economics* 25(5):821–46.

Wacziarg, R., and K. Welch. 2003. *Trade Liberalization and Growth: New Evidence.* NBER Working Paper 10152. Cambridge, MA: National Bureau of Economic Research.

World Bank. 2003. *Breaking the Conflict Trap: Civil War and Development Policy.* New York: Oxford University Press.

———. 2005. *World Development Indicators.* Washington, D.C.: World Bank.

Postconflict Monetary Reconstruction

Christopher Adam, Paul Collier, and Victor A.B. Davies

During civil wars governments typically resort to inflation to raise revenue. A model of this phenomenon is presented, estimated, and applied to the choices and constraints faced during the postconflict period. The results show that far from there being a fiscal peace dividend, postconflict governments tend to face even more pressing needs after than during war. As a result, in the absence of postconflict aid, inflation increases sharply, frustrating a more general monetary recovery. Aid decisively transforms the path of monetary variables in the postconflict period, enabling the economy to regain peacetime characteristics. Postconflict aid thus achieves a monetary "reconstruction" analogous to its more evident role in infrastructure. JEL codes: H56, F35, O10

War is expensive and so has powerful economic consequences. Civil war—now by far the most common form of war—is particularly damaging, reducing income, increasing capital flight, and diverting activity into subsistence. All of these effects can be expected to reduce the demand for money. The resulting decline in seigniorage revenue collides with increased government fiscal needs, with both effects tending to raise inflation. A likely economic legacy of war is thus a deterioration in the tradeoff between seigniorage and inflation. Just as the postwar government faces a hard choice between continued military spending and the reconstruction of infrastructure, so too it faces a choice between continued inflation and the "reconstruction" of the monetary base.

Section I of this article sets out the decision problems facing households and governments. The demand for money on the part of households is the constraint against which the government maximizes. Section II applies the model to the data. The expansion in data on civil wars has recently made it a researchable phenomenon using standard quantitative techniques (Miguel, Satyanath, and Sergenti 2004; Collier and Hoeffler 2004b). Estimates are made

Christopher Adam (corresponding author) is a research associate at the Centre for the Study of African Economies, Department of Economics, University of Oxford; his email address is christopher.adam@economics.ox.ac.uk. Paul Collier is the director of the Centre for the Study of African Economies; his email address is paul.collier@economics.ox.ac.uk. Victor A.B. Davies is a senior research economist at the African Development Bank and a research associate at the Centre for the Study of African Economies; his email address is v.davies@afdb.org.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 87–112 doi:10.1093/wber/lhm020
Advance Access Publication January 14, 2008

of both how money demand is affected by civil war and its aftermath and how the revealed preferences of governments change when governments are faced with new constraints and needs. Potentially, the harsh policy tradeoffs that governments face in postconflict situations can be alleviated by aid—indeed, this is the context for which aid was invented. Section III introduces aid into the analysis, estimating how it affects constraints and choices and showing the path of inflation, money demand, and seigniorage with and without postconflict aid. Section IV discusses the policy implications of the results.

I. The Government Decision

Assume that before civil war, the government conducts its monetary policy on a sustainable basis. This may well involve the choice of a positive rate of inflation; it does not, however, involve the government attempting to fool private agents by delivering more inflation than they expect. The government is not assumed necessarily to try to maximize social welfare. The actual choice of inflation will depend on how costly it is to the government relative to other sources of revenue, where the costs taken into account by the government may differ from those that concern society.

Civil war takes the society and the government by surprise. This is a reasonable characterization because, although civil wars can to some extent be expected, no model has been able to predict the actual outbreak of civil war with any certainty: the main news occurs around the outbreak.

Consider the effect of the outbreak of civil war on private agents. Civil war reduces GDP growth. A typical estimate of the economic loss is that growth is reduced by about 2.3 percent over a period of seven years (Collier 1999). Heightened insecurity tends to divert economic activity toward relatively sheltered sectors, notably subsistence, and agents attempt to protect assets through capital flight. In the postconflict decade the economy usually recovers, but slowly, with GDP typically growing about 1.1 percent more than normal (Collier and Hoeffler 2004a). Hence, for a prolonged period, the demand for money is likely to be reduced both directly, as a result of the fall in income, and indirectly, as a result of activity and asset substitution.

The decline in the demand for money exacerbates the seigniorage-inflation tradeoff facing the government. However, the government will want to increase its spending for the duration of the conflict. Military spending typically increases by nearly 2 percent of GDP during civil war (Collier and Hoeffler 2007). This need for increased military spending raises the government discount rate. Because borrowing is difficult during civil war, the government chooses a higher rate of inflation.

This government's problem can be set out more formally using a simple model in which a forward-looking government chooses how much conflict-related expenditure is to be financed at the margin through seigniorage. The model, built around a simple Cagan (1956) characterization of the private

sector's demand for money, is similar to that found in Bruno and Fischer (1990), Adam (1995), and Marcet and Nicolini (2003).

A number of simplifying assumptions are made to sharpen the exposition. First, and least important, private income and other sources of financing are held constant, except for the direct changes caused by war itself. The discussion thus abstracts from broader questions of the optimal fiscal response to expenditure shocks (see, for example, Mankiw 1987; Cashin, Ul Haque, and Olekalns 2002).[1]

Second, a simple monetarist framework is assumed, in which the authorities' monetary instrument is the volume of nominal base money, which is the only domestic financial liability. Money is held by both the bank and nonbank private sectors. In the empirical analysis seigniorage earned on currency in circulation is distinguished from seigniorage earned on bank reserves; without loss of generality the two are combined in the model presented in this section.

Third, the private sector's inflation expectations are assumed to be formed adaptively, albeit in a manner consistent with learning. Given the context, this has an intuitive appeal, because private agents could be expected to respond with a lag, possibly a very short one, to conflict-related changes in public expenditure. Employing an adaptive expectations framework has other merits. Specifically, given the Cagan-form money demand function, inflation equilibria on the "good side" of the seigniorage Laffer curve are dynamically stable under the assumption of adaptive expectations, whereas those on or above the top of the Laffer curve are unstable. As Bruno and Fischer (1990) show, the opposite holds under rational expectations, a feature that gives rise to the "high inflation trap" analyzed in their article.[2] Given that the analysis here starts from a position at which the economy is in an initial equilibrium on the "good side" of the Laffer curve, it makes sense that this initial equilibrium is dynamically stable.

Finally, time-inconsistency problems are assumed away, in the strict sense that the initial and any subsequent long-run inflation equilibria are credible.

The Model

Government preferences are defined as

$$V = V(g(\pi_t), k(\pi_t)) \tag{1}$$

where g denotes government expenditure, π_t the inflation rate at time t, and $k(\pi_t)$ the discounted future costs of current inflation distinct from the inflation-tax distortion on the demand for money. For example, $k(\pi_t)$ could reflect the reduction in investment efficiency associated with higher inflation. The model

1. Equivalently, the government's problem can be characterized in terms of the change in expenditure requirements net of other financing items.

2. This property of adaptive expectations is replicated under rational expectations if there is lagged adjustment of money demand.

assumes that $V_g > 0$, $V_k < 0$, and $k'(\pi_t) > 0$. Both $g(.)$ and $k(.)$ are measured as shares of GDP.

The government's period budget constraint in nominal terms is given by

$$G_t = \Delta M_t + A_t + T_t \tag{2}$$

where M_t denotes the nominal base money, A_t the domestic value of aid inflows, and T_t conventional tax revenues. Dividing through by nominal GDP, $Y_t = P_t\, y_t$, allows equation (2) to be expressed as

$$g_t - a_t - \tau_t = \Delta m_t + \left(\frac{\pi_t}{1 + \pi_t}\right) m_{t-1} \tag{3}$$

where g_t denotes the real value of government expenditure, a_t real aid, τ_t real conventional taxation, while the terms on the right side denote total seigniorage (consisting of the growth in real money balances plus the inflation tax).

Aid and tax revenue are treated as fixed, so that at the margin, changes in government expenditure are financed by changes in domestic deficit financing. The private sector's demand for money is characterized by a Cagan money demand function of the form

$$m_t = c_t y_t \exp(-\alpha \tilde{\pi}_t^e) \tag{4}$$

where c denotes a constant that may shift over time (in response to the onset or cessation of conflict, for example), π_t^e denotes expected inflation, and $\tilde{\pi}_t^e = \pi_t^e / (1 + \pi_t^e)$.[3] Defined in this manner, the inflation term, $\tilde{\pi}_t$, is bounded above by 1 as the conventional measure of inflation becomes arbitrarily large, giving it a natural interpretation as a tax rate at which a value of 1 implies complete confiscation.

The private sector adjusts its inflation expectations, defined in terms of the inflation factor $\tilde{\pi}_t^e$, in response to the deviation of actual inflation from the level anticipated in the previous period:

$$\dot{\tilde{\pi}}_t^e = \beta_t(\pi_t - \pi_t^e), \tag{5}$$

where a dot ($\dot{}$) denotes the derivative with respect to time and $0 < \beta_t < 1$ measures the speed of adjustment, which could vary over time, as a result of learning, for example. See Marcet and Nicolini (2003) for a discussion of alternative learning algorithms. For the most part, β_t is assumed to equal β.

3. Calvo and Leiderman (1992) show that under specific restrictions on functional form, equation (4) derives directly from the dynamic first-order condition for a representative agent maximizing utility of the form $U = \int_{t=0}^{\infty} [u(c_t) + v(m_t)] e^{-\rho t}\, dt$.

Equilibrium

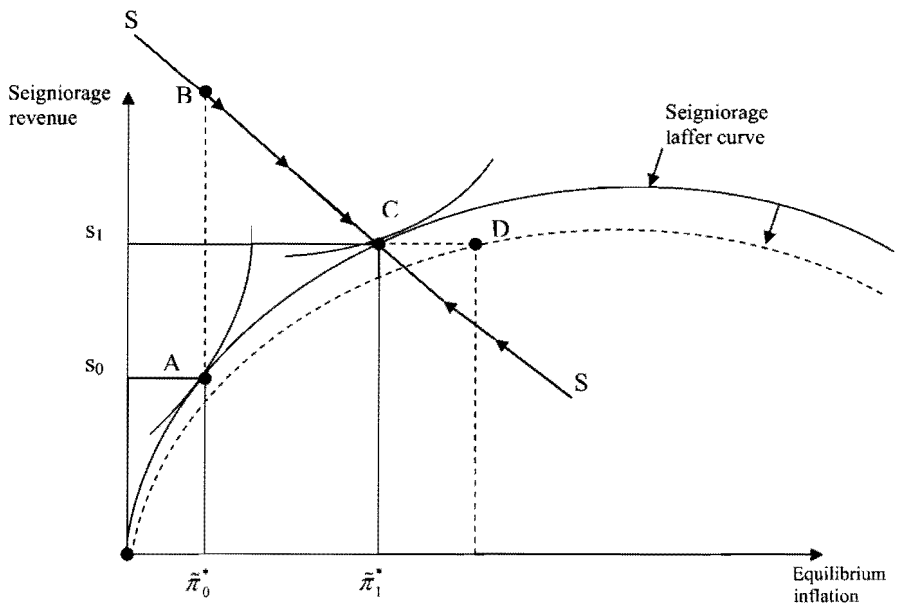
The government's problem is to maximize equation (1) subject to equations (2) and (4). Given the assumption that the government can credibly commit to a given inflation rate, in equilibrium inflation expectations are correct. Assuming no growth in real income, this implies a constant rate of inflation $\tilde{\pi}^e_{t+1} = \tilde{\pi}_{t+1} = \tilde{\pi}_t$ and hence a constant growth rate of the money supply. From equation (5) it follows that $\dot{\tilde{\pi}}^e_{t+1} = 0$. In these circumstances the first-order condition with respect to inflation is given by

$$\frac{-V_k k'(\tilde{\pi}_t)}{V_g} = m_t(1 - \alpha\tilde{\pi}_t) \tag{6}$$

The solution to equation (6) defines the optimal (constant) inflation rate, $\tilde{\pi}_t{}^* = \tilde{\pi}_0{}^*$, and hence the optimal rate of growth of the money supply. Substituting equation (6) into equations (4) and (3) yields optimal seigniorage, shown as point A in figure 1.

The right side of equation (6) is simply the slope of the seigniorage Laffer curve defined by the demand for money, equation (4). The seigniorage revenue–maximizing point is attained at $\tilde{\pi}^{\max} = 1/\alpha - \hat{y}$, where $\hat{y}$ denotes the growth rate of real income. The left side of equation (6) is the slope of the government's indifference curve, measuring the rate at which the government trades off present government consumption against future damage to the economy. This can be thought of as a quasi-discount rate. It follows from

FIGURE 1. Inflation and Seigniorage Revenue during Conflict



equation (6) that

$$\frac{\partial \tilde{\pi}_t^*}{\partial V_g} > 0 \text{ and } \frac{\partial \tilde{\pi}_t^*}{\partial V_k} < 0. \tag{7}$$

Dynamics

The short-run dynamics of the model in response to anticipated and unanticipated changes in the fiscal deficit emerge directly from equations (3)–(5). The model assumes that in the short run, c_t in equation (4) is constant. Taking the log derivative with respect to time and substituting yields

$$\frac{d \ln m_t}{dt} = \hat{m}_t = \hat{y}_t - \alpha \dot{\tilde{\pi}}_t^e, \tag{8}$$

where a hat (^) represents a proportionate change. Using the definition $\hat{M}_t = (\hat{m}_t + \pi_t + \hat{y}_t)$ and denoting the growth in the nominal money supply by $\sigma_t = \hat{M}_t$ allows inflation to be expressed as

$$\pi_t = \sigma_t + \alpha \dot{\tilde{\pi}}_t^e - \hat{y}_t \tag{9}$$

Substituting equation (9) into equation (5) leads to the following differential equation for inflation expectations:

$$\dot{\tilde{\pi}}_t^e = \left(\frac{\beta}{1 - \alpha\beta}\right)(\sigma_t - \pi_t^e - \hat{y}_t). \tag{10}$$

When inflation expectations adjust sufficiently slowly, such that $\beta < 1/\alpha$, equation (10) is dynamically stable and the economy's adjustment to an increase in the nominal growth of the money supply is denoted by the saddle path SS in figure 1.

Responses During Conflict

During conflict the government faces increased pressure to spend in order to confront its opponents. At the same time, however—and not necessarily independently—private sector demand for money declines.

Consider first the government's choice, assuming for the moment no change in the private sector's demand for money. Additional expenditure needs temporarily increase the marginal utility of government consumption, V_g. This implies an increase in the quasi-discount rate for the duration of the conflict, and, from equation (7), a higher optimal rate of inflation, chosen in order to generate a higher rate of seigniorage. Thus, the government seeks to move along the Laffer curve. It does so by increasing the growth of the money supply, which generates an initial jump in seigniorage to point B. From equation (5), expected inflation

rises, inducing a decline in real money balances. This continues along SS to the new equilibrium at C, at which point inflation expectations have fully adjusted, such that $\tilde{\pi}_t^e = \tilde{\pi}_1{}^*$. The initial real resource flow at B occurs regardless of how rapidly inflation expectations subsequently adjust, because the increased growth in the money supply is exchanged for real private resources at the price level prevailing in the initial equilibrium. How much additional transitional seigniorage revenue accrues thereafter depends on the private sector's speed of adjustment: the more slowly inflation expectations adjust to the new rate of growth of the money supply, the larger the windfall.

Extracting this higher level of seigniorage assumes, however, no change in the underlying demand for money. For the reasons noted in the introduction, however, this is unlikely. Conflict reduces incomes: it therefore lowers the demand for money and hence seigniorage for any inflation rate. But conflict also induces private agents to disengage from the formal economy and to seek opportunities for capital flight and currency substitution, entailing both an autonomous shift out of domestic currency (a decline in c_t in equation (4) and an increase in the inflation semi-elasticity of the demand for money, α. These effects on private sector money demand combine to unambiguously shift the Laffer curve downward, so that the level of seigniorage previously feasible at point C cannot now be achieved at inflation rate $\tilde{\pi}_1{}^*$. This level could be generated at a higher rate of inflation, such as that prevailing at D, but clearly the shift inward could be sufficiently large as to render this level of seigniorage infeasible in equilibrium, forcing the government to accept a lower (credible) seigniorage yield.[4]

Learning

The onset of conflict often comes as a surprise. However, it is reasonable to assume that as conflict endures, the private sector learns about the government's policy rule and adjusts its expectation algorithms accordingly (see, for example, Marcet and Nicolini 2003). Learning algorithms are not explicitly introduced here, but those considered in the literature would be consistent with a gradual increase in β_t in the face of rising inflation. As learning progresses, inflation expectations adjust more rapidly, the saddle path in figure 1 gets steeper (more so if at the same time the semi-elasticity of money demand increases), and the transitional seigniorage revenue shrinks. Eventually, when $\beta_t \geq 1/\alpha$, the polarity of the saddle path reverses and the economy suddenly experiences an explosive path for inflation expectations.[5]

4. Depending on the precise changes in the demand for money, the seigniorage-maximizing rate of inflation may increase or decrease. As noted earlier, this inflation rate is defined as $\tilde{\pi}^{\max} = (1/\alpha) - \hat{y}$. Whether it rises or falls during and after conflict depends on the inflation semi-elasticity of money demand, the income elasticity of demand, and the growth of income.

5. Marcet and Nicolini (2003) use adaptive learning algorithms of this class to analyze hyperinflations in Latin America, which tend to follow the same trajectory of periods of high but stable inflation followed by hyperinflationary bursts that are brought under control by aggressive exchange rate–based stabilization.

Postconflict Responses

The postconflict period is in some respects a halfway house between peace and war. Although GDP starts to recover, the process takes many years; given the legacy of conflict, money demand is likely to remain below its peacetime level. Moreover, there is evidence that episodes of the loss of fiscal control tend to reduce the post-crisis income elasticity of the demand for money, thereby slowing the remonetization of the economy once inflation pressures have passed and growth has recovered (Adam and Bevan 2004). Similar patterns are likely to be present in the wake of conflict-induced increases in inflation.

While the constraint on raising seigniorage remains tight, government spending needs remain higher than before the war and are indeed likely to increase. Postconflict reconstruction cannot generally be financed by a fiscal peace dividend, because the high risk of conflict reversion typically keeps military spending close to wartime levels (Collier and Hoeffler 2006). Both the persistence of the reduced demand for money and the increased demand for spending imply that the government would choose a higher rate of inflation, such as that entailed by a point in the region of D on the new Laffer curve in figure 1.

II. Empirical Analysis

The empirical analysis uses annual data for a panel of 66 developing economies, 30 of which experienced at least one episode of civil war between 1964 and 2002 (appendix table A-1). Organization for Economic Co-operation and Development countries are excluded, because they are generally free of civil war and tend not to rely on seigniorage to nearly the same extent as other countries. Former communist countries are also excluded, principally because of lack of data. Finally, countries in currency unions are excluded, because union membership constrains their scope for seigniorage. (South Africa is retained, because its dominant role in the Common Monetary Agreement of Southern Africa means that it enjoys de facto full monetary independence.)

Countries with more than 1000 battle-related deaths are classified as being in a state of civil war in that year.[6] A recent innovation in data on civil war has been the development of measures of conflict intensity based on the extent of combat-related mortality (Lacina and Gleditsch 2005). A priori, it is unclear whether the monetary effects of a conflict will be more closely related to a measure of its intensity (such as combat-related mortality) or to a state-dependent measure. The core results reported here are based on a state-dependent measure, on the grounds that it better reflects the quantum effects a shift from a state of peace to a state of war may have on expectations relevant for asset demands. (The robustness of the results is tested to alternative measures of conflict; see table A-2.)

6. All war-related data are from the Uppsala Conflict Data Program/International Peace Research Institute Armed Conflict Dataset Version 4-2006 (see Gleditsch and others 2002).

Four stylized facts characterize the data (table 1). On average, inflation rises during conflict and falls following conflicts, but it remains higher than before the conflict. Seigniorage follows a similar pattern, rising by more than 1 full percentage point of GDP during the conflict (against a preconflict level of 1.8 percent of GDP) before falling back toward, but not quite achieving, its preconflict level. The composition of seigniorage revenue changes markedly. During conflict, governments rely more heavily on seigniorage raised through reserve requirements on the banking system. The differential responses of the (unconstrained) nonbank private sector and the (highly constrained) bank sector are reflected in the summary statistics on real balances. Currency holdings of the nonbank private sector fall from 8.2 percent to 6.2 percent of GDP, while bank reserves rise by almost 3 percentage points of GDP. As a result, the average seigniorage yield on bank reserves rises sharply during conflict. Following a conflict, both revert toward their preconflict values.

Money Demand

The regression analysis begins with money demand, which constitutes the constraint on government choices. Currency in circulation and reserve holdings by the banking system are distinguished, because reserve holdings may be subject to government control rather than being a choice variable for the private sector. Cagan-style money demand functions of the form

$$
\begin{aligned}
\ln(m)_{it} = {} & \gamma_0 + \gamma_1 \ln(y)_{it-1} + \gamma_2 \ln(pop)_{it} \\
& + \gamma_3 \tilde{\pi}_{it-1} + \gamma_4 war_{it} + \gamma_5 postwar_{it} \\
& + \gamma_6[\ln(y)_{it-1} \cdot war_{it}] + \gamma_7[\ln(y)_{it-1} \cdot postwar_{it}] \\
& + \gamma_8[\tilde{\pi}_{it-1} \cdot war_{it}] + \gamma_9[\tilde{\pi}_{it-1} \cdot postwar_{it}] + \varepsilon_{it}
\end{aligned} \tag{11}
$$

are estimated (table 2), where countries are denoted by i and time by t. The dependent variable, corresponding to each of the measures of money, is defined as a share of GDP. The analysis expands on the specification in equation (4) by

TABLE 1. Descriptive Statistics (Means)

Statistic	Full sample (1964–2002)	Prewar	War	Postwar
Inflation (percent per year)	14.3	13.3	20.4	15.3
Reserve money (percent of GDP)	11.4	10.7	11.6	11.7
Currency	6.6	8.2	6.2	6.8
Reserves	4.8	2.6	5.4	4.8
Seigniorage (percent of GDP)	2.1	1.8	3.0	2.3
Currency	1.0	1.2	1.2	1.1
Reserves	1.0	0.5	1.8	1.2
Aggregate GDP growth (percent per year)	3.6	4.4	2.6	4.5
Per capita GDP growth (percent per year)	1.2	1.9	0.4	2.3
Aid (percent of GDP)	6.0	4.5	4.6	6.5

Source: See table A-3.

TABLE 2. Fixed-Effects Estimates of Log Money Demand (percent of GDP). OLS Estimates.

Variable	Log reserve money (1)	Log currency (2)	Log bank reserves (3)	Log currency (4)	Log currency (5)	Log bank reserves (6)
Constant	−1.690 (9.67)	−1.953 (8.21)	−5.130 (15.80)	−2.700 (11.88)	−2.640 (11.85)	−4.975 (12.11)
Log real GDP				0.007 (2.21)	0.005 (1.87)	0.008 (1.09)
lrgdp*war				−0.007 (0.93)	−0.008 (1.01)	−0.008 (0.48)
lrgdp*postwar				−0.054 (4.59)	−0.025 (2.06)	−0.018 (0.92)
Log population	−0.020 (0.32)	0.190 (2.16)	−0.122 (1.07)	0.449 (5.34)	0.432 (5.21)	−0.168 (1.14)
infl	−0.239 (2.77)	−0.541 (8.72)	−0.026 (0.17)	−0.484 (7.06)	−0.461 (6.60)	−0.204 (1.07)
infl*war				−0.337 (2.53)	−0.257 (1.92)	0.745 (2.36)
infl*postwar				0.089 (0.49)	0.036 (0.22)	0.416 (1.32)
war	−0.126 (3.89)	−0.074 (2.73)	−0.251 (4.28)	0.044 (0.62)	−0.038 (0.50)	−0.268 (1.79)
postwar	−0.189 (6.14)	−0.178 (5.38)	−0.287 (5.33)	0.264 (2.53)	−0.019 (0.18)	−0.167 (0.93)
aid*war					1.001 (6.74)	−0.098 (0.15)
aid*postwar					0.831 (4.40)	−0.679 (1.74)
Pooling *F*-test (country = 0) Probability	124.4 (0.000)	213.97 (0.000)	55.71 (0.000)	202.36 (0.000)	208.6 (0.000)	53.116 (0.000)
F-test (war = postwar)	3.5 (0.0129)	8.65 (0.003)	0.4 (0.529)	3.59 (0.058)	0.02 (0.882)	0.24 (0.626)
F-test (infl*war = infl*postwar)				4.52 (0.034)	2.59 (0.107)	0.66 (0.417)
F-test (lrgdp*war = lrgdp*postwar)				12.06 (0.001)	1.35 (0.245)	0.22 (0.643)
F-test (aid*war = aid*postwar)				0.56 (0.456)	0.62 (0.431)	
R-squared	0.704	0.832	0.607	0.848	0.855	0.612
Number of observations	2,009	2,004	2,004	1,925	1,908	1,908

Note: All specifications include country and year dummy variables. infl denotes inflation factor (see Table A3). The interaction effect between log real GDP and the war dummy is denoted lrgdp*war and similarly for all other interaction effects. Figures in parenthesis beside coefficient estimates are *t*-statistics; figures in parenthesis beside *F*-statistics are probability values.

Source: Authors' calculations based on analysis in text.

allowing for the possibility that the per capita income elasticity of the demand for money deviates from 1: it thus includes population (*pop*) and real income (y) as regressors. To avoid potential problems of endogeneity, $\ln(y)$ enters equation (11) with a lag. Inflation, $\tilde{\pi}$, is as defined in equation (4). The dummy variable *war* takes the value of 1 if country i is in a state of civil war at time t and 0 otherwise; *postwar* takes the value of 1 in the first 10 years following the ending of hostilities. The equation residual $\varepsilon_{it} = \mu_i + \omega_t + v_{it}$ is a conventional two-way error component residual. Each equation is estimated using a within-groups/fixed-effects estimator with a full set of common time and year dummies. Pooling tests reject the null hypothesis of a common intercept.[7]

Controlling for income and inflation, the analysis first introduces civil war and its aftermath as dummy variables (columns 1–3). Both are highly significant and negative for reserve money as a whole and for each of its components. Over and above any effects through income and inflation, conflict reduces the demand for money. More surprisingly, this direct erosion in the demand for money appears to intensify during the postconflict decade, when money demand declines by 19 percent relative to peacetime and 6 percent relative to wartime. The decline relative to wartime is statistically significant, overall and for currency demand.

Bank reserves are not a direct choice variable for the private sector; they are determined by the interaction of government policy on reserve requirements, the banking sector's liquidity preference, and the private sector's demand for inside money. It is, then, not surprising that except for the autonomous *war* and *postwar* shift effects, the results for bank reserves are weak: neither income nor inflation is statistically significant, and the overall fit of the equation is markedly lower than for currency demand. What follows therefore concentrates on the demand for currency (the question of how the authorities balance their seigniorage extraction between these two sources is returned to later).

To investigate the transmission paths for this erosion in the private sector's money demand, interaction terms are introduced between both the *war* and *postwar* dummy variables and inflation (column 4). Similarly, the possibility that the income elasticity of money systematically differs from unity during and after conflict is allowed for. With the introduction of these interaction terms, the direct effects of both dummies cease to be significant. As suggested by the model in section I, the wartime erosion of currency demand works principally through a heightened sensitivity to inflation: not only does the (absolute) inflation semi-elasticity rise significantly during war, the increase in inflation generates a disproportionate reduction in money demand. Postwar the inflation semi-elasticity of demand is not significantly different from the prewar environment. However, the decline in the income elasticity of money demand, as attested by the postwar interaction term, means that the recovery in postwar

7. Whether behavior differed between the first and second five-year postconflict subperiods was investigated. Because the data do not reject pooling across the subperiods, however, only a single "postwar" effect is reported.

income does not rebuild money demand proportionally, because the private sector continues to reduce its need for currency per unit of income.

These results are tested for robustness to various measures of the intensity of conflict (the results are reported in appendix table A-2). Columns 1–3 in table A-2, in which conflict is measured directly by the number of combat-related deaths, correspond directly to the same columns in table 2. The wartime dummy variable is replaced by the number of deaths in each year, and the postwar dummy variable is replaced by the cumulated number of these deaths during the preceding conflict. In columns 4–6 a variant is investigated in which the number of deaths is scaled by population. For both specifications the new variables are significant, the size of the effect on money demand being very close to that found using the dummy variables. Non-nested encompassing tests suggest that from a purely statistical perspective, no measure of conflict dominates. Changing the way in which conflict is measured does not affect the other determinants of money demand, in particular the inflation semi-elasticity, which is the focus here. Column 7 in table A-2 extends the robustness check by interacting the combat death measure of conflict with inflation, with the same results as in table 2, column 4. What follows therefore relies on the results of table 2, in which conflict is treated as a state variable.

The Government's Choice of Seigniorage

In raising seigniorage, the government has two instruments at its disposal, the supply of currency and the reserve requirements on the banking system. The supply of currency induces inflation; the reserve requirement is a tax on the banking system, which is liable to reduce the allocative efficiency of finance. The government's chosen level of seigniorage can be expressed as:

$$s_{it} = \beta_0 + \beta_1 war_{it} + \beta_2 postwar_{it} + \lambda_1 \ln(m)_{it} + u_{it}, \tag{12}$$

where s is the seigniorage revenue corresponding to the three measures of money, expressed as percentage points of GDP, and u_{it} is a two-way error term, as defined above. Substituting for ln (m) from the money demand equations allows choices to be expressed as determined by the direct and indirect "structural" effects of war and postwar. Collecting the war and postwar terms allows seigniorage outcomes to be expressed as follows:

$$\begin{aligned} s_{it} &= (\hat{\beta}_0 + \hat{\lambda}_1 \hat{\gamma}_0) + \hat{\lambda}_1 [\hat{\gamma}_1 \ln(y)_{it-1} + \hat{\gamma}_2 \ln(pop)_{it} + \hat{\gamma}_3 \pi_{it-1}] \\ &+ [\hat{\beta}_1 + \hat{\lambda}_1 (\hat{\gamma}_4 + (\hat{\gamma}_1 + \hat{\gamma}_6) \ln(y)_{it-1} + (\hat{\gamma}_3 + \hat{\gamma}_8) \bar{\pi}_{it-1})] war_{it} \\ &+ [\hat{\beta}_2 + \hat{\lambda}_1 (\hat{\gamma}_5 + (\hat{\gamma}_1 + \hat{\gamma}_7) \ln(y)_{it-1} + (\hat{\gamma}_3 + \hat{\gamma}_9) \bar{\pi}_{it-1})] postwar_{it} + w_{it} \end{aligned} \tag{13}$$

where w_{it} is the composite error term from regressions 11 and 12. The parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ reflect the direct "choice" effects; the terms multiplied by $\hat{\lambda}_1$ reflect the indirect effects of war and postwar on the demand for money.

Consider the combined seigniorage from currency in circulation and bank reserves. Columns 1–3 in table 3 report the results of the seigniorage regressions. Controlling for the level of real money balances, regressions introduce dummy variables for the war and postwar periods. Both are highly significant: wartime and postwar governments resort more to seigniorage than peacetime governments do. The coefficients on the two dummies are not significantly different from each other: postwar governments are as desperate for revenue as wartime governments. Conditional on the level of the constraint, war conditions increase seigniorage extraction by about 1.1 percent of GDP. Given that prewar seigniorage is about 1.8 percent of GDP (see table 1), this is a substantial increase in needs. In the postconflict decade the direct effect is virtually as large as during the war, at 0.8 percent of GDP. This is consistent with the observed continuing high levels of government military spending in postconflict conditions.

While the total direct effects of wartime and postwar conditions on the government's resort to seigniorage are the same, their composition differs. During war governments rely predominantly on taxing the banking system, with four-fifths of total seigniorage generated from this source. Postwar, although taxation of the banking system still dominates, there is some shift toward greater reliance on printing currency. This shift may be appropriate: because of the collapse of investment, allocative financial efficiency may temporarily be unimportant during conflict.

War and postwar conditions alter seigniorage not only through direct effects on government choices but also through their effects on income and money demand. While government needs increase pressure to resort to seigniorage, the tightening of the constraint resulting from lower income and the erosion in the demand for money reduces the amount of seigniorage that can be raised. The direct effect combined with the two offsetting indirect effects determines the overall effect of war and its aftermath on seigniorage. The decline in overall seigniorage between war and postwar noted in table 1 reflects both a decline in needs, of about 0.3 percentage points of GDP on average, and a reduction in the seigniorage tax base of about the same amount (table 4).

III. Postconflict Assistance

During the postwar period, the government is faced with a harsh tradeoff. The need for revenue increases, but the capacity to raise it through seigniorage deteriorates, as the demand for money erodes. From the long-run perspective there is a case for reducing resort to seigniorage, thereby investing in the reconstruction of money demand and so restoring the future potential for sustainable seigniorage. From the short-run perspective there is a case for further resort to seigniorage, despite its rising cost in terms of inflation and damage to the banking system. Does aid resolve this dilemma?

TABLE 3. Fixed-Effects Estimates of Seigniorage Choice (Percent of GDP)

Variable	Seigniorage from reserve money (ordinary least squares) (1)	Seigniorage from currency (ordinary least squares) (2)	Seigniorage from bank reserves (ordinary least squares) (3)	Seigniorage from reserve money (instrumental variable) (4)	Seigniorage from reserve money (instrumental variable) (5)	Seigniorage from currency (instrumental variable) (6)	Seigniorage from currency (instrumental variable) (7)
Constant	0.069 (10.65)	0.039 (11.88)	0.062 (7.44)	0.068 (9.60)	0.068 (9.59)	0.037 (10.65)	0.070 (5.90)
war	0.011 (3.83)	0.003 (3.00)	0.008 (3.24)	0.009 (0.96)	0.008 (3.29)	0.002 (1.41)	0.011 (3.03)
postwar	0.008 (3.38)	0.002 (1.78)	0.006 (3.05)	0.027 (1.78)	0.027 (2.02)	0.007 (1.18)	0.016 (1.30)
ln money demand	0.030 (8.38)	0.008 (6.56)	0.013 (8.28)	0.027 (7.15)	0.027 (7.54)	0.008 (5.62)	0.012 (6.78)
aid*war				−0.024 (0.15)			
aid*postwar				−0.320 (1.23)	−0.332 (1.50)	−0.087 (0.93)	−0.198 (0.88)
Pooling *F*-test (country = 0)	6.04 (0.000)	7.71 (0.000)	3.1 (0.000)	5.31 (0.000)	5.28 (0.000)	9.28 (0.000)	2.97 (0.000)
F-test (war = postwar)	1.35 (0.246)	1.16 (0.283)	1.08 (0.299)	0.67 (0.414)	2.58 (0.108)	1.04 (0.306)	0.22 (0.636)
R-squared	0.348	0.377	0.284	0.2453	0.2404	0.3195	0.2381
Number of observations	2,200	2,193	2,193	2,003	2,003	1,996	1,652
Cragg-Donald Identification Test (Ho: underidentified)				(0.008)	(0.000)	(0.0004)	(0.0005)
Shea Weak Instrument Test (Ho: instrument is weak)							
aid*war				0.0227 (0.000)			
aid*postwar				0.0177 (0.0396)	0.0177 (0.0396)	0.0148 (0.0742)	0.0092 (0.0178)

Note: All specifications include country and year dummy variables. The interaction effect between aid and the war dummy and between aid and the postwar dummy are denoted aid*war and aid*postwar respectively. These variables are treated as endogenous regressors and are instrumented. See text for details. Figures in parenthesis beside coefficient estimates are *t*-statistics; figures in parenthesis beside *F*-statistics and tests for identification and weak instruments are probability values.

Source: Authors' calculations based on analysis in text.

TABLE 4. Decomposition of War and Postwar Seigniorage Yields (Percent of GDP)

Period/cause of change in seigniorage	Reserve money	Currency	Bank reserves
War			
Direct needs	1.10	0.30	0.80
Constraint shift	−0.34	−0.06	−0.33
Net seigniorage	0.76	0.24	0.47
Postwar			
Direct needs	0.80	0.20	0.60
Constraint shift	−0.59	−0.14	−0.37
Net seigniorage	0.21	0.06	0.23
Mean prewar seigniorage	1.80	1.20	0.50

Source: Authors' calculations based on analysis in text.

Postconflict reconstruction is the original rational for aid. Indeed, the original name for the World Bank was the International Bank for Reconstruction; the words "and Development" were added as an afterthought. Following the end of conflict, there is typically a surge in aid, as donors respond to perceived postconflict needs. Aid has indeed been found to be significantly more effective in enhancing growth during the postconflict decade than at other times (Collier and Hoeffler 2004a). Postwar aid can also have monetary effects, which can affect the government's need for seigniorage.

Aid and the Government Choice of Seigniorage

Aid, expressed as a share of GDP, is introduced into the seigniorage regression as follows:

$$s_{it} = \beta_0 + \beta_1 war_{it} + \beta_2 postwar_{it} + \beta_3[aid_{it} \cdot war_{it}] + \beta_4[aid_{it} \cdot postwar_{it}] + \lambda_1 \ln(m)_{it} + u_{it} \tag{14}$$

The supply of aid cannot be assumed to be exogenous to government fiscal choices: donors might plausibly either increase aid flows in response to fiscal desperation or reduce them in response to fiscal irresponsibility. Aid is therefore instrumented using a vector of political, cultural, and economic measures of distance between each recipient country and its principal Development Assistance Committee aid donors (see Tavares 2003).[8] The underlying idea is that to some extent, bilateral donor governments provide aid based on historical ties and

8. The aid instrument is defined as $\tilde{a}_{it} = \sum_j \theta_{ij} A_{jt}$, where $\tilde{a}_{it}$ denotes instrumented aid for recipient i in period t; $\theta_{ij} = (1/D_{ij}, L_{ij}, R_{ij})$ is a vector of time-invariant measures of "distance" between donor j and recipient i, where D_{ij} is the distance between the capital cities of i and j; L_{ij} is a dummy variable taking the value of 1 when i and j share the same official language and 0 otherwise; R_{ij} is a dummy variable taking the value of 1 when i and j share the same dominant religion and 0 otherwise; and A_{jt} is donor i's aid to GNI ratio in time t. This measure is calculated for the five principal aid donors: the United States, the United Kingdom, Japan, France, and Germany.

domestic budgetary circumstances that are unrelated to circumstances in the recipient country. As the instrument validity tests reported at the bottom of table 3 suggest, the instrumenting strategy adopted here appears reasonably robust.

Columns 4 and 5 in table 3 introduce aid instrumented in this way into the seigniorage regression for reserve money. During wartime aid is unsurprisingly negligible; instrumented aid is therefore insignificant when interacted with the wartime dummy variable. In effect, the aversion of donors to funding warfare overrides the proclivities of bilateral donors to provide aid. When the interaction of aid with this wartime dummy variable is dropped, the interaction between aid and the postwar dummy variable is negative and substantial. Although it is only on the borderline of significance, it is fully consistent with theory: it is not surprising that aid reduces the resort by government to seigniorage. An increase in aid equivalent to 1 percentage point of GDP reduces seigniorage by about 0.33 percentage points of GDP.

Because aid usually surges following the end of conflict, the addition of aid makes a substantial difference to the other components of the regression. In particular, the direct effect of the postwar dummy variable is now about four times its previous value. Although the regression involves a small reduction in the sample, this is not the explanation for the change.[9] Thus, controlling for aid, postwar governments appear to be far more desperate for revenue even than they are during war. The surge in aid postwar accommodates these needs and thereby reduces resort to seigniorage.

Columns 6 and 7 in table 3 again show how the resort to seigniorage is split between currency and bank reserves, in this case controlling for post-conflict aid. Although the differences between seigniorage extracted from currency and from bank reserves seen in columns 2 and 3 remain, there is no statistically significant difference in the impact of post-conflict aid.

Aid and Money Demand

Because aid reduces the resort to seigniorage, it reduces inflation. Indirectly, aid therefore raises the demand for money. Aid also significantly augments growth in the postwar context (Collier and Hoeffler 2004a), which directly raises the demand for money. The theoretical specification of the demand function suggests that conditional on income and inflation, aid should have no direct effect on private sector demand for money. However, in the context of the aftermath of civil war, there may be such an effect. One likely explanation for the severe erosion of the demand for money in the postwar period is a "peso effect," in which the high risk of a reversion to conflict supports expectations of further inflation in excess of that directly implied by current experience. Conceivably, large aid programs might help reassure citizen, thereby directly increasing the demand for money. This hypothesis is tested, by modifying the

9. Reestimating the regression in column 1 on the reduced sample does not significantly affect the coefficient estimates.

demand for money function as follows:

$$
\begin{aligned}
\ln(m)_{it} = {} & \gamma_0 + \gamma_1 \ln(y)_{it-1} + \gamma_2 \ln(pop)_{it} + \gamma_3 \tilde{\pi}_{it-1} \\
& + \gamma_4 war_{it} + \gamma_5 postwar_{it} \\
& + \gamma_6[\ln(y)_{it-1} \cdot war_{it}] + \gamma_7[\ln(y)_{it-1} \cdot postwar_{it}] \\
& + \gamma_8[\tilde{\pi}_{it-1} \cdot war_{it}] + \gamma_9[\tilde{\pi}_{it-1} \cdot postwar_{it}] \\
& + \gamma_{10}[aid_{it} \cdot war_{it}] + \gamma_{11}[aid_{it} \cdot postwar_{it}] + \varepsilon_{it}.
\end{aligned} \tag{15}
$$

In a reduced-form regression of the type estimated here, any impact of aid inflows must reflect a direct effect, because the regression controls for the indirect effects.

It is reasonable to assume that aid is weakly exogenous with respect to private sector demand for money; aid is therefore introduced into the demand for money functions using ordinary least squares (table 2, columns 5 and 6). Aid has a small and statistically insignificant effect on the demand for reserve money (not shown in table 2), but this is because it has strong but offsetting effects on the two components of money. Aid significantly increases the demand for currency in circulation both during war and in the postconflict period (column 5); it also reduces the demand for bank reserves by the same order of magnitude (column 6). The "demand" for bank reserves in the context of civil war and its aftermath is most reasonably interpreted as being a coerced demand: banks must comply with central bank regulations. Hence, the decline in "demand" as a result of postwar aid is likely to reflect the reduced pressure for government revenue. These aid effects are not large, but they are significant: as the regression controls for both income and inflation, the result is consistent with the hypothesis that aid reassures and so reduces the "peso problem" arising from fears of collapse.

The inclusion of aid in the money demand regression changes the direct effect of the postwar dummy variable quite substantially. In the analysis of section II, which omitted aid, it appeared that the demand for money eroded considerably more during the postwar period than during the war itself. Controlling for aid, this is no longer the case. The "pure" effect of the postwar period on money demand continues to be substantially adverse in comparison to the prewar peace, but it is now significantly and substantially better than during the war itself.

Postwar aid "reconstructs" the demand for money and hence seigniorage capacity through three distinct routes. First, postconflict aid directly substitutes for seigniorage revenue, enabling the government to reduce its reliance on inflationary finance and thereby stimulating recovery in the demand for money. Second, postconflict aid restores money demand indirectly, through its effect on income growth. While the overall effect of aid on growth is controversial, there is evidence that it is particularly effective in postwar situations (Collier and Hoeffler

2004a). Finally, as the regression results suggest, aid also appears to play a role in supporting a modest portfolio shift in favor of domestic money demand.

Applications to Postconflict Monetary Reconstruction

The regression results are applied to two pertinent questions concerning postconflict aid. The first concerns the marginal effects of postconflict aid; the second concerns the paths of reconstruction of the monetary base postconflict, with and without aid.

The analysis draws on the evidence from the seigniorage and money demand regressions, combined with evidence from Collier and Hoeffler (2004a), to simulate the marginal impact on inflation and money demand of an increase in postconflict aid. Given the peculiar nature of the demand for bank reserves, the focus here is exclusively on currency demand. Using the sample estimates for the levels of inflation and money demand at the start of the postconflict period, the analysis measures the marginal impact of an increase in aid of 1 percentage point of GDP, sustained over a 10-year period. The impact consists of two distinct components. The first is the change in inflation arising from the aid-induced fall in seigniorage needs, which determines the end-of-conflict demand for currency (this represents a movement along the immediate postconflict Laffer curve). The second is the improvement in postconflict currency demand, including the induced effect of the further fall in inflation associated with rising currency demand given the marginal change in seigniorage needs. The results (table 5) suggest that if the fall in desired seigniorage noted in table 3 were sustained over the 10-year postwar period, the aid inflow would lead to a substantial restoration in real money balances of about 2.9 percentage points of GDP (an increase of almost 50 percent over the end-of-conflict level) and an almost halving of inflation, from about 20 percent to just over10 percent a year (table 5).

The regression results are also used to track the evolution of the monetary base, inflation and seigniorage for the typical conflict-affected country under two scenarios: no aid and aid at the level typical of postconflict situations. Each dynamic simulation is computed as a recursive forecast (figures 2–4). Exploiting the adaptive inflation expectations specification used in the estimations, and given exogenous paths for real income growth (from tables 1 and 5), the demand for currency at time t is predetermined, given the relevant coefficient values from table 3. With currency demand predetermined, changes in seigniorage needs from table 2 then imply changes in inflation and the current seigniorage yield. The change in inflation is used to update the demand for money in $t + 1$ and so forth. The coefficients on the war and postconflict dummy variables are derived from a cross-section of observations; they therefore only approximate a genuinely dynamic analysis. In particular, they produce discrete jumps upon the onset and end of war that exaggerate the likely actual speed of adjustment. These artificial jumps are juxtaposed against more genuinely dynamic adjustments to inflation and income.

TABLE 5. Marginal Impact of 1 Percent Increase in Aid Sustained Over 10-Year Period

Channel	Impact	Impact on: Seigniorage (percent of GDP)	Inflation (percent per year)	Money demand (percent of GDP)
Inflation effect				
Impact of aid on aggregate seigniorage needs[a]	−0.33			
Decline in seigniorage from currency[b]		−0.12		
Initial postconflict currency balances[c]	6.20			
Inflation (factor)[c]	0.20			
Inflation semi-elasticity[d]	−0.72			
Change in annual inflation			−3.08	
Induced increase in currency demand				0.10
Income effect				
Impact of aid on postconflict growth (percent per year)[e]	0.26			
Postwar income elasticity of currency demand[d]	0.98			
Increase in currency demand	0.26			2.56
Postwar inflation semi-elasticity of currency demand[d]	−0.43			
Change in annual inflation)	−0.88			
Change in inflation over 10-year post war period			−677	
Induced increase in currency demand				0.14
Portfolio effect				
Portfolio shift coefficient[d]	0.83			
Currency balances, including Inflation effect[f]	6.30			
Increase in currency demand				0.05
Change in inflation over 10-year post war period			−0.18	
Induced increase in currency demand				0.00
Total effects				
Total direct reduction in seigniorage requirements		−0.12		
Total postwar increase in money demand				2.86
Total postwar reduction in inflation			−10.04	

[a]Table 3, column 5.

[b]Assumes that 36 percent of seigniorage revenue is to be raised from currency in circulation.

[c]Table 1

[d]Table 2, column 5.

[e]From Collier and Hoeffler (2004a)

[f]Includes increase in currency demand due to initial postconflict inflation reduction.

Source: Authors' calculations based on analysis in text.

FIGURE 2. Impact of Postconflict Aid on Inflation (percent per year)

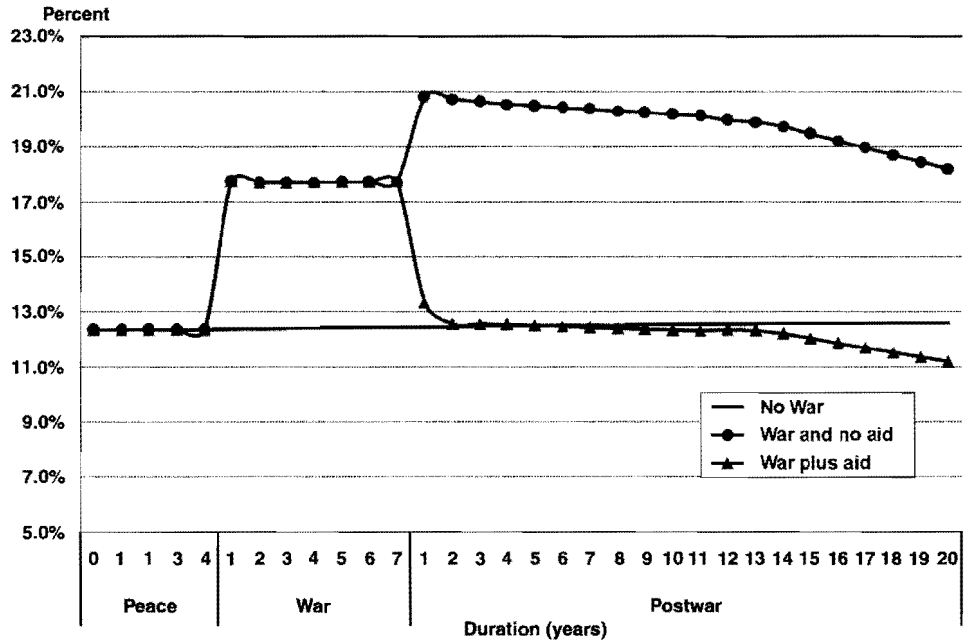


FIGURE 3. Impact of Postconflict Aid on Currency Demand (percent of GDP)

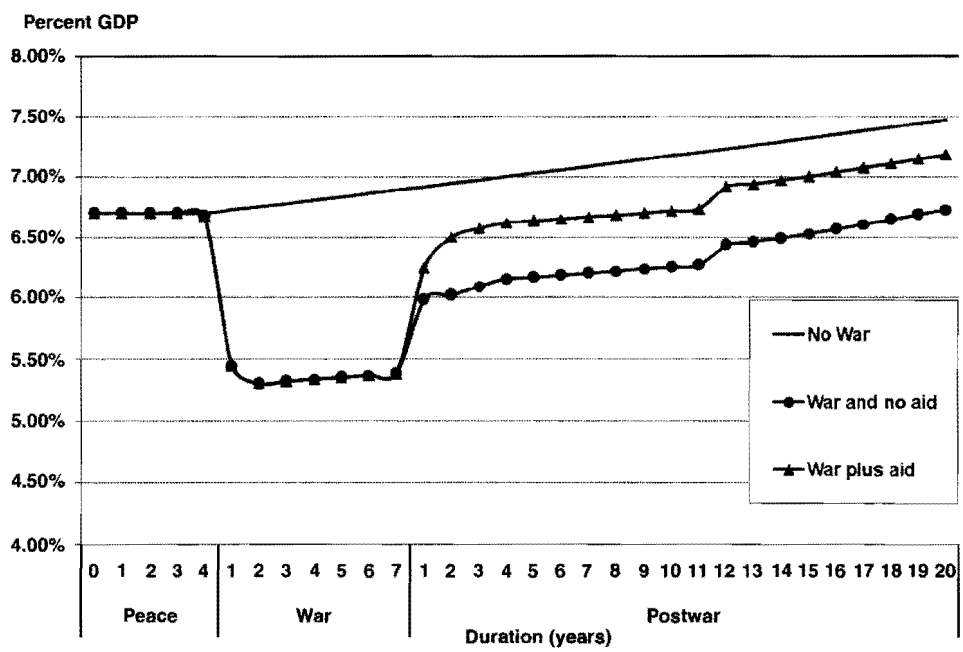


Figure 2 plots the rate of inflation. The typical country in the sample had a preconflict rate of inflation of about 12 percent. During war this rate rises to about 18 percent, although not as abruptly as depicted in the figure. In the absence of aid the regressions reported in tables 2 (column 5) and 3 (column 6) imply that there would be a substantial increase in inflation to 21 percent with the onset of peace, an adjustment that is likely to be more gradual than depicted. This increase reflects the heavy fiscal needs facing postconflict governments. Postconflict aid at typical levels is sufficient not only to meet these needs but to enable the government to invest in monetary reconstruction. The inflation rate with aid rapidly reverts to its peacetime level and indeed starts to dip below it, as is necessary to rebuild money demand. As postconflict aid typically surges immediately after the end of conflict, this fiscally driven effect may well be rapid.

Figure 3 plots the consequences for the demand for currency as a share of GDP. During war currency demand collapses from about 6.7 percent to about 5.3 percent of GDP. In the absence of aid currency demand nevertheless rebounds somewhat following the end of the conflict, to about 6 percentage points of GDP, but thereafter it stalls near this level: money demand is never rebuilt to its preconflict level. The rebound followed by stall is likely to be a spurious artifact of the postconflict dummy variable; the fact that the ceiling to the recovery is well short of the preconflict level is not. With aid the initial postconflict rebound is a little higher, at about 6.5 percent of GDP; the key difference is that thereafter it gradually recovers toward its peacetime level. Even after a decade, recovery is not complete, but it is substantially accomplished.

FIGURE 4. Impact of Postconflict Aid on Seigniorage (percent of GDP)

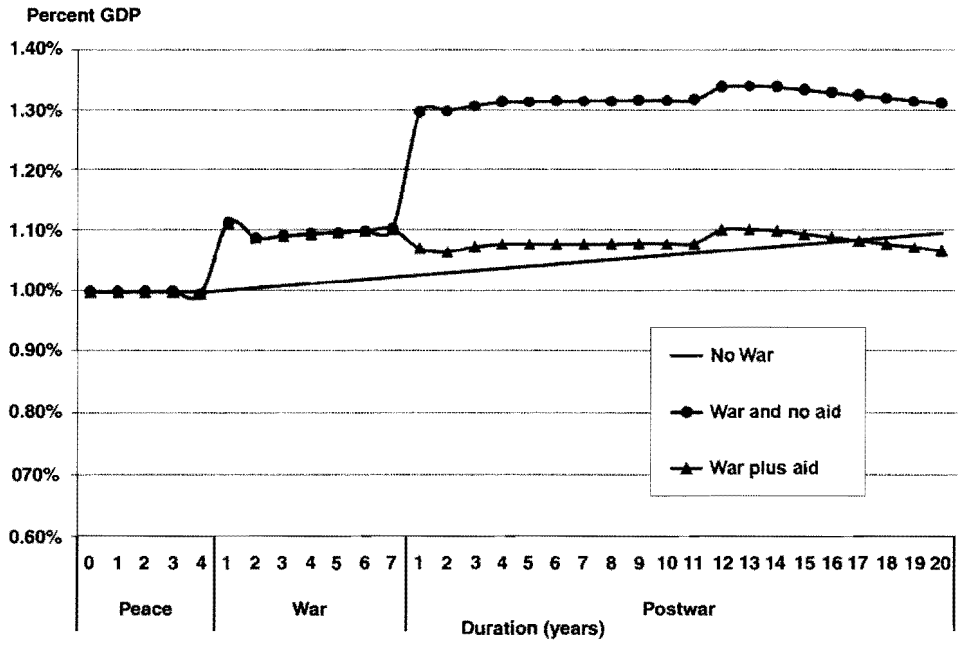


Figure 4 tracks seigniorage from currency as a percent of GDP. Such seigniorage increases during the war; if postconflict aid is not forthcoming, a dramatic exploitation of the currency occurs with the onset of peace. While the speed of the increase is surely exaggerated by the dummy variable, the onset of increased fiscal needs is indeed liable to be rapid. Aid enables the government not only to avoid this jump in seigniorage but gradually to bring its taxation of the currency down to its peacetime level. The process of reversion is slow, taking almost 20 years.

IV. Conclusion

Postconflict situations are characterized by an unusually wide range of outcomes. While on average economies rebound from wartime decline, decline continues in some countries, and about 40 percent of countries revert to conflict within a decade. Policy choices concerning the economic recovery of these hopeful but fragile situations have received far less attention than issues of political design and humanitarian needs.

Seigniorage is strategic, both because as revenue of last resort it reveals government preferences and because the ability to raise it reflects the degree of confidence of private actors in a fundamental government commitment. The results imply a rationale for aid that is peculiar to the postconflict macroeconomic situation. In effect, just as infrastructure needs to be reconstructed, so too does the demand for money. Even controlling for inflation and income, private demand for money erodes sharply in the postwar period. Yet in the absence of aid, postconflict governments resort to seigniorage far more heavily after war than during it.

While the restoration of the demand for money is beyond the capacity of the typical postconflict government to finance out of its own resources, it is both an important objective in itself and a useful indicator of the broader restoration of confidence. Aid is effective in reconstructing the long-term scope for seigniorage, acting through three distinct routes. The most obvious one is that for which aid is primarily intended: it raises the growth of income, thereby raising the demand for money. Unfortunately, this effect is relatively weak in postwar conditions, because the income elasticity of the demand for money is lower than in normal times. However, two other effects occur. First, aid reduces the need for the government to resort to seigniorage and so reduces inflation. Second, and more surprisingly, over and above the effects through income and inflation, postwar aid has a direct effect, perhaps through strengthening confidence in the maintenance of peace. This article treats aid as a single aggregate, abstracting from different types and uses. The core result, however, implies that it is aid to the budget that achieves monetary reconstruction. This need not necessarily imply the superiority of budget support. Because much project aid is likely to be fungible, it indirectly relieves the budget, even though it is ostensibly earmarked.

Postconflict aid is decisive in achieving monetary reconstruction. Rather than deteriorating, inflation and seigniorage—the monetary variables under the control of the government—revert to peacetime levels. The demand for currency, which is the constraint on government choices, takes longer to recover, because civil war severely damages confidence in the currency. Aid helps facilitate a gradual recovery. These monetary effects of postconflict aid have been an unsung success: attention has focused on the more televisual roles of aid in humanitarian relief and the reconstruction of physical infrastructure. The effects on monetary reconstruction are no less real or substantial.

Appendix

Table A-1. Civil War Episodes in Sample Countries

Country	Years
Algeria	1991–present
Argentina	1973–77
Bangladesh	1985–92
Burundi	1995–present
Colombia	1978–present
Egypt, Arab Republic of	1967, 1969–70
El Salvador	1979–91
Ethiopia	1966–91 2002–present
Guatemala	1966–95
India	1985–present
Indonesia	1975–92, 1997–present
Iran	1966–68, 1979–88, 1990–93, 1996–97, 1999–2001
Israel	1964–present
Lebanon	1975–91
Morocco	1975–89
Nepal	1999– present
Nicaragua	1978–79, 1981–89
Nigeria	1966–70
Pakistan	1971, 1974–77
Peru	1980–99
Philippines	1970– present
Rwanda	1991–94, 1997–2002
Sierra Leone	1992–2000
South Africa	1975–88
Sri Lanka	1971, 1983–2001
Sudan	1963–72, 1983–present
Syria	1979–82
Thailand	1974–82
Uganda	1978–79, 1981–91, 1994–present
Zimbabwe	1974–79

Source: Gleditsch and others 2002.

Note: Nonconflict countries in the sample include Bahrain, Barbados, Belize, Bhutan, Bolivia, Botswana, Brazil, Cape Verde, Chile, Costa Rica, Ecuador, Fiji, The Gambia, Ghana, Haiti, Honduras, Jamaica, Jordan, Kenya, Kuwait, Madagascar, Malawi, Mauritius, Oman, Paraguay, Saudi Arabia, the Seychelles, the Solomon Islands, Suriname, Tonga, Trinidad and Tobago, Tunisia, Uruguay, Vanuatu, República Bolivariana de Venezuela, and Zambia.

TABLE A-2. Robustness of Money Demand Results to Alternative Measures of Conflict (Percent of GDP)

	Battle deaths			Battle deaths per capita			Battle deaths
Variable	Log reserve money (1)	Log currency (2)	Log bank reserves (3)	Log reserve money (4)	Log currency (5)	Log bank reserves (6)	Log currency (7)
Constant	−1.634 (24.07)	−1.880 (7.80)	−5.064 (15.15)	−1.590 (9.03)	−1.930 (8.27)	−4.960 (11.74)	−1.880 (7.76)
log population	0.002 (4.72)	0.184 (2.10)	−0.113 (0.99)	−0.037 (0.59)	0.177 (2.05)	−0.164 (1.10)	0.188 (2.12)
infl	−0.245 (2.91)	−0.547 (8.76)	−0.050 (0.33)	−0.242 (2.86)	−0.571 (8.98)	−0.034 (0.23)	−0.594 (9.44)
infl*war							0.001 (0.67)
infl*postwar							0.008 (1.56)
war	−0.004 (3.11)	−0.002 (2.05)	−0.004 (1.96)	−0.028 (2.28)	0.001 (0.02)	−0.043 (2.78)	−0.003 (2.08)
postwar	−0.004 (2.85)	−0.004 (2.56)	−0.003 (1.58)	−0.038 (2.72)	−0.020 (1.53)	−0.048 (2.70)	−0.005 (2.86)
F-test (war = postwar)	0.02 (0.8853)	1.59 (0.207)	1.13 (0.288)	4.58 (0.033)	19.2 (0.000)	0.57 (0.449)	2.57 (0.109)
F-test (infl*war = infl*postwar)							1.54 (0.215)
R-squared	0.703	0.829	0.604	0.704	0.834	0.613	0.831
Number of observations	1,973	1,968	1,968	1,973	1,968	1,892	1,968
Encompassing test Ho: M2 encompasses M1	9.31	5.57	42.9	21.76	44.51	60.05	5.98

Note: All specifications include country and year dummies. The interaction effect between inflation and the war dummy and between inflation and the postwar dummy are denoted infl*war and infl*postwar respectively. Figures in parenthesis beside coefficient estimates are *t*-statistics; figures in parenthesis beside *F*-statistics and tests for identification and weak instruments are probability values. The Cox test is distributed N(0,1) under the null.

Source: Authors' estimates based on analysis in text.

TABLE A-3. Variable Definitions and Data Sources

Variable	Meaning	Definition and source
m	Money aggregate (percent of GDP)	Defined for reserve money (IMF 2006 line 14) and its components, currency in circulation (IMF 2006 line 14a) and bank reserves (IMF 2006 line 20), all measured as share of current price GDP in local currency (World Bank 2006)
s	Seigniorage (percent of GDP)	Defined for each money aggregate as $s_t = (M_t - M_{t-1})/((1/2)(Y_t + Y_{t+1}))$ for nominal money aggregate, M, and nominal GDP, Y.
y	Real GDP	Constant price GDP (World Bank 2006)
pop	Population (millions)	World Bank (2006)
$\tilde{\pi}$	Inflation factor	Defined as $\tilde{\pi} = \pi/(1 + \pi)$, where π denotes annual change in consumer price index (World Bank 2006)
war	Civil war indicator	See text for explanation (Gleditsch and others 2002)
$postwar$	Postwar indicator	See text for explanation (Gleditsch and others 2002)
aid	Aid (percent of GDP)	Net official development assistance (excluding technical assistance) as percent of GDP (World Bank 2006)

Source: Authors' compilation.

REFERENCES

Adam, Christopher. 1995. "Fiscal Adjustment, Financial Liberalization and the Dynamics of Inflation: Some Evidence from Zambia." *World Development* 23(5):735–50.

Adam, Christopher, and David Bevan. 2004. "Fiscal Policy Design in Low-Income Countries." In Tony Addison and Alan Roe eds., *Fiscal Policy for Development: Poverty, Reconstruction and Growth.* London: Palgrave Macmillan.

Bruno, Michael, and Stanley Fischer. 1990. "Seigniorage, Operating Rules and the High Inflation Trap." *Quarterly Journal of Economics* 105(2):353–74.

Cagan, Philip. 1956. "The Monetary Dynamics of Hyperinflation." In Milton Friedman, ed., *Studies in the Quantity Theory of Money.* Chicago: University of Chicago Press.

Calvo, Guillermo, and Leonardo Leiderman. 1992. "Optimal Inflation Tax under Pre-Commitment: Theory and Evidence." *American Economic Review* 82(1):179–94.

Cashin, Paul, Nadeem Ul Haque, and Nils Olekalns. 2002. "Tax Smoothing, Tax Tilting and Fiscal Sustainability in Pakistan." *Economic Modelling* 20(1):47–67.

Collier, Paul. 1999. "On the Economic Consequences of Civil War." *Oxford Economic Papers* 51(1):168–83.

Collier, Paul, and Anke Hoeffler. 2004a. "Aid, Policy and Growth in Post-Conflict Societies." *European Economic Review* 48(4):1125–45.

Collier, Paul, and Hoeffler Anke. 2004b. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–95.

———. 2006. "Military Spending in Post-Conflict Societies." *Economics of Governance* 7(1):89–107.

———. 2007. "Unintended Consequences: Does Aid Finance Military Spending?" *Oxford Bulletin of Economics and Statistics* 69(1):1–27.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenburg, and Havard Strand. 2002. "Armed Conflict 1956–2001: A New Data Set." *Journal of Peace Research* 39(5):615–37.

IMF (International Monetary Fund). 2006. *International Financial Statistics.* Washington, D.C.

Lacina, Bethany, and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21(2–3):2–3.

Mankiw, Gregory. 1987. "The Optimal Collection of Seigniorage: Theory and Evidence." *Journal of Monetary Economics* 20(2):327–41.

Marcet, Albert, and Juan Nicolini. 2003. "Recurrent Hyperinflations and Learning." *American Economic Review* 93(5):1476–98.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(112):725–54.

Tavares, Juan 2003. "Does Foreign Aid Corrupt?" *Economic Letters* 79(1):99–106.

World Bank. 2006. *World Development Statistics*. Washington, D.C.

Foreign Aid, the Real Exchange Rate, and Economic Growth in the Aftermath of Civil Wars

Ibrahim A. Elbadawi, Linda Kaltani, and Klaus Schmidt-Hebbel

Foreign aid, the real exchange rate (RER), and economic growth are three key variables that shape the aftermath of civil wars in many developing countries. Panel estimations drawn from a sample of 39 conflict and 44 nonconflict countries between 1970 and 2004 indicate that although postconflict countries receive larger aid flows and exhibit moderate RER overvaluation after peace is attained, overvaluation cannot be traced to aid. Yet foreign aid is among the significant determinants of the equilibrium RER. Aid is also an important determinant of economic growth, particularly after peace is reached. Aid exhibits decreasing returns, however, and interacts negatively with RER overvaluation. RER overvaluation reduces growth, but this effect is ameliorated by financial development. Postconflict policies should therefore aim to use aid prudently, avoid RER misalignment, and support financial and capital market development to achieve high and stable growth in the aftermath of war and beyond. JEL codes: F5, F3, F43

Countries emerging from civil wars usually have enormous humanitarian and developmental needs. Foreign aid can play an important role in the postconflict reconstruction of these economies, as well as in the consolidation of peace and the reduction in the risks of future conflicts.

Postconflict countries tend to experience surges of aid flows, which have potential consequences for their competitiveness, measured by the real

Ibrahim Elbadawi (corresponding author) is a lead economist in the Development Economics Research Group of the World Bank; his email address is ielbadawi@worldbank.org. Linda Kaltani is an economist at the International Monetary Fund; her email address is lkaltani@imf.org. Klaus Schmidt-Hebbel is professor of economics at the Catholic University of Chile; his email address is kschmidt@bcentral.cl. A preliminary version of this article was presented at the first collaborative research project workshop, "Political Institutions, Development and a Domestic Civil Peace," held in Oslo, June 19–20, 2006. The conference was organized by the Development Economic Research Group of the World Bank; the Centre for the Studies of African Economies, at the University of Oxford; and the International Peace Research Institute (PRIO), Oslo. The authors thank Majak d'Agoot, Paul Collier, Jim de Melo, and workshop attendants for valuable comments and suggestions. Ibrahim Elbadawi and Linda Kaltani also thank Aart Kraay and Norman Loayza for helpful discussions. The authors acknowledge the valuable research support of Nahla Hilmi and Marcelo Ochoa. They are indebted to the editor and three anonymous referees for valuable comments and suggestions.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 113–140 doi:10.1093/wber/lhm024
Advance Access Publication February 7, 2008

exchange rate (RER) and growth performance. Recipient governments face three choices regarding their use of aid: whether to spend or save the aid, how to divide it across sectors, and how to define the currency composition of public saving. RER appreciation will be greater the more governments spend, the larger the share of spending on nontraded goods, and the larger the share of public sector saving in net domestic-currency assets.

As an economywide relative price affecting both inter- and intratemporal expenditure and resource allocation decisions, the RER has a potentially important influence on export growth, export diversification, and economic growth.[1] Large temporary aid inflows may cause domestic spending booms, leading to RER overvaluation and lower growth (Edwards 1989; Cottani, Cavallo, and Kahn 1990; Ghura and Grennes 1993; Razin and Collins 1997; Easterly and Levine 1997; Aguirre and Calderón 2005). This result—termed "Dutch disease"—has often been associated with commodity price booms; it may also be observed during postconflict aid surges.[2] The implication of the literature on the effects of resource booms is that postconflict governments receiving aid surges have reasons to be concerned about potential aid-induced RER appreciation, which may reduce competitiveness and growth prospects.

The indirect adverse effect of aid on growth caused by RER overvaluation is one of several channels of transmission from foreign aid to growth.[3] A direct and positive channel of aid is that it embodies access to foreign resources that may lead to higher growth through higher investment, better policies, and aggregate efficiency gains.[4] However, aid is unlikely to increase growth if recipient countries' institutions and policies are weak (in the extreme, growth could decline); aid resources are misspent; or absorptive capabilities of large aid inflows are limited (in the extreme, excessive aid could weaken domestic

1. The RER concept adopted here reflects the economywide price of nontraded goods relative to traded goods. A rise in the ratio leads to an appreciation of the RER, while a fall in the ratio leads to a depreciation of the RER. The RER is considered misaligned when it diverges from its equilibrium level, derived from a behavioral model that accounts for equilibrium in the nontraded goods markets as well as the intertemporal implications of current account sustainability. The RER is characterized as overvalued (undervalued) when it is larger (smaller) than the equilibrium level. The definition and measures of RER misalignment are examined in section III.

2. The traditional literature on Dutch disease deals with the macroeconomic and growth effects of commodity discoveries and price booms. Following the early theoretical work of Van Wijnbergen (1984) and Krugman (1987), more recent work (Sachs and Warner 1995; Spatafora and Warner 1999) focusses on the empirical evidence of resource booms on resource allocation, corruption and rent-seeking behavior, and growth effects.

3. Another indirect effect of aid on growth could occur if permanently higher aid raises the long-term equilibrium RER and a higher equilibrium RER reduces growth. This link is unconventional, as there is little theoretical and empirical support for the notion that a key relative price at its equilibrium level may hurt growth. An exception is Rodrik (2007), who provides a theoretical justification linking RER undervaluation to growth.

4. The academic and operational literature on aid effectiveness includes Collier and Hoeffler (2002); Addison (2003); Clemens, Radelet, and Bhavnani (2004); Kang and Meernik (2004); IMF (2005); and Schwartz, Halkyard, and Smith (2006).

incentives). This would be reflected in a nonmonotonic relation between growth and aid. Aid could interact with third variables in affecting growth. It could have a larger positive growth impact in conflict countries after peace is attained (Collier and Hoeffler 2002). Aid could also interact negatively with RER misalignment in affecting growth, having an effect beyond its effect on RER misalignment.

This article focusses on the impact of large aid flows—a typical postconflict phenomenon—on the RER and economic growth. It addresses the following questions:

- What is the time profile of aid, RER misalignment, and growth during the postconflict cycle?
- Is postconflict aid associated with disequilibrium RER appreciation?
- How serious is RER misalignment in postconflict, and how much of RER misalignment is explained by large postconflict aid flows?
- Is growth in postconflict greater than during civil war or in nonconflict countries?
- How do aid and RER misalignment affect growth? Is postconflict aid particularly effective in spurring growth?

There are three approaches to assessing the macroeconomic impact of aid. The first identifies the correlations between aid on the one hand and government spending/saving flows and asset stocks on the other. These correlations are not very meaningful, because they are based on budget constraints (not behavioral models) and therefore do not control for other factors, including private sector reactions to and government use of aid.[5] These limitations are overcome by the second approach, based on assessing aid within a fully specified macroeconomic general equilibrium model. The third method is based on specifying behavioral models for individual variables affected by aid—such as the RER and growth—that embody the aforementioned channels of transmission and are consistent with theory. This approach is adopted here.

The article is organized as follows. Section I analyzes the size and timing of postconflict aid during the postconflict cycle, as well as aid–RER misalignment–growth correlations, in a sample of 39 postconflict countries. Section II reports estimation results for the RER based on annual 1980–2004 data for a sample of 83 countries. By focussing on the role of aid on long-term RER behavior, this analysis allows measures of RER misalignment to be derived for different subperiods in the pre- and postconflict cycle and for the influence of aid flows on RER overvaluation in postconflicts to be traced. Section III reports estimation results for growth based on a sample of 77 countries and five-year

5. Aiyar, Berg, and Hussain (2005) and IMF (2005) examine the accounting features and qualitative aspects of the macroeconomics of foreign aid.

data spanning the period 1970–2004.[6] The growth specification allows testing for the influence of the postconflict cycle, foreign aid, and RER misalignment on growth, controlling for standard growth determinants and allowing for key interactions between postconflict periods, aid, RER misalignment, and financial development. Section IV summarizes the article's main conclusions and draws some policy implications.

I. Aid, RER Misalignment, and Growth Over the Conflict Cycle

This section analyzes the experience of a sample of postconflict aid-recipient countries and, when appropriate, compares these data with those for a control group of developed and developing countries that have not experienced conflict. The sample covers 39 postconflict countries and 44 nonconflict countries, using annual data for 1960–2004. In the many countries that experienced multiple conflicts, all conflicts that ended between 1960 and 2004 are taken into account. The start and end dates of wars are based on the definition of civil war and the data presented in Sambanis (2004) and PRIO (2005).[7]

Following Collier and Hoeffler (2002), three postconflict subperiods are distinguished; in addition a prepeace period (before the end of conflict) is examined to allow for a more comprehensive assessment of the key variables and relations over the full conflict cycle. The conflict cycle is defined as follows, where year 0 indicates the year in which peace is attained:

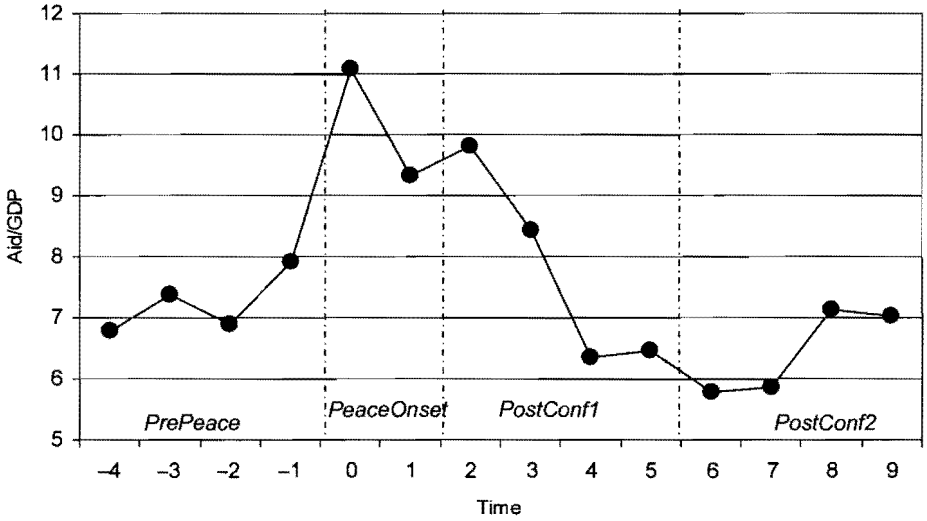
- *PrePeace* {yr(−4), . . . , yr (−1)} covers the four conflict years before the year peace begins (if the conflict is ongoing during part of this period, only that part is coded as prepeace).
- *PeaceOnset* {yr(0), yr (1)} covers the first two peaceful years, including the year in which peace is attained (yr(0)).
- *PostConf1* {yr(2), . . . yr (5)}covers the four years of peace after the onset of peace (unless another conflict starts, in which case this period accounts only for the peaceful years).
- *PostConf2* {yr(6), . . . yr (9)}covers the four consecutive peace years following *PostConf1* (unless another conflict starts, in which case this period accounts only for the peaceful years).

The time pattern of three key variables (annual aid flows as a share of GDP, RER misalignment, and per capita real GDP growth) observed in the average

6. The RER model can be estimated only for the shorter period (1980–2004), because the error-correction panel estimation methodology requires almost full data availability. The estimated model can be used to construct indexes of RER equilibrium and RER misalignment for the longer (1970–2004) period.

7. Elbadawi, Kaltani, and Schmidt-Hebbel (2007) provide a list of conflict countries that shows the end date (or dates) of each conflict.

FIGURE 1. Annual Aid Flows as a Share of GDP (average values, in percent)



Note: Figure is based on all conflict countries for which RER misalignment data are available (see the supplemental appendix, available at http://wber.oxfordjournals.org/, for details). The four conflict periods are defined in section I.

Source: Authors' analysis based on data described in the text.

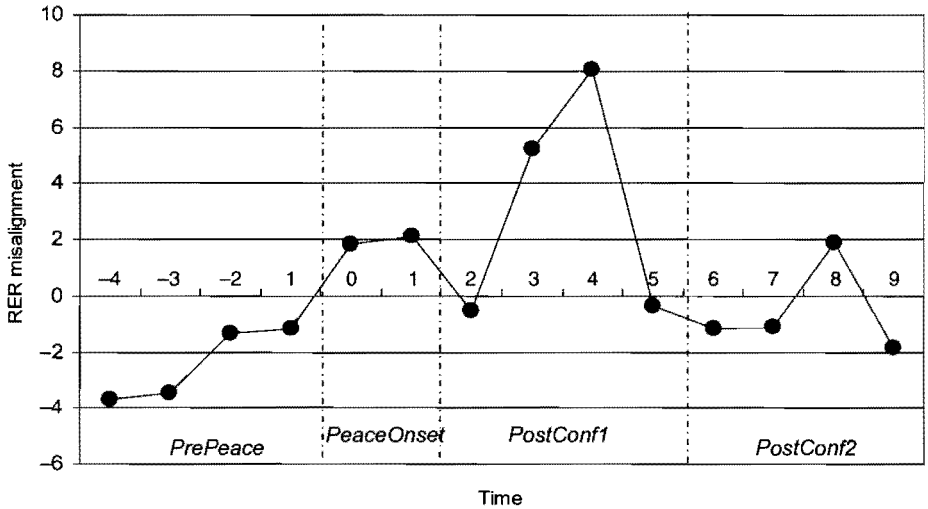
conflict country is tracked over the conflict cycle (figures 1–3).[8] The evidence indicates that postconflict countries receive larger aid shares with the start of peace. Aid peaks early during *PeaceOnset*, at 11 percent of GDP in the year peace is attained. Aid in *PostConf1* remains higher than in the *PrePeace* period, but it declines during *PostConf2* to the *PrePeace* average of almost 7 percent of GDP.

The RER tends to be undervalued during the period leading to the attainment of peace and overvalued during *PeaceOnset* and *PostConf1*, returning close to the equilibrium level in *PostConf2*. This suggests a lack of large and persistent exchange rate overappreciation in the average postconflict experience. Per capita real GDP growth turns negative the year peace is attained, rising in

8. The values shown are averages for each year of the conflict cycle (from year –4 to year +9). Because several countries in the sample experienced more than one conflict, some averages may be generated from more than the 39 cross-sectional units (some countries enter the average more than once). Many countries do not experience the full conflict cycle, either because they fall back into conflict or because the conflict ended less than nine years before 2004. The country sample thus changes for each yearly average. The samples used in figures 1–3 may also differ slightly because of data availability. Sri Lanka, for example, had a year of peace in 2002 and hence did not experience all the years of the conflict cycle. Because RER misalignment data were not available for 2004, the sample generating RER misalignment averages is slightly different from that for aid or growth.

Averages are presented rather than medians because individual observations present large and unstable differences across countries. Median values suggest a similar time profile for all variables over the conflict cycle but exhibit much larger volatility over time.

FIGURE 2. Real Effective Exchange Rate Misalignment (average values, in percent)



Note: Figure is based on all conflict countries for which RER misalignment data are available (see the supplemental appendix, available at http://wber.oxfordjournals.org/, for details). The four conflict periods are defined in section I.

Source: Authors' analysis based on data described in the text.

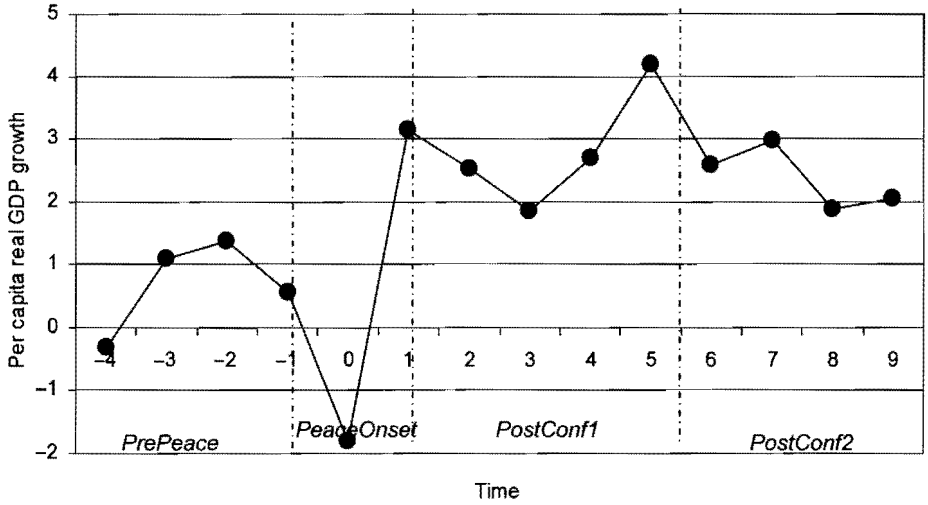
subsequent peace years to levels above those observed during conflict. In *PostConf2*, however, growth shows a slightly downward trend, suggesting an end of the output recovery during *PeaceOnset* and *PostConf1*.[9]

An event-study approach, reporting fixed-effect estimations of differences in country means between *PrePeace* and *PeaceOnset*, *PeaceOnset* and *PostConf1*, and *PostConf1* and *PostConf2* periods, is used to determine the statistical significance of the differences in the three key variables across the different conflict cycle periods (table 1).[10] It shows that aid is significantly higher (by 3.4 percent of GDP) during *PeaceOnset* than it is during *PrePeace*. Aid declines by 2.2 percent of GDP during *PostConf1* and by another 2.0 percent in *PostConf2*. The RER exhibits a temporary trend appreciation between conflict years −4 and +4 (which is subsequently reversed) (see figure 1). The statistical tests do not reveal significant differences in RER misalignment between different pairs of conflict periods, however. Growth is not statistically different

9. The average at time 0 is pulled down by one outlier: the year of conflict end in Rwanda (1994), when the growth rate was −64 percent. Without Rwanda the average would be −0.3 percent. In the growth regressions this value is not too worrisome, because the model uses five-year averages and the effect is significantly reduced by other, more-normal growth years.

10. Figures 1–3 depict unconditional averages for each conflict cycle year; table 1 reports estimated mean differences between conflict period averages, which are conditional on country fixed effects. The differences between conflict-period averages of the unconditional annual data depicted in figures 1–3 therefore do not match the differences reported in table 1.

FIGURE 3. Per Capita GDP Growth (average values, in percent)



Note: Figure is based on all conflict countries for which RER misalignment data are available (see the supplemental appendix, available at http://wber.oxfordjournals.org/, for details). The four conflict periods are defined in section I.

Source: Authors' analysis based on data described in the text.

TABLE 1. Event-Study Estimations of Differences in Variable Means across Pairs of Conflict and Postconflict Periods (in percent)

Item	Aid/GDP (1)	RER misalignment (2)	Growth (3)
PeaceOnset–PrePeace	3.41*** (1.10)	4.21 (3.11)	0.02 (1.09)
Number of observations	201	186	195
Number of countries	28	27	27
PostConf1–PeaceOnset	−2.16* (1.15)	−0.02 (3.33)	2.09* (1.11)
Number of observations	204	173	195
Number of countries	30	30	30
PostConf2–PostConf1	−1.95*** (0.69)	−2.09 (2.50)	−1.07 (0.72)
Number of observations	218	188	209
Number of countries	31	30	31

***Significant at the 1 percent level.

*Significant at the 10 percent level.

Note: Results show fixed-effects regressions based on annual data for 1960–2004. Numbers in parentheses are standard errors.

Source: Authors' analysis based on data described in the text.

during *PeaceOnset* and *PrePreace*, but a large output recovery in year 1 follows a recession in year 0 when peace is attained. During *PostConf1* GDP rises 2.1 percentage points over *PeaceOnset*, remaining high during *PostConf 2*.

Pairwise comovements of the key variables are examined by measuring correlations for three different country samples: countries that experienced no conflict between 1960 and 2004, countries that experienced conflicts for which data are available for years −4 to −1 (the prepeace years), and countries that experienced conflicts for which data are available for the full peace period (years 0–9). For robustness both cross-section and panel correlations for each country group are reported (table 2).

Aid and RER misalignment are not significantly correlated in most cases, and they are never positively and significantly correlated. This provides simple, unconditional evidence against generalized exchange-rate overvaluation associated with aid. The simple correlation between aid and growth is negative and significant in most cases, suggesting at least that the relation between aid and growth is not unambiguously positive. Growth and RER misalignment are

TABLE 2. Pairwise Correlations between Aid/GDP, RER Misalignment, and Growth, 1960–2004

Variable	Aid/GDP (1)	RER misalignment (2)	Per capita real GDP growth (3)
No-conflict countries			
Aid/GDP	1.000	0.036 (0.857)	−0.383** (0.044)
RER misalignment	−0.012 (0.727)	1.000	0.001 (0.997)
Per capita real GDP Growth	−0.097*** (0.003)	−0.043 (0.123)	1.000
Conflict countries during prepeace			
Aid/GDP	1.000	0.026 (0.888)	−0.542*** (0.001)
RER misalignment	−0.075 (0.400)	1.000	−0.051 (0.782)
Per capita real GDP Growth	−0.136* (0.104)	0.042 (0.639)	1.000
Conflict countries during peace			
Aid/GDP	1.000	−0.135 (0.421)	0.097 (0.556)
RER misalignment	−0.095* (0.094)	1.000	−0.424*** (0.008)
Per capita real GDP Growth	−0.123** (0.017)	−0.123** (0.030)	1.000

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Note: Upper diagonal portion of the table (above the diagonal string of 1.000's without *p*-values) illustrates cross-sectional correlations; lower diagonal portion illustrates panel correlations. Numbers in parentheses are *p*-values.

Source: Authors' analysis based on data described in the text.

negatively and significantly correlated only during the peace years in conflict countries, hinting that RER overvaluation could undermine output recovery during peace years.

This evidence is illustrative but should be taken with much care. Simple associations are not informative about causality (which could run either or both ways), do not consider the influence of other variables, and do not reflect nonmonotonic relations and interactions with third variables. To overcome these limitations, one needs to test the relations among the three key variables in a multivariate framework, based on behavioral equations that address dynamics, third variables, reverse causality, nonlinearities, and interactions. This is the task of the next two sections.

II. RER Misalignment During the Conflict Cycle

The concept of RER equilibrium most commonly used in the literature is based on a single-equation reduced-form behavioral model that accounts for current account flow variables as well as factors influencing longer-run stock equilibrium (see, for example, Edwards 1989; Elbadawi 1994; Elbadawi and Soto 1997a, b; Baffes, Elbadawi, and O'Connell 1999; Clark and MacDonald 1999; and Edwards and Savastano 1999). The underlining notion of RER equilibrium in this approach is intertemporal, as the path of the equilibrium RER is influenced not only by the current value of fundamental determinants but also by expectations about the future evolution of these variables.

An empirical model for the RER that is a close version of the micro-founded general equilibrium RER model developed by Elbadawi and Soto (2005) is specified and estimated here. This model reflects the influence of standard RER fundamentals, including aid, as reflected by the following equation:

$$\begin{aligned}\log(\mathrm{RER})_{it} = {} & \beta_{0i} + \beta_1 \underset{(+)}{\log(\mathrm{TOT})_{it}} + \beta_2 \underset{(+)}{\log(\mathrm{PROD})_{it}} \\ & + \beta_3 \underset{(-)}{\mathrm{OPEN}_{it}} + \beta_4 \underset{(+)}{\log(\mathrm{GOV})_{it}} \\ & + \beta_5 \underset{(+)}{\log(\mathrm{TAX})_{it}} + \beta_6 \underset{(+)}{\mathrm{AID}_{it}} + \beta_7 \underset{(+)}{\mathrm{NFI}_{it}} + \varepsilon_{it}\end{aligned} \tag{1}$$

where subscripts i and t represent country and time indexes; β_{0i} is the country-specific intercept; TOT is the terms of trade; PROD is the ratio of per capita income in country i to the average in OECD countries (a proxy for sector productivity differences); OPEN is a measure of trade openness (computed as the residual from a regression of trade as a share of GDP on geographic variables); GOV is the share of government consumption in GDP; TAX is the share of taxes on nontraded goods in GDP; AID is the GDP share of foreign aid net of

changes in foreign exchange reserves; NFI is the share of net foreign income in GDP; and ε_{it} is a disturbance term.

The expected signs of RER determinants, noted below the corresponding variables in equation 1, are consistent with theory and earlier studies. The model predicts the equilibrium RER to be higher with better terms of trade, higher productivity in the traded goods sector relative to the nontraded sector, less trade openness, higher government consumption, higher taxes on nontraded goods, more aid, and larger net foreign income.

An error-correction version of equation 1 is estimated for a panel comprising annual data for 83 countries (including 36 postconflict countries) for 1980–2004.[11] (Data definitions and sources are reported in the appendix.)

Three econometric estimation methods appropriate for an error-correction specification are applied to the panel data: pooled mean group, mean group, and dynamic fixed-effects estimators.[12] The pooled mean group estimator imposes the restriction that all countries share the long-run coefficients; the more general mean group model assumes that economies differ in their short- and long-run parameters. The pooled mean group estimator is more general than the dynamic fixed-effects estimator, which assumes that all parameters are constant across countries, except for the intercept, which is allowed to vary across countries. The choice between the three estimators entails a tradeoff between consistency and efficiency. The dynamic fixed-effects estimator dominates the other two in terms of efficiency if the restrictions are valid.[13] If they are not valid, the dynamic fixed-effects will generate inconsistent estimates and be dominated by the pooled mean group and mean group estimates. For this study the pooled mean group offers the best compromise between consistency and efficiency, because one would expect the long-run path of the RER to be driven by a similar process across countries while the short-run dynamics around the long-run equilibrium path may differ from one country to another because they are likely to be driven by idiosyncratic news and shocks to fundamentals.

11. Unlike section II, in which the sample comprises only postconflict countries, section III presents the results of an empirical estimation of the equilibrium RER for both postconflict and peaceful countries. The country sample size and period coverage are therefore determined by availability of time series data. Every country must have at least 20 annual observations in order to be included; data are missing for many countries before 1980.

12. Recent literature (for example, Pesaran, Shin, and Smith 1999) shows that the existence of a long-run relationship in equation 1 is not contingent on cointegration. Because right-hand-side variables can combine stationary and nonstationary variables, the equation can be embedded in a dynamic error-correction model. Pooled mean group estimation hence does not require pretesting for unit roots and cointegration. All variables in equation 1 were constructed as index numbers, trend deviations, or shares, implying that they are stationary in the long run.

13. The mean group estimator is derived from the fully heterogeneous coefficient model, which imposes no cross-country parameter restrictions and can be estimated on a country-by-country basis, provided that the time-series dimension of the data is sufficiently large. When the cross-country dimension is also large, the mean of short- and long-run coefficients across countries can be consistently estimated by the unweighted average of the individual country coefficients, which is the mean group estimator proposed by Pesaran, Smith, and Im (1996).

TABLE 3. Short- and Long-Run Determinants of RER, 1980–2004

Variable	Pooled mean group (1)	Mean group (2)	Hausman test	Dynamic fixed effects (3)
Long-run coefficients				
Terms of trade (in logs)	0.2082*** (0.031)	0.6880 (1.038)	0.7400 (0.390)	0.0783 (0.055)
Productivity (in logs)	0.5184*** (0.017)	0.3150*** (0.113)	3.3000 (0.070)	0.5596*** (0.040)
Trade openness (in logs)	−0.5578*** (0.031)	−0.3370* (0.193)	1.3400 (0.250)	−0.4543*** (0.047)
Government consumption/GDP (in logs)	2.6253*** (0.258)	−23.4960 (23.438)	1.2400 (0.270)	0.2271 (0.399)
Net foreign income/GDP	0.0037** (0.002)	0.0710 (0.072)	0.8700 (0.350)	0.0130*** (0.003)
Foreign aid net of international reserves accumulation/GDP	0.0020** (0.001)	−0.0070 (0.005)	3.4100 (0.060)	0.0017 (0.002)
Taxes on nontraded goods (in logs)	2.0308*** (0.403)	8.5500 (7.317)	0.8000 (0.370)	1.1015 (0.730)
Error correction coefficient	−0.2040*** (0.022)	−0.6990 (0.041)		−0.2509*** (0.013)
Short-run coefficients				
Δ (terms of trade, logs)	0.0820** (0.037)	−0.0280 (0.035)		0.0093 (0.022)
Δ (productivity, logs)	0.4290*** (0.029)	0.1950*** (0.043)		0.3918*** (0.020)
Δ (net foreign income/GDP)	0.0040** (0.002)	0.0020 (0.003)		0.0014 (0.001)
Δ (taxes on nontraded goods, logs)	0.4470 (0.470)	−0.4860 (0.489)		1.0929*** (0.317)
Intercept	0.0210 (0.029)	0.0140 (0.099)		

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Note: Number of countries = 83; number of observations = 875. Numbers in parentheses are standard errors, except in the column displaying the Hausman test, where they are p-values. All estimates control for country and time effects. Dynamic specification is autoregressive distributed lag (1,1,1,0,0,1,0,1).

$\Delta X_t = X_t - X_{t-1}$

Source: Authors' analysis based on data described in the text.

The restriction of the pooled mean group against the mean group model can be tested by performing Hausman tests (table 3). The null hypothesis of equality of coefficients cannot be rejected at the 1 percent level, except for productivity and foreign aid (net of international reserves), for which the null hypothesis can be rejected at the 10 percent level. This evidence favors the pooled mean group model against the mean group estimator. The dynamic fixed-effects results confirm in part those obtained by the pooled mean group estimator. However, for the reason discussed above, the pooled mean group is favored over the dynamic fixed-effects method. The rest of the discussion therefore focusses on the pooled mean group findings.

The pooled mean group results are consistent with those reported in other studies.[14] All long-run coefficient estimates are highly significant (at the 1 and 5 percent significance levels) and display the expected signs. With respect to the speed of adjustment and the short-term elasticities, the pooled mean group results suggest that several fundamentals (the terms of trade, relative productivity, and net foreign income) have highly significant short-run effects on the RER. The estimated adjustment parameter (−0.20) is equal to that obtained by Edwards (1989) using a partial adjustment model for 12 developing countries. The adjustment parameter estimate reported here implies that the half-life of an RER deviation from equilibrium is about three years. This finding is consistent with the three- to five-year consensus range of international estimates of the half-life of deviations of the RER from purchasing power parity (Rogoff 1996; Cashin and McDermott 2003).

The results show that higher long-term foreign aid (net of international reserve accumulation) contributes significantly to RER appreciation.[15] In contrast, short-term changes in aid do not have significant effects on the short-term behavior of the RER in preliminary results obtained for the three methods. Aid was therefore omitted from the short-run variables included in the results reported in table 3.

Based on the model results and the results of Elbadawi, Kaltani, and Soto (2006), indexes can be constructed for the equilibrium RER (ERER) and RER misalignment. The ERER index is obtained from the estimated model using permanent components of the fundamental RER determinants (estimated using the Hodrick–Prescott procedure) substituted into the estimated RER equation. The permanent components of the fundamentals are characterized as sustainable levels and are therefore consistent with the concept of RER equilibrium. The ERER index is normalized (through the country-specific intercept) so that

14. Comparable findings in the literature include Chinn (1997) for productivity; Elbadawi and Soto (1997b) and Drine and Rault (2004) for the terms of trade; Elbadawi and Soto (1997b) and Maeso-Fernandez, Osbat, and Schnatz (2002) for government consumption; and Elbadawi and Soto (2005) for most variables.

15. Gross foreign aid (without deducting reserve accumulation) is a less robust determinant of the RER than aid net of accumulation of reserves. This finding confirms that whether aid is spent or saved affects the behavior of the RER.

TABLE 4. RER Misalignment and Its Decomposition in Postconflict Countries (average values, in percent)

Variable	(1) RER misalignment (2) + (6)	(2) Deviation of fundamentals from long-term trend (3) + (4)	(3) Policy variables[a]	Structural variables: (4) All structural variables[b]	Structural variables: (5) Aid net of international reserve accumulation/ GDP	(6) Error correction term
PrePeace	−5.789	−2.839	−1.396	−1.443	−0.148	−2.950
PeaceOnset	7.788	−0.247	0.422	−0.669	0.277	8.035
PostConf1	2.490	0.403	0.613	−0.210	0.044	2.087
PostConf2	−4.744	−0.970	−0.556	−0.413	−0.106	−3.775

Note: RER misalignment and its decomposition are based on the underlying data of column 1 in table 3 and their coefficients.

[a]Policy variables are government consumption/GDP, taxes on nontraded goods, and trade openness.

[b]Structural variables are terms of trade, productivity, net foreign income/GDP, and aid net of international reserve accumulation as a share of GDP.

Source: Authors' analysis based on data described in the text.

the average RER misalignment (the mean logarithmic difference between the RER and ERER for each country) is set equal to zero. The log of the resulting normalized ERER is then subtracted from the log of the RER to obtain the RER misalignment time series for each country.[16]

Setting the average RER misalignment at zero for the full 25-year sample period imposes the plausible identification condition that no country is allowed to exhibit RER misalignment for the full 25-year sample period. This identification restriction is justified by the long-term equilibrium model for the RER, which nests temporary RER disequilibria with a reported half-life of three years. This relatively rapid correction of temporary RER misalignments suggests that several episodes of RER over- and undervaluation are likely to be observed over the 25-year sample period in any country.

The average RER misalignment and its decomposition according to the contribution of the deviations of fundamentals from their trend values and of short-term dynamics are reported in table 4. The results are based on the methodology discussed above applied to the pooled mean group results for the RER equation presented in the first column of table 3. Observed RER misalignment (column 1) is broken down into two components: RER disequilibrium caused by the deviation of fundamentals from their long-term trends (column 2) and RER disequilibrium caused by short-term shocks in fundamentals (labeled "error correction term," column 6). The combined contribution of deviations of fundamentals from long-

16. Elbadawi, Kaltani, and Schmidt-Hebbel (2007) provide a detailed description of the methodology for computing RER equilibrium levels and misalignment measures.

term trends is divided into the contribution of the three policy fundamentals (column 3) and the four structural fundamentals (columns 4 and 5).

The RER misalignment estimates suggest that the average conflict country exhibited a moderately undervalued (−5.8 percent) exchange rate during *PrePeace*. After peace was attained, the average postconflict country experienced an RER overvaluation of 7.8 percent during *PeaceOnset*, which declined to 2.5 percent during *PostConf1*. During *PostConf2* the RER became undervalued again (by −4.7 percent).

It is not surprising that short-term changes in fundamentals (reflected by the error correction term) explain most of the deviations from equilibrium RER levels. With the sole exception of the *PrePeace* period, deviations of fundamentals from their long-term trend levels contribute little to RER misalignment in each conflict-cycle period. Among the seven fundamentals, aid deviations from trend values contribute particularly little to misalignment, explaining only 0.1 percentage points of the observed RER undervaluation of 2.8 percent and 0.3 percentage points of the observed RER overvaluation of 7.8 percent during *PrePeace* (*PeaceOnset)* period and even less in the subsequent postconflict periods. The minor contribution of aid to RER disequilibrium over the conflict cycle results from two factors: the small to moderate changes in aid observed over the conflict cycle (documented in section I) and the relatively small magnitude of the long-run coefficient of aid in the RER estimation (reported in table 3).

III. Growth During the Postconflict Cycle

A general result from the cross-country econometric research on aid effectiveness is that aid does not influence economic performance, at least not unambiguously. Some research (World Bank 1998; Burnside and Dollar 2000) shows that aid to countries with good policy environments may increase growth but that its effect is subject to diminishing returns. In evaluating government use of aid, Barder (2006) notes that limits on aid absorption may explain why the costs of aid increase faster than its benefits. Collier and Hoeffler (2002) find that aid is more effective in postconflict transition than in normal peaceful environments. Roodman (2007), who tests the robustness of empirical studies on the effect of aid on growth, concludes that most results appear fragile to changes in model specification, especially sample expansion.

A strand of the empirical growth literature that focusses on the impact of RER misalignment finds a robust negative association between RER overvaluation and growth. While moderate undervaluation may support growth, extreme undervaluation may reduce it (Razin and Collins 1997; Aguirre and Calderón 2005). This literature also finds that the growth impact of RER overvaluation and volatility depends on the level of financial development, with overvaluation generally reducing growth in countries with low levels of financial development

while having no significant effect in financially advanced economies (Aghion and others 2006).

An empirical growth model is specified here that nests the four strands of the growth literature relevant to postconflict growth—conflict periods, aid, RER competitiveness, and financial development—as well as their possible interactions. Controlling for conventional growth determinants that are robustly identified in the empirical cross-country growth literature, the model addresses the following questions, which have not been addressed in the aid and postconflict transition literature:

- What are the separate effects of peace and postconflict periods, aid, RER misalignment, and financial development on growth?
- Does aid reduce or augment the impact of RER misalignment on growth?
- Does aid increase the contribution of postconflict peace to growth?
- Is the growth loss from RER misalignment ameliorated by financial development?

To answer these questions, the following growth model is specified and estimated for five-year periods:

$$\begin{aligned} y_{it} = {} & \underset{(+)}{\beta_0 A_{it}} + \underset{(-)}{\beta_1 A_{it}^2} + \underset{(-)}{\beta_2 RERMIS_{it}} + \underset{(+)}{\beta_3 PeaceOnset_{it}} \\ & + \underset{(+)}{\beta_4 PostConf1_{it}} + \underset{(+)}{\beta_5 PostConf2_{it}} + \underset{(+)}{\beta_6 FD_{it}} \\ & + \underset{(?)}{\beta_7 A_{it} * \mathrm{RERMIS}_{it}} + \underset{(?)}{\beta_8 A_{it} * \mathrm{PeaceOnset}_{it}} + \underset{(?)}{\beta_9 A_{it} * \mathrm{PostConf1}_{it}} \\ & + \underset{(?)}{\beta_{10} A_{it} * \mathrm{PostConf2}_{it}} + \underset{(+)}{\beta_{11} RERMIS_{it} * \mathrm{FD}_{it}} + \beta_{12} CV_{it} + \mu_t + \eta_i + \varepsilon_{it} \end{aligned}$$

(2)

where i and t are country and time indexes; y is per capita GDP growth; A is aid as a share of GDP;[17] RERMIS is RER misalignment; *PeaceOnset, PostConf1, and PostConf2* are the postconflict period dummy variables;[18] *FD*

17. Aid is measured at the beginning of each five-year period, as recipient countries tend to absorb and spend aid with some time lag. Start-of-period aid is also more likely to be exogenous, because aid is partially responsive to recipient countries' past economic performance, including past growth. Aid is set at zero for developed countries.

18. Defining the three postconflict periods in the growth sample, which consists of five-year averages, calls for some discretion. The rule of thumb used here was to code a period in accordance with the number of years in each group (PeaceOnset, PostConf1, PostConf2) that has a majority. When a five-year period consists of one year of war, two years of PeaceOnset, and 2 years of PostConf1, for example, the period is coded as PeaceOnset. A five-year period that consists of two years of war, two years of PeaceOnset, and one year of PostConf1 is coded as PeaceOnset.

is a measure of financial development; CV is a set of standard control variables that are robustly associated with cross-country growth (initial per capita GDP, initial per capita GDP cyclical component, inflation, government expenditure as a share of GDP, human capital investment as a share of GDP, a rule of law index, and trade openness);[19] and μ_t and η_i are time and country fixed effects.

For the individual variables the expected signs of the coefficients of growth determinants, noted below the corresponding variables in equation 2, are consistent with theory and earlier studies. The interaction effect between aid and RER misalignment (or overvaluation) on growth could have either sign, depending on the relative allocation of aid to traded and nontraded sector activities and on sector differences between rates of return during periods of RER overvaluation. Growth effects of interactions between aid and the three postconflict period dummy variables could also go either way, depending on the country-specific scarcity of funds and absorptive capacity of aid. Finally, one would expect a positive growth effect from the interaction between RER misalignment and financial development, because financial markets are more likely to supply bank loans and hedging instruments (such as foreign exchange derivatives) when they are deeper, ameliorating the negative direct impact of RER overvaluation.

Policy fundamentals are likely to be jointly determined with growth and responsive to expected future growth performance. The generalized method of moments system dynamic panel estimation method is therefore used, in order to properly account for endogeneity and country-specific unobserved characteristics. This system (developed in Arellano and Bover 1995 and Blundell and Bond 1997) uses lagged values of the dependent and independent variables as instruments and combines regressions in differences with regressions in levels to better address the issue of weak instrumentation. Under the assumed moment conditions, the generalized method of moments system accounts for the combined problems of potential endogeneity and unobserved country effects.

The data panel comprises 77 countries, including both conflict and nonconflict developing economies as well as industrial countries, for five-year nonoverlapping averages spanning 1970–2004. The regression and model consistency test results are reported in table 5. The first column reports the results of estimating equation 2 excluding interaction terms; the second column reports the results using the full specification, including interaction terms; the third column reports a more parsimonious regression that excludes nonsignificant variables. Both the Sargan tests for overidentifying restrictions and the serial-correlation tests for error terms, reported at the bottom of table 5, validate the specification. Six of the seven standard growth fundamentals exhibit the expected sign and are statistically very significant in the three regression results (the exception is trade openness, which is dropped in the last regression).

19. Loayza and Soto (2002) provide a comprehensive review of the evidence on the empirical relevance of these standard growth correlates.

TABLE 5. Economic Growth and the Role of Aid, RER Misalignment, and Financial Development

Variable	(1)	(2)	(3)
Aid/GDP	0.2675*** (0.030)	0.1987** (0.043)	0.2738*** (0.031)
Aid/GDP squared	−0.0042*** (0.001)	−0.0024** (0.001)	−0.0037*** (0.001)
RER misalignment	−0.0292*** (0.004)	−0.0277* (0.015)	−0.0164* (0.009)
Peace onset	−0.0066 (0.004)	−0.0056 (0.008)	−0.0030 (0.004)
Postconflict 1	0.0196*** (0.004)	0.0150** (0.006)	0.0378*** (0.006)
Postconflict2	−0.0282*** (0.007)	−0.0342** (0.009)	−0.0273*** (0.008)
Financial development (in logs)	−0.0088*** (0.002)	0.0058** (0.002)	0.0064*** (0.002)
Interactions			
RER misalignment*aid/GDP		−0.2919** (0.134)	−0.3139*** (0.114)
Aid/GDP*peace onset		0.0941* (0.057)	
Aid/GDP*postconflict1		0.1022** (0.041)	
Aid/GDP*postconflict2		0.2963 (0.225)	
RER misalignment*financial development		0.0095* (0.005)	0.0054* (0.003)
Standard Control Variables			
Initial GDP per capita (in logs)	−0.0083*** (0.002)	−0.0085** (0.002)	−0.0073*** (0.002)
Initial GDP per capita cyclical component	−0.1619*** (0.019)	−0.1790** (0.019)	−0.1836*** (0.014)
Inflation (in logs)	−0.0165*** (0.002)	−0.0080** (0.004)	−0.0184*** (0.002)
Government expenditures/ GDP (in logs)	−0.0286*** (0.003)	−0.0230** (0.004)	−0.0350*** (0.004)
Human capital investment (in logs)	0.0.2233*** (0.004)	0.0269** (0.004)	0.0217*** (0.004)
Rule of law	0.0160*** (0.001)	0.0159** (0.002)	0.0184*** (0.002)
Trade openness (in logs)	−0.0011 (0.003)	0.0052 (0.004)	
Period shifts			
Intercept (base period: 1975–79)	−0.0543***	−0.1116***	−0.0669***
1980–84	−0.0172***	−0.0197***	−0.0198***
1985–89	−0.0199***	−0.0216***	−0.0201***
1990–94	−0.0240***	−0.0274***	−0.0258***
1995–99	−0.0298***	−0.0358***	−0.0357***
2000–04	−0.0201***	−0.0275***	−0.0245***
Specification tests (*p*-values)			
Sargan test	0.61	0.72	0.38
Second-order serial correlation	0.23	0.12	0.29

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Note: Table displays cross-country panel data results, based on nonoverlapping five-year averages spanning 1970–2004, estimated using the generalized method of moments–IV system estimator. Dependent variable is growth rate of real GDP per capita. Number of countries = 77; number of observations = 367. Numbers in parentheses are robust standard errors.

Source: Authors' analysis based on data described in the text.

The baseline regression results reported in column 1 show that aid is positively but nonmonotonically associated with growth. The positive and significant effect of the linear aid term is increasingly offset by the negative and significant quadratic aid term, implying declining marginal growth returns of aid. RER misalignment (appreciation) has a significant negative effect on growth. There is evidence of a direct peace dividend for growth in *PostConf1* but neither before (in *PeaceOnset*) nor after (in *PostConf2*). In fact, growth declines significantly in *PostConf2*. This strong result suggests that there is significant but temporary catch-up growth some years after peace is attained but that growth subsequently relapses to prepeace levels, after controlling for all other growth determinants. Financial development has a small and significant but negative effect on growth; subsequent results overturn this result, however.

The full specification in column 2 allows for key interaction terms among variables. Signs, sizes, and significance levels of the coefficients of individual variables of interest (the first seven variables in table 5) are similar to those reported in column 1. The one important difference is financial development, which now exhibits a positive and significant contribution to growth.

While a priori the interaction term between RER misalignment and aid could have either sign, it is empirically found to have a negative and significant effect on growth. Therefore aid reduces growth in a macroeconomic environment of exchange rate misalignment (overvaluation), controlling for its other effects captured in this specification. The negative impact of RER misalignment (both directly and when interacted with aid) is ameliorated by financial development, as demonstrated by the positive and significant coefficient of the interaction between RER misalignment and financial development. Aid has a positive impact on growth when disbursed during the three postconflict periods, reflected by the corresponding interaction effects between aid and period dummies; however, its effect is significant at the 5 percent level only during *PostConf1*.

For robustness purposes, a parsimonious version of the estimation equation (reported in column 3) removes several variables that were not robustly significant in columns 1 and 2. Significance levels and coefficient sizes of single or direct effects of several variables of interest are reinforced under this specification. The negative interaction between RER misalignment and aid is more significant, while the positive interaction between RER misalignment and financial development is smaller.

These results provide robust evidence on the questions posed at the beginning of this section. Growth is significantly higher in *PostConf1* and significantly lower in *PostConf2*. Aid affects growth positively, but its effect is characterized by diminishing returns. RER misalignment has a detrimental effect on growth. Financial development contributes positively to growth.

The effects of individual variables of interest to postconflict growth (complemented by the contributions of six other standard robust growth determinants)

are extended by interaction effects between key variables. There is some evidence that aid disbursed after attainment of peace has higher growth returns, but the evidence is not entirely robust. There is, however, significant evidence that the interaction of RER misalignment and aid has a negative effect on growth. This may reflect the fact that more aid contributes more strongly to nontraded activities with lower growth returns. The level of aid also contributes to RER appreciation, and its deviation from trend contributes marginally to RER misalignment, as documented in section II. Hence, there is also a small direct effect of aid on RER misalignment, which strengthens the interaction effect on growth. The fact that the interaction between RER misalignment and financial development has a positive effect on growth may confirm the notion that deeper financial markets provide better protection to (traded-goods) firms against RER overvaluation.

The results of column 3 can be used to probe more deeply into the analysis of the impact of RER misalignment on growth. From equation 2 the overall growth effect of a change in RER misalignment is given by the following expression:

$$\Delta Y = (\beta_2 + \beta_7 A + \beta_{11} FD)\Delta RERMIS \tag{3}$$

where β_2 is the direct effect of RER misalignment on growth and $\beta_7 A$ and $\beta_{11} FD$ are the nonlinear effects of RER misalignment on growth that depend on the levels of aid and financial development. The estimated parameters imply that a one standard deviation increase in RER misalignment (that is, a one standard deviation RER appreciation for a given RER equilibrium level) leads to a 0.35 percent decline in economic growth. In addition, if aid and financial development are held at their median sample values, the indirect growth effects from the interactions between RER misalignment and the levels of aid and financial development lead to an additional growth loss of 0.2 percent and a growth gain of 0.4 percent.

Many postconflict countries are highly dependent on aid and have shallow financial markets. Two simple simulation exercises are conducted, based on the regression result, to reflect this reality and obtain more-relevant results. For the first exercise the effects of a one standard deviation increase in RER misalignment on growth are simulated at different levels of aid dependency. Second, the effects of a one standard deviation increase in RER misalignment on growth are estimated at different levels of financial development.

For the first exercise, three levels of aid—the 25th, 50th, and 75th percentile of the distribution of aid among developing countries in the sample—are examined, holding the level of financial development at its median sample value (table 6). High aid levels can be the most detrimental to growth in the presence of RER overvaluation. Ethiopia, Guinea-Bissau, Jordan, Kenya, Mali, Mozambique, Senegal, Sierra Leone, the Syrian Arab Republic, and Uganda

TABLE 6. Impact of One Standard Deviation Increase in RER Misalignment on Growth at Different Levels of Aid and Financial Development, 1975–2004 (in percent)

	Variable level		
Variable	Low	Medium	High
Aid	−0.02	−0.17	−0.53
Financial development	−0.23	−0.17	−0.10

Note: A low level of any variable represents the 25th percentile of the distribution of aid or financial development in the sample of developing countries; a medium level, the 50th percentile; and a high level, the 75th percentile.

Source: Authors' analysis based on data described in the text.

experienced aid levels at or above the 75th percentile. The level of aid plays a major role in the growth impact of RER misalignment: growth declines by 0.02 percent under a low-aid scenario and 0.53 percent under a high-aid scenario. Growth reduction would be even larger in many highly aid-dependent countries, where aid represents close to 20 percent of GDP.

For the second exercise the growth impact of the same increase in RER misalignment is estimated at different levels of financial development while holding aid at its median sample value for developing countries. Postconflict countries differ widely in their financial market development. Many African countries have very shallow financial markets, as captured by the ratio of private credit to GDP. The Republic of Congo, the Democratic Republic of Congo, Ethiopia, the Gambia, Guatemala, Mali, Mozambique, Nigeria, Papua New Guinea, Sierra Leone, Sri Lanka, Sudan, Syria, and Uganda experienced financial development levels at or below the 25th percentile. For 2000–04, financial development in Burkina Faso, Cameroon, and Tanzania is lower than that of the sample's 25th percentile, while it is close to that of the 75th percentile in Egypt and Morocco. The results suggest that financial development mitigates some of the detrimental effects of overvaluation. Growth in countries in the 75th percentile in financial market development declines 0.10 percent as a result of a one standard deviation increase in RER misalignment. In contrast, growth falls 0.23 percent in countries in the 25th percentile in response to the same increase in RER overvaluation. This figure may seem relatively small, but when considering the unique features of postconflict countries—which have greater variation in RER misalignment and less financial development than other countries—the repercussions on growth are substantially larger. In the Republic of Congo, Guinea-Bissau, or Niger, for example, the decline in growth from RER disequilibrium appreciation could be as high as 0.5 percent. Financial sector reform could both foster economic development directly and make growth more resilient to RER overvaluation.

IV. Conclusions

New empirical evidence on the time pattern and behavior of foreign aid, the RER, and growth suggests that aid rises significantly when peace is attained in the average conflict country but declines in the postconflict years. The RER shifts from being undervalued during civil war to being overvalued in the first few years of peace and returning to equilibrium thereafter, but these changes are not statistically significant. Per capita GDP growth rises significantly after the year peace is attained and for the full postconflict period.

Simple correlations—performed separately for nonconflict and conflict countries—reveal several patterns of associations among these variables. Aid and RER misalignment (or overvaluation) are not positively and significantly correlated, providing prima facie evidence against exchange rate overvaluation when the level of aid is high. The simple correlation between aid and growth is negative and significant in most cases, suggesting that the relation between aid and growth is not positive. Growth and RER misalignment are negatively and significantly correlated only during the peace years in conflict countries, hinting that RER overvaluation could undermine output recovery during peace years.

Simple correlations may be highly misleading, however, because they are not informative about causality, do not control for third variables, and do not consider nonmonotonic relations and interactions with other variables. For these reasons this study focusses on the behavior of the RER and growth in a multivariate framework, identifying their response to variables, particularly aid, that are crucial during the conflict cycle.

The RER misalignment data are generated from error-correction estimations for the equilibrium RER, based on structural determinants, including foreign aid. The empirical results for the model—based on a sample of annual 1980–2004 data for 83 countries—show that long-run coefficients of all structural variables and short-run coefficients of some structural variables are significant and display the signs predicted by theory. Higher long-term foreign aid contributes significantly to long-term RER appreciation; short-term changes in aid do not have significant effects on short-run RER behavior.

Regression results are used to compute RER misalignment series for each country and for the average conflict country over the four periods of its conflict cycle. RER misalignment is decomposed based on short-term dynamics and the contribution of deviations of fundamentals from their long-term trends. Not surprisingly, short-term shocks explain most of the RER deviations from equilibrium levels. Deviations of fundamentals from their long-term trend levels contribute little to RER misalignment in each conflict-cycle period. Aid deviations from trend values contribute particularly little to misalignment.

This empirical evidence on growth provides new insights into the pattern of growth following conflict and the role of key determinants of postconflict growth (conflict periods, aid, RER misalignment, and financial development), both individually and through their interactions. The growth equation nests

these variables within a standard specification, controlling for growth fundamentals that are robustly identified in the empirical growth literature.

Empirical estimations, based on the dynamic system generalized method of moments estimator, were performed on a panel sample of 77 conflict and nonconflict countries for five-year averages spanning 1970–2004. The results indicate that aid initially contributes positively to growth but reduces growth (as captured by a quadratic specification) at higher levels, confirming declining growth returns to aid. RER misalignment (overvaluation) reduces growth. Some years after peace, a peace dividend is attained, but it turns negative at the end of the postconflict cycle.

These results are complemented by important interaction effects. There is some evidence that aid disbursed after attainment of peace promotes higher growth than during conflict or at the very end of the postconflict period. Other interaction effects suggest that the negative growth impact of RER misalignment is intensified by aid and weakened by financial development. Hence aid reduces growth in a macroeconomic environment that allows for RER overvaluation. Deeper financial markets provide better protection to (traded-goods) firms against RER overvaluation, possibly by supplying more credit or offering hedging instruments against RER risk.

These results provide empirical support for several channels of transmission from aid to growth. The linear positive effect of aid on growth supports the notion that recipient countries with access to foreign resources are likely to use aid to finance investment, improve policies, and raise aggregate efficiency. The negative nonlinear effect shows that aid has decreasing growth benefits, reflecting the growing misuse of or weakening capabilities of absorbing larger aid inflows. By showing that very little postconflict RER appreciation can be traced to exceptionally large aid inflows, the results suggest that aid has no indirect negative effect on growth through RER misalignment, an effect that is potentially very relevant to postconflict countries. Postconflict aid thus does not appear to contribute to Dutch disease-type RER misalignment and lower growth. Aid does interact with other variables in affecting growth: the timing of aid matters, with aid reducing growth when it interacts with RER overvaluation.

These results lead to some policy inferences for the management of aid and the conduct of macroeconomic and structural policies in the years following the attainment of peace. Peaks of aid inflows at peace onset and shortly thereafter should be avoided. Spreading aid commitments out over time brings actual aid flows closer to permanent aid flows, so that spending aid is more consistent with permanent aid resources. Building up the quality of policies and institutions is a gradual process after the onset of peace. To maximize effectiveness, aid flows should also increase gradually. Avoiding excessive government spending raises growth both directly and indirectly (by reducing RER overvaluation). These findings suggest the need to strengthen international competitiveness through appropriate macroeconomic policies.

In the realm of structural policies the results suggest that financial development and deepening have direct positive effects on long-term growth and an additional growth bonus by reducing the adverse growth impact of RER overvaluation. Hence financial development should be high on the priority of postconflict policies. This could include strengthening domestic banking, supporting the development of domestic capital markets, and promoting the development of financial instruments to protect against exchange rate risk.

Funding

Klaus Schmidt-Hebbel thanks Fondecyt for support provided by research project 1060175.

Appendix

Table A-1. Definitions and Sources of Variables Used in Regression Analysis

Variable	Definition and construction	Source
ERER estimation		
Aid	Ratio of official development assistance to GDP (both in current US$)	*World Development Indicators* (World Bank 2006)
Real effective exchange rate	Nominal effective exchange rate (a measure of the value of a currency against a weighted average of several foreign currencies) divided by price deflator or index of costs	*International Financial Statistics* (IMF 2006), supplemented with internal data from IMF Information Notice System
Terms of trade	National accounts exports price index divided by imports price index	*World Development Indicators* (World Bank 2005, 2006); Loayza, Fajnzylber, and Calderón (2005)
Postconflict episodes	Prepeace: years −4 to −1 before the end of conflict, denoted as time 0; peace onset: years 0 and 1; postconflict1: years 2–5; postconflict2: years 6–9	Authors' construction using conflict data from Sambanis (2004), complemented with data from PRIO armed conflict data set (PRIO 2005) for 2002–04
Inflation	As measured by consumer price index	*World Development Indicators* (World Bank 2006)
Productivity	Ratio of per capita GDP at factor cost to average GDP at factor cost for industrial countries, both in US$	Authors' construction based on data from *World Development Indicators* (World Bank 2005)

(Continued)

TABLE A-1. Continued

Variable	Definition and construction	Source
Openness	Residual of regression of log of ratio of exports and imports to GDP (both in current local currency) on logs of area and population and dummies for oil-exporting and landlocked countries	*World Development Indicators* (World Bank 2004); Loayza, Fajnzylber, and Calderón (2005)
Government consumption	Ratio of government consumption to GDP (both in current local currency)	Authors' construction based on data from *International Financial Statistics*, December 2004 CD-ROM Release (IMF 2004); *World Development Indicators* (World Bank 2004); and Africa Live Database (World Bank 2004)
Net foreign income	Net labor income and net property and entrepreneurial income	*World Development Indicators* (World Bank 2006)
Taxes on nontraded goods	Log of 1 plus the ratio of tax revenue to GDP	Authors' construction based on data from *World Development Indicators* (World Bank 2004), Africa Live Database (World Bank 2004), and *Government Finance Statistics* (IMF 2004)
Growth estimation per capita real GDP growth	Log difference in per capita real GDP	Authors' construction based on data from *World Development Indicators* (World Bank 2006)
Initial GDP per capita	Initial value of ratio of total real GDP to total population	Authors' construction based on data from *World Development Indicators* (World Bank 2006)
Initial GDP per capita cyclical component	Difference between log of actual GDP per capita and log of potential (trend) GDP; Hodrik–Prescott filter used to estimate trend GDP	Authors' construction based on data from *World Development Indicators* (World Bank 2006)
Inflation	Log of 100 plus inflation rate	Author's construction based on data from *World Development Indicators* (World Bank 2006)
Government expenditure	Ratio of government expenditures to GDP (both in current local currency)	Data from *International Financial Statistics* (IMF 2006), complemented by data from *World Development Indicators* (World Bank 2006) and *UN National Accounts Statistics* (United Nations 2006)
Human capital investment	Ratio of total secondary enrollment, regardless of age, to population of age group that officially corresponds to level of education	Easterly and Sewadeh (2002), *World Development Indicators* (World Bank 2006), and UNESCO (2006)

(Continued)

TABLE A-1. Continued

Variable	Definition and construction	Source
Rule of law	Presence of law and order on scale of 0–6	International Country Risk Guide (ICRG), Political Risk Services (www.icrgonline.com)
Trade openness	Ratio of exports and imports to GDP (both in current local currency)	Data primarily from *International Financial Statistics* (IMF 2006), complemented by data from *World Development Indicators* (World Bank 2006) and *UN National Accounts Statistics* (United Nations 2006)
RER misalignment	Percentage difference between real effective exchange rate and estimated equilibrium value	Authors' calculation based on method presented in Elbadawi, Kaltani, and Schmidt-Hebbel (2007)
Postconflict episodes	Coded based on number of years pertaining to each group (PeaceOnset, PostConf1, PostConf2) that dominate during the five-year period. When five-year period consists of one year of war, two years of PeaceOnset, and two years of PostConf1, it is coded as PeaceOnset. When there are two years of war, two years of PeaceOnset, and one year of PostConf1 within a five-year period, the period is coded as PeaceOnset.	Authors' construction using conflict data from Sambanis (2004) and data from PRIO (PRIO 2005) armed conflict data set for 2002–04
Financial depth	Ratio of bank credit to private sector to GDP (both in current local currency)	*World Development Indicators* (World Bank 2006)
Period-specific shifts	Time dummy variables	Authors' construction

Source: Authors' compilation.

REFERENCES

Addison, T. ed. 2003. *From Conflict to Recovery in Africa*. New York: Oxford University Press for United Nations University/World Institute for Development Economics Research.

Aghion, P., P. Bacchetta, R. Ranciere, and K. Rogoff. 2006. "Exchange Rate Volatility and Productivity Growth: The Role of Financial Development." Department of Economics, Harvard University, Cambridge, Mass.

Aguirre, A., and C. Caldern. 2005. "Real Exchange Rate Misalignments and Economic Performance." Working Paper. Central Bank of Chile, Santiago.

Aiyar, S., A. Berg, and M. Hussain. 2005. "The Macroeconomic Challenge of More Aid." *Finance and Development* 42(3):2831.

Arellano, M., and O. Bover. 1995. "Another Look at the Instrumental-Variable Estimation of Error-Components Models." *Journal of Econometrics* 68(1):29–52.

Baffes, J., I.A. Elbadawi, and S.A. O'Connell. 1999. "Single-Equation Estimation of the Equilibrium Real Exchange Rate." In L. Hinkle and P. Montiel, eds., *Exchange Rates Misalignment: Concepts and Measurement for Developing Countries*. Washington, D.C.: World Bank.

Barder, O. 2006. "Are the Planned Increases in Aid Too Much of a Good Thing?" Working Paper 90. Center for Global Development, Washington, D.C.

Blundell, R., and S. Bond. 1997. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." University College London Discussion Paper in Economics 97/07. London, UK.

Burnside, C., and D. Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):847–68.

Cashin, P., and C.J. McDermott. 2003. "An Unbiased Appraisal of Purchasing Power Parity." *IMF Staff Papers* 50(3):321–51.

Chinn, M. 1997. *Sectoral Productivity, Government Spending, and Real Exchange Rates: Empirical Evidence for OECD countries*. NBER Working Paper 6017. Cambridge, Mass.: National Bureau of Economic Research.

Clark, P., and R. MacDonald. 1999. "Exchange Rates and Economic Fundamentals: A Methodological Comparison of BEERs and FEERs." In J. Stein, and R. MacDonald, eds., *Equilibrium Exchange Rates*. Amsterdam: Kluwer.

Clemens, M.A., S. Radelet, and R. Bhavnani. 2004. "Counting Chickens When They Hatch: The Short-Term Effect of Aid on Growth." Working Paper 44. Center for Global Development, Washington, D.C.

Collier, P., and A. Hoeffler. 2002. Aid, Policy, and Growth in Post-Conflict Societies. Policy Research Paper 2902. Washington, D.C.: World Bank.

Cottani, J., D. Cavallo, and S. Kahn. 1990. "Real Exchange Rate Behavior and Economic Performance in LDCs." *Economic Development and Cultural Change* 39(1):61–76.

Drine, I., and C. Rault. 2004. "On the Long-Run Determinants of Real Exchange Rates for Developing Countries: Evidence from Africa, Latin America and Asia." Department of Economics, Sorbonne University, Paris.

Easterly, W., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–50.

Easterly, W., and M. Sewadeh. 2002. *Global Development Network Growth Database*. Washington, D.C.: World Bank.

Edwards, Sebastian. 1989. Real Exchange Rate, Devaluation and Adjustment: Exchange Rate Policy in Developing Countries. Cambridge, Mass.: MIT Press.

Edwards, S., and M. Savastano. 1999. *Exchange Rates in Emerging Economies: What Do We Know? What Do We Need to Know?* NBER Working Paper 7228. Cambridge, Mass.: National Bureau of Economic Research.

Elbadawi, I. 1994. "Estimating Long-Run Equilibrium Real Exchange Rates." In J. Williamson ed., *Estimating Equilibrium Exchange Rates*. Washington, D.C.: Institute for International Economics.

Elbadawi, I., L. Kaltani, and K. Schmidt-Hebbel. 2007. "Post-Conflict Aid, Real Exchange-Rate Adjustment, and Catch-Up Growth." World Bank Policy Discussion Paper 4187. Washington, D.C.: World Bank.

Elbadawi, I., L. Kaltani, and R. Soto. 2006. "Real Exchange Rate Misalignment in SSA: How Serious, How Dangerous?" World Bank, Development Economics Research Group, Washington, D.C.

Elbadawi, I., and R. Soto. 1997a. "Capital Flows and Long-Term Equilibrium Real Exchange Rates in Chile." *Revista de Analisis Economico* 12(1):35–62.

———. 1997b. "Real Exchange Rates and Macroeconomic Adjustment in Sub-Saharan Africa and Other Developing Countries." *Journal of African Economies* 6(3):74–120.

———. 2005. "Theory and Empirics of Real Exchange Rates in Sub-Saharan Africa and Other Developing Countries." World Bank, Development Economics Research Group, Washington, D.C.

Ghura, D., and T.J. Grennes. 1993. "The Real Exchange Rate and Macroeconomic Performances in Sub-Saharan Africa." *Journal of Development Economics* 42(1):55–74.

IMF (International Monetary Fund). Government Finance Statistics. Washington, D.C.: International Monetary Fund.

———. 2005. "The Macroeconomics of Managing Increased Aid Inflows: Experiences of Low-Income Countries and Policy Implications." [http://www.imf.org/external/np/pp/eng/2005/080805a.pdf].

———. 2004 and 2006. International Financial Statistics. Washington, D.C.: International Monetary Fund.

Kang, S., and J. Meernik. 2004. "Determinants of Post-Conflict Economic Assistance." *Journal of Peace Research* 41(2):149–66.

Krugman, P. 1987. "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies." *Journal of Development Economics* 27(1):41–55

Loayza, N., P. Fajnzylber, and C. Caldern. 2005. "Economic Growth in Latin America and the Caribbean: Stylized Facts, Explanations, and Forecasts." Central Bank of Chile Working Paper 265. Santiago.

Loayza, N., and R. Soto. 2002. "The Sources of Economic Growth: An Overview." In N. Loayza and R. Soto, eds., *Economic Growth: Sources, Trends, and Cycles*. Series on Central Banking, Analysis, and Economic Policies. Santiago, Chile: Central Bank of Chile.

Maeso-Fernandez, F., C. Osbat, and B. Schnatz. 2002. "Towards the Estimation of Equilibrium Exchange Rates for CEE Acceding Countries: Methodological Issues and a Panel Cointegration Perspective." Working Paper 353. European Central Bank, Frankfurt, Germany.

Pesaran, H., P. Smith, and S. Im. 1996. "Dynamic Linear Models for Heterogeneous Panels." In L. Matyas, and P. Sevestre, eds., *Economics of Panel Data: A Handbook of Theory with Applications*. Dordrecht, The Netherlands: Kluwer.

Pesaran, H., Y. Shin, and R. Smith. 1999. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association* 94(446):621–34.

PRIO (International Peace Research Institute). 2005. UCDP/PRIO Armed Conflict Dataset Version 3.0. [http://new.prio.no/CSCW-Datasets/Data-on-Armed-Conflict/UppsalaPRIO-Armed-Conflicts-Dataset/Armed-Conflicts-Old-Versions/Armed-Conflicts-Version-30/].

Razin, O., and S. M. Collins. 1997. *Real Exchange Rate Misalignments and Growth*. NBER Working Paper 6174. Cambridge, Mass.: National Bureau of Economic Research.

Rodrik, D. 2007. "The Real Exchange Rate and Economic Growth: Theory and Evidence." John F. Kennedy School of Government, Harvard University, Cambridge, Mass.

Rogoff, K. 1996. "The Purchasing Power Parity Puzzle." *Journal of Economic Literature* 34(June):647–68.

Roodman, D. 2007. "The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics." *World Bank Economic Review* 21(2):255–77.

Sachs, J.D., and A.M. Warner. 1995. *Natural Resource Abundance and Economic Growth*. NBER Working Paper 5398. Cambridge, Mass.: National Bureau of Economic Research.

Sambanis, N. 2004. "List of Civil Wars: 1945–1992." [http://pantheon.yale.edu/~ns237/index/research.html#Data]

Schwartz, J., P. Halkyard, and S. Smith. 2006. "Post-Conflict Infrastructure: Trends in Aid and Investment Flows." News Letter Note 305. World Bank, Private Sector Development Vice Presidency, Washington, D.C.

Spatafora, N., and A. Warner. 1999. "Macroeconomic and Sectoral Effects of Terms of Trade Shocks: The Experience of the Oil-Exporting Developing Countries." IMF Working Paper WP134. International Monetary Fund, Washington, D.C.

UNESCO (United Nations Educational, Scientific, and Cultural Organization) Institute for Statistics. 2006. Table 5 Enrolment ratios by ISCED level. [http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx].

United Nations. 2006. National Accounts Statistics. New York: United Nations.

Van Wijnbergen, S. 1984. "The Dutch Disease: A Disease after All?" *Economic Journal* 94(373):41–55.

World Bank. 1998. Assessing Aid. New York: Oxford University Press for the World Bank.

———. 2004, 2005, and 2006. World Development Indicators. Washington, D.C.: World Bank.

Disability, Poverty, and Schooling in Developing Countries: Results from 14 Household Surveys

Deon Filmer

Analysis of 14 household surveys from 13 developing countries suggests that 1–2 percent of the population have disabilities. Adults with disabilities typically live in poorer than average households: disability is associated with about a 10 percentage point increase in the probability of falling in the two poorest quintiles. Much of the association appears to reflect lower educational attainment among adults with disabilities. People of ages 6–17 with disabilities do not live in systematically wealthier or poorer households than other people of their age, although in all countries studied they are significantly less likely to start school or to be enrolled at the time of the survey. The order of magnitude of the school participation deficit associated with disability—which is as high as 50 percentage points in 3 of the 13 countries—is often larger than deficits related to other characteristics, such as gender, rural residence, or economic status differentials. The results suggest a worrisome vicious cycle of low schooling attainment and subsequent poverty among people with disabilities in developing countries. JEL codes: O15, J14, I32, I20, I10

With more than 100 million primary school–age children not attending school worldwide (UNESCO 2005), the target of universal education, endorsed by more than 180 countries as a part of the Millennium Development Goals, remains elusive. Children with disabilities face particular hurdles in attending and completing school in developing countries. While there has been much discussion about policy interventions to increase access to schooling for children with disabilities (see, for example, Peters 2003; World Bank 2003), little systematic empirical analysis has been conducted on which to base this policy.

The lack of analysis largely reflects the lack of appropriate and comparable data. Almost a decade ago, Elwan (1999) described the lack of empirical work

Deon Filmer is a senior economist in the Development Research Group at the World Bank; his email address is dfilmer@worldbank.org. Tomoki Fuji, Johannes Hoogeveen, Elizabeth King, Daniel Mont, Susan Peters, and Dan Rees provided helpful comments on an earlier draft. Dilip Parajuli provided excellent research assistance. A supplemental appendix to this article is available at http://wber.oxfordjournals.org/.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 141–163 doi:10.1093/wber/lhm021
Advance Access Publication January 15, 2008

on the association between disability and poverty in the developing world; such work is still missing.[1] This study aims to start filling some knowledge gaps using existing data on the prevalence of disability and its association with poverty and schooling in 12 developing, and 1 transition, countries.

Defining disability is complicated—and controversial. The purely medical definitions used in the past are giving way to definitions that incorporate continuous measures of the activities that people can undertake, the extent of their participation in society and social and civic life, and the role of adaptive technologies (Mont 2007). The World Health Organization's International Classification of Functioning, Disability and Health (ICF) describes disability as an umbrella term for impairments, activity limitations, and participation restrictions as part of a broader classification scheme covering three main domains: body functioning and structure, activities and participation, and environmental factors.[2] The interaction of aspects of all three of these domains determines individual welfare and social policy choices facing governments.[3]

The main goal of this article is descriptive. Many of the basic facts about disability, poverty, and schooling in developing countries are unknown or have not been systematically addressed. To contribute to the foundations of policy development, this article analyzes the data to investigate the interactions between impairment and schooling and their relation with poverty. It finds that disability among youth is not typically associated with household poverty but that it is systematically and significantly related to lower school participation, which in turn increases poverty in adulthood.

The article is organized as follows. Section I compares definitions and the prevalence of disability across the household surveys covered. Section II investigates the association with poverty by examining the extent to which young people with disabilities live in households with lower economic status and the extent to which disability and schooling are related to poverty in adulthood. Section III investigates the association between disability and school participation among school-age youth. Section IV draws conclusions and makes the case for better data.

I. Data

The data come from 14 nationally representative household surveys in 13 countries. Five surveys—from Bolivia (1997), Cambodia (2000), Chad (2004), Colombia (1995), and India (1992)—are Demographic and Health Surveys

1. An early exception is Afzal (1992), who analyzes disability and its correlates in Pakistan. Yeo and Moore (2003) review some of the literature on poverty and disability, but the work they refer to is typically not based on large-scale surveys.

2. A guide to the ICF is available at http://www3.who.int/icf/.

3. Haveman and Wolfe (2000, p. 998) emphasize that an economic definition of disability refers to characteristics that "constrain normal daily activities or cause substantial reduction in productivity on the job."

(DHS). Two surveys—from Jamaica (1998) and Romania (1995)—are associated with the Living Standards Measurement Study (LSMS) surveys. Two other surveys—from Burundi (2000) and Mongolia (2000)—are End of Millennium Multiple Indicator Cluster Surveys (MICS2) carried out under the guidance of the United Nations Children's Fund (UNICEF). Five surveys—from Cambodia (1999), Indonesia (2000), Mozambique (1996), South Africa (1995), and Zambia (2003)—are national socioeconomic surveys (SES).[4] These types of surveys are typically used to calculate poverty statistics or to derive basic health indicators, such as child mortality, or the use of health services; they underlie much empirical poverty and social analysis in developing economies. Most of the surveys have a sample size of about 5,000–25,000 households, with India (88,512 households) and Indonesia (65,762 households) as outliers (table 1).

All DHS, LSMS, and MICS2 surveys were reviewed for questions on disability, with all surveys with a clear question on disability for the relevant age range included.[5] The SES from Cambodia, Indonesia, Mozambique, South Africa, and Zambia are some of the most recent in the world with information on disability. There are relatively few data of this kind in developing countries: the datasets, and therefore the countries, for this analysis were selected on the basis of data availability and are not necessarily representative of developing countries in general.

This is clearly a heterogeneous group of countries. The population living on less than $1 a day ranges from 2 percent in Jamaica and Romania to 55 percent in Burundi; under-five mortality—an indicator of basic health status—ranges from 20 per 1,000 live births in Jamaica to 212 per in Mozambique (see table 1). The sample includes five countries in Sub-Saharan Africa, four in Asia, three in Latin America and the Caribbean, and one in Eastern Europe. While heterogeneity has the advantage that the results will reflect a range of underlying conditions, it has the disadvantage that little draws these countries together other than the availability of data for this analysis. The definitions of disability in these datasets are most closely consistent with a focus on impairment—and as such fall mostly under ICF's "body functioning and structure" domain. This is arguably an advantage, because impairment such as blindness or the lack of a limb is typically easy to verify.

Selective misreporting of morbidity has long been recognized as a potential problem in studies of the relation between health and other socioeconomic

4. DHS data are available at http://www.measuredhs.com; LSMS data are available at http://www.worldbank.org/lsms; MICS2 data are available at http://www.childinfo.org; national SES are available from the countries' national statistics offices. Despite the general consistency of DHS surveys across countries, disability is not a part of the "core" DHS module; information on disability is therefore not typically collected as a part of DHS. Identifying questions relating to disability required reviewing the country-specific components of the datasets.

5. Surveys with fewer than 50 disabled people between the ages of 6 and 17 were dropped from the analysis, because they represented too few observations on which to draw inferences. DHS data from Mozambique and MICS2 data from Myanmar and Sierra Leone were excluded on this basis.

TABLE 1. Basic Statistics on Countries and Surveys

Country/year of survey	Type of survey	GDP per capita (2000 PPP dollars)	Population living on less than a $1 a day (percent)	Under-five mortality rate (per 1,000 live births)	Number of households surveyed	Type of disability covered	Size of population ages 6–17	Number of people ages 6–17 with disabilities
Bolivia 1997	DHS	2,349	20	105	12,028	All	16,605	75
Burundi 2000	MICS2	584	55	190	3,979	Physical	5,865	73
Cambodia 1999	SES	1,741	34	135	6,001	All	10,881	96
Cambodia 2000	DHS	1,859	34	135	12,236	Physical	23,765	214
Chad 2004	DHS	1,241	—	208	5,366	All	9,952	57
Colombia 1995	DHS	6,207	3	31	10,107	All	11,951	130
India 1992	NFHS (DHS)	1,692	42	123	88,512	Visual, physical	140,297	1,337
Indonesia 2003	SES	3,204	8	48	65,762	All	64,136	326
Jamaica 1998	LSMS	3,611	2	20	7,375	Physical, mental	6,964	58
Mongolia 2000	MICS2	1,610	27	65	6,000	Visual, hearing	7,645	245
Mozambique 1996	SES	700	38	212	8,250	All	14,520	156
Romania 1995	LSMS	6,210	2	26	24,560	All	13,777	82
South Africa 2005	SES	11,044	11	68	28,192	All	30,151	454
Zambia 2003	SES	823	76	182	9,713	All	19,075	223

— not available.

Note: PPP is purchasing power parity; DHS is Demographic and Health Survey; MICS2 is End of Millennium Multiple Indicator Cluster Survey; SES is Socioeconomic Survey; NFHS is National Family Health Survey; LSMS is Living Standards Measurement Study survey. Data for all countries except Burundi cover 6- to 17-year-olds; age range in Burundi is 6- to 14-year-olds. Poverty rates are for the following years: Bolivia 1997; Burundi 1998; Cambodia 1997; India 1993; Indonesia 2002; Jamaica 1999; Mongolia 1998; Mozambique 1996; Romania 1998; South Africa 2000; and Zambia 2003. Under-five mortality data on Burundi, Cambodia, Colombia, Indonesia, Jamaica, and South Africa are for 2000. Data on India are for 1990. Data on Bolivia, Mozambique, and Romania are for 1995. Data on Chad, South Africa, and Zambia are for 2005. "All" types of disabilities include visual, hearing, speech, physical, and mental disabilities. See appendix table A-1 for the precise wording and types of disabilities covered by each survey.

Source: World Bank (various years); author's analysis based on data described in the text.

characteristics (Gertler, Rose, and Glewwe 2000). To overcome this problem Gertler and Gruber (2002) use responses to questions regarding activities of daily living when analyzing the impact of major illness on household consumption in Indonesia. Yount and Agree (2005) use activities of daily living in analyzing gender differences in disability among the elderly in Egypt and Tunisia.

Despite the relative ease of verifying the types of disabilities in the study datasets, it is nevertheless possible that there is selective reporting. Some respondents and interviewers, for example, might interpret blindness as partial sight, whereas others might interpret it to mean complete inability to see. It is also possible that mental disability is selectively recognized and reported by some respondents. Selective reporting is usually assumed to result in higher reporting of disabilities by wealthier socioeconomic groups. Under this assumption the estimates reported here would underestimate the relation between disability and poverty.[6]

Despite the fact that all 14 surveys have an impairment definition of disability, substantive differences remain across datasets. Nine surveys use an "extensive" definition that includes visual, hearing, speech, physical, and mental disability. But even within this group, the definition of each type of impairment varies. In the Cambodia SES, for example, the physical disability category contains a detailed list of potential cases ("amputation of one limb; amputation of more than one limb; unable to use one limb; unable to use more than one limb; paralyzed lower limbs only; paralyzed all four limbs"). In contrast, in Jamaica a single category covers "physical or mental disability." More generally, in some countries the definition is stricter than in others. In Mongolia sight and hearing are described as "with difficulty"; in other surveys they are characterized as "blind" and "deaf" (the wording of the questions on disability in these surveys is provided in appendix table A-1).

The second main data constraint is the fact that the surveys do not typically identify large numbers of people with disabilities. Any subsequent analysis therefore suffers from imprecision. The smallest number of cases of disability among 6- to 17-year-olds are in Chad (57) Jamaica (58), Burundi (73), and Bolivia (75). Standard errors are often large for the results reported below, although the main finding—the deficit in school participation among people with disabilities—is consistently statistically significant.

6. A large body of literature covers the selective reporting of disability in the context of social programs targeting disability. Higher benefits are typically hypothesized to result in higher rates of reported disability. For a recent empirical demonstration, see Duggan, Rosenheck, and Singleton (2006). Some program aspects, such as hurdles in accessing benefits, may reduce self-declared disability (see the discussion in Parsons 1991). At least one study (Benitez-Silva and others 2004) finds no systematic bias in self-reported disability compared with bureaucratic assessment among adult applicants for Social Security benefits in the United States. Reported disability might also be an unintended consequence of a different set of programs. Figlio and Getzler (2002) argue that increases in the use of learning achievement tests for school accountability in the United States has led to an increase in reported disability among students, because schools can exclude these students from average scores.

A last data constraint concerns the measurement of household poverty. All LSMS and SES surveys include household per capita consumption expenditures, the variable typically used in poverty analysis. For these datasets economic status quintiles based on per capita consumption expenditures can therefore be used. DHS and MICS2 data do not include consumption expenditures. For these datasets, an index of household consumer assets and housing characteristics (an economic status index) are used to classify households into quintiles (following Filmer and Pritchett 2001). In the Cambodia SES, which encountered problems collecting consumption data, and the South Africa SES, which does not include a full consumption module, an index based on assets and housing quality variables is also used.[7]

II. Prevalence of Disability and Association with Household Economic Status

The first issue these data are used to explore is the prevalence of disability and its association with household economic status among youth 6- to 17-years-old. The next issue is the relation among disability, poverty, and education attainment among adults.

Prevalence of Disability among 6- to 17-Year-Olds

Estimates of the prevalence of disability range from 0.49 percent (Chad) to 3.2 percent (Mongolia) (table 2). These figures are consistent with those that appear in the United Nations statistical database on disability (DISTAT).[8] In that database, which compiles results from more than 65 surveys and censuses in developing economies between 1970 and 1992, the mean prevalence rate is 1.7 percent for the entire population and 0.7 for children under 14. Using a definition of disability consistent with the one adopted in the datasets analyzed here, LeRoy, Evans, and Deluca (2000) find a prevalence of disability among 5- to 15-year-olds of 2.07–2.62 percent in Ireland, Italy, Switzerland, and the United States.[9]

Perhaps surprisingly, of the 14 surveys analyzed here, those that list more types of impairments do not systematically identify a higher percentage of the population as disabled. In countries that include visual, hearing, speech, physical, and mental disabilities, for example, prevalence ranges from 0.49 in Bolivia and Chad to 1.38 in South Africa—close to the entire range across all

7. Consistent with typical poverty analysis, quintiles are derived on the basis of the distribution of people across the economic status measure.

8. The database is available at http://unstats.un.org/unsd/demographic/sconcerns/disability/disab2.asp. A summary of the country-level DISTAT information is available in the supplemental appendix to this article, accessible at http://wber.oxfordjournals.org/.

9. The U.S. rate of 2.1 rises to 4.4 percent if "speech and language disability" (a separate category from "mute and deaf/mute") is included. When people with difficulty in learning, remembering, or concentrating are added, the rate increases to about 6 percent (Freedman, Martin, and Schoeni 2004).

TABLE 2. Prevalence of Disability among 6- to 17-Year-Olds, by Household Economic Status Quintile

Country	All	Quintile 1 (poorest)	Quintile 2	Quintile 3	Quintile 4	Quintile 5 (richest)	*p*-values on economic status variables[a]	
							Dummy variables	Continuous
Bolivia	0.49 (0.07)	0.45 (0.12)	0.43 (0.12)	0.54 (0.15)	0.64 (0.19)	0.39 (0.14)	0.82	0.78
Burundi[b]	1.24 (0.21)	1.28 (0.50)	1.19 (0.69)	1.19 (0.42)	1.28 (0.36)	1.28 (0.30)	1.00	0.62
Cambodia (Socioeconomic Survey)	0.87 (0.12)	0.91 (0.24)	0.84 (0.24)	0.87 (0.34)	0.81 (0.28)	0.94 (0.18)	1.00	0.95
Cambodia	0.86 (0.07)	1.00 (0.17)	1.01 (0.16)	0.77 (0.15)	0.65 (0.14)	0.90 (0.16)	0.40	0.39
Chad	0.49 (0.09)	0.46 (0.19)	0.32 (0.14)	0.64 (0.22)	0.49 (0.20)	0.55 (0.12)	0.72	0.70
Colombia	1.08 (0.10)	1.24 (0.24)	1.34 (0.25)	1.03 (0.22)	0.72* (0.17)	1.05 (0.26)	0.24	0.15
India	1.03 (0.04)	1.20 (0.09)	1.13 (0.09)	0.92** (0.08)	1.03 (0.07)	0.84*** (0.08)	0.01	0.01
Indonesia	0.50 (0.03)	0.70 (0.08)	0.55 (0.08)	0.41** (0.07)	0.50* (0.08)	0.38*** (0.06)	0.02	0.08
Jamaica	0.82 (0.11)	1.01 (0.28)	1.06 (0.26)	0.48 (0.18)	0.68 (0.24)	0.88 (0.30)	0.29	0.63
Mongolia	3.20 (0.27)	3.40 (0.57)	3.01 (0.50)	2.88 (0.56)	2.81 (0.50)	3.92 (0.63)	0.62	0.14
Mozambique	1.19 (0.13)	0.87 (0.17)	0.81 (0.20)	1.58* (0.36)	1.39 (0.29)	1.29 (0.28)	0.14	0.60
Romania	0.60 (0.07)	0.91 (0.19)	0.47* (0.13)	0.54 (0.16)	0.47* (0.13)	0.58 (0.14)	0.38	0.13
South Africa	1.38 (0.09)	1.50 (0.20)	1.46 (0.19)	1.67 (0.21)	1.22 (0.17)	1.06 (0.24)	0.27	0.26
Zambia	1.32 (0.11)	1.46 (0.26)	1.32 (0.22)	1.40 (0.28)	1.24 (0.23)	1.16 (0.22)	0.88	0.36

Difference with poorest quintile: ***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.

Note: The Numbers in parentheses are robust standard errors.

[a]The *p*-values report of the test of joint siginificance of the set of dummy variables for quinitiles 2–5 and of the continuous measure of economic status and its square in a probit regression of disablity on economic status.

[b]Age range Burundi is 6–14.

Source: Author's analysis based on data described in the text.

the surveys. The highest prevalence rate in this collection of datasets is observed in Mongolia (3.20 percent), which includes only visual and hearing impairments, while the rates are lower in Burundi (1.24 percent) and the Cambodia DHS (0.86 percent), which cover only physical disabilities.

Of course, this variability combines both actual differences in prevalence and differences in survey techniques. In Cambodia, which conducted two surveys separated by only one year, the survey with the more extensive definition of disability does not yield the higher prevalence: the SES in 1999, with a broad definition of disability, identifies 1.51 percent of the population as disabled, whereas the DHS in 2000, with a narrow definition (restricted to physical disabilities), identifies 1.57 percent of the population as disabled. Clearly, there is substantial variation across surveys in how people with disabilities are identified; cross-country comparisons in prevalence can be made only with caution.[10]

Despite the lack of cross-country comparability in the definitions and measurement of disability, these surveys are still useful in describing the association of disability with other characteristics. That is, conditional on a particular definition, the analysis is valid for a given survey because the definition is common to all individuals in that survey. Moreover, it is less likely that cross-country comparisons of the association between disability and other characteristics suffer from these problems. But even this comparison needs to be treated with some caution: if, for example, some types of disabilities are more closely associated with a correlate than others, then surveys that include that type of disability will show a higher association with the correlate than those that do not. If loss of a limb is more closely associated with poverty than are other types of impairments, for example, then (everything else being equal) a survey that includes loss of a limb in its definition of disability will yield a higher correlation between disability and poverty.

Do Youth with Disabilities Live in Poorer Households?

The prevalence of disability among 6- to 17-year-old tends to be slightly lower in richer quintiles, but the association is not always monotonic. Moreover, India and Indonesia are the only countries in which the prevalence of disability in the richest quintile is statistically significantly different from that in the poorest quintile. In India, prevalence in the richest quintile is about a third of that in the poorest quintile; in Indonesia it is about half (see table 2).

Two additional tests of the association between disability and poverty were carried out. The first entailed estimating a probit regression of disability on dummy variables for living in a household of each economic status quintile and then determining the joint statistical significance of these dummy variables (that is, a joint version of the quintile-by-quintile tests). This approach allows

10. Developing good data on disability is difficult: United Nations (2001) contains a guide to doing so. See Mont (2007) for a recent review.

for a great degree of nonlinearity in the association, because the coefficient can be different for each quintile and the approach does not impose monotonicity. However, it is possible that the small number of young people with disabilities means that there will be very few cases within each quintile and that therefore even a joint test of the coefficients on the quintile dummy variables may not have enough power to identify a significant association.

To address this potential problem, the second approach entailed a probit regression of disability on the continuous variable measuring economic status (per capita household expenditures or an index of assets and housing characteristics) and its square. This approach allows less flexibility, but it does not rely on quintile-specific estimates of prevalence, which may be imprecise.

In both tests, India and Indonesia are the only countries in which either the joint test on the dummy variables or on economic status and its square yield a statistically significant association. These are the countries with the largest sample sizes, which gives rise to the concern that it is simply the power of the test that is low in the other countries. However, as discussed below, the same datasets yield large and statistically significant gaps in schooling in all countries, suggesting that it is not simply an issue of power.[11] In sum, these results do not suggest that, as a general rule, youth with disabilities are more likely to live in poorer households, although this is the case in two of the 14 datasets.

Disability, Poverty, and Schooling in Adulthood

Disability is a strong correlate of poverty in adulthood. Haveman and others (1999) show that income in the United States in 1991 of households headed by a person with a disability was roughly half the mean for the population as whole (even after accounting for transfers from social programs) and the rate of poverty about twice as high as the overall population's. Hoogeveen (2005) estimates that 42 percent of households headed by a person with a disability were poor in Uganda in 1991 but that just 25 percent of other households were.[12]

The analysis of the relation between disability and economic status should be interpreted as an association and not necessarily a cause or consequence. Disability is both a determinant of poverty, because it lowers earning power and consumption expenditures (Haveman and Wolfe 2000; Gertler and Gruber 2002), and a consequence of poverty, because the cumulative deprivations of

11. The results are consistent if (as suggested by a referee) one estimates the reverse regression of the probability of being poor as a function of whether a school-age child has a disability controlling for other covariates (that is, a specification analogous to the schooling model estimated below). In this approach, disability is statistically insignificantly related to poverty in all the countries when poverty is defined as being in the poorest two quintiles. The association is statistically significant in India and Indonesia if poverty is defined as being in the poorest quintile (results are available in the supplemental tables of this article, available at http://wber.oxfordjournals.org/).

12. Hoogeveen (2005, p. 606) defines disability among household heads differently than the datasets used here. In that survey, a head of household is considered disabled if the disability "prevents him or her from being actively engaged in labor activities during the past week."

poverty such as inadequate infant or child development, or exposure to dangerous working conditions, can manifest themselves in disability. Moreover, the presence of a person with a disability entails direct costs, which lower standards of living (Jones and O'Donnell 1995; Haveman and Wolfe 2000; Zaidi and Burchardt 2005).

Twelve of the datasets include information on the disability status of adults. A probit regression model was estimated in which the dependent variable is a dummy variable equal to 1 if a person of 20–50 lives in a household in the two poorest quintiles.[13] The first model estimates the association with disability after controlling for a set of basic characteristics: age, age-squared, a dummy variable for being male, urban residence, and dummy variables for region of residence. Disability is statistically significantly related to an increase in the probability of being poor in eight of the datasets (table 3). Among these, having a disability is associated with a 5.0–14.5 percentage point increase in the probability of being in the two poorest quintiles. In the remaining four datasets the association is positive but not statistically significant.

The second model adds years of schooling completed to the set of correlates of poverty, thereby controlling for the effect of schooling on poverty. The coefficient on disability becomes statistically insignificant in several countries—and turns from being significantly positive to significantly negative in several others.[14] In all cases, the coefficient on years of schooling is statistically significantly negative: each additional year of schooling is associated with about a 2–5 percentage point reduction in the probability of being in the two poorest quintiles.[15]

These results suggest that disability and poverty are related in adulthood and that much of this association is mediated by education: after accounting for the lower educational attainment of adults with a disability there is no longer a systematic positive relation between disability and poverty. This finding suggests that ensuring that youth with disabilities do not have lower educational attainment could be a powerful way to reduce the likelihood that they live in poverty as adults. The next section shows just how great a challenge this is.

III. Disability and Schooling

Consider now the relation between disability and schooling among youth (table 4). Six- to seventeen-year-olds with disabilities are almost always substantially less likely to be in school than their peers without disabilities.

13. Qualitatively similar results are obtained if just the poorest quintile is used.

14. This effect is caused by years of schooling alone. The results are qualitatively similar if only years of schooling enter the model and all other correlates are dropped.

15. The results presented in the supplemental appendix, available at http://wber.oxfordjournals.org/, do not support the notion that there is an interactive effect between disability and schooling: when included, an interaction term is always insignificantly different from zero and small in magnitude in all countries.

TABLE 3. Association between Disability, Schooling, and the Probability of Being in the Bottom Two Economic Status Quintiles among Adults 20–50

Country	Model 1, excluding schooling but including other variables	Model 2, including schooling and other control variables		Number of observations (4)	Number of people with disabilities (5)
	Coefficient on: disability (1)	Coefficient on: disability (2)	Coefficient on: years of schooling (3)		
Bolivia	0.145*** (0.052)	0.013 (0.057)	−0.037*** (0.002)	18,632	179
Cambodia (Socioeconomic Survey)	0.015 (0.040)	0.013 (0.046)	−0.027*** (0.003)	9,195	251
Cambodia	0.129*** (0.028)	0.092*** (0.029)	−0.047*** (0.002)	23,191	504
Chad	0.056 (0.077)	0.021 (0.074)	−0.037*** (0.005)	7,128	138
Colombia	0.131*** (0.033)	−0.052* (0.031)	−0.050*** (0.002)	18,807	309
India	0.068*** (0.012)	0.052*** (0.012)	−0.049*** (0.001)	199,140	309
Indonesia	0.083*** (0.024)	−0.030 (0.023)	−0.042*** (0.001)	121,964	774
Jamaica	0.133*** (0.048)	0.084* (0.048)	−0.024*** (0.002)	10,197	188
Mozambique	0.015 (0.032)	0.006 (0.032)	−0.019*** (0.003)	13,909	366
Romania	0.050* (0.028)	−0.068** (0.028)	−0.046*** (0.002)	30,584	356
South Africa	0.097*** (0.018)	−0.020 (0.020)	−0.061*** (0.002)	44,539	1,766
Zambia	0.015 (0.034)	−0.039 (0.035)	−0.032*** (0.002)	18,596	377

***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.

Note: The Numbers in parentheses are robust standard errors. Explanatory varible differ across models. Model 1 includes disability, age, age-squared, and dummy varibles for gender, urban residence, and region. Model 2 includes years of schooling and all the variable from model 1.

Source: Author's analysis based on data described in the text.

The shortfall among children aged 6–11 ranges from 10 percentage points in India to almost 60 percentage points in Indonesia. The gap is also large among older children, ranging from 15 percentage points in Cambodia to 58 percentage points in Indonesia (exceptions are India, where the gap is just 4 percentage points, and Burundi, where there is no gap). The gaps are typically larger among the older group: the median shortfall is 21 percentage points among 6- to 11-year-olds and 25 percentage point among 12- to 17-year olds.

Schooling Deficits Controlling for Individual, Household, and Community Characteristics

To the extent that disability is correlated with other factors that affect schooling, such as poverty, age, and urban or rural residence, the raw difference in school participation between children with and without disabilities may give a misleading picture. For each survey, an adjustment was carried out by estimating a multivariate probit model with school participation as the dependent variable and an indicator of disability as the explanatory variable (table 5). The estimates also include, as explanatory variables, the potentially confounding variables—age and age-squared, a dummy variable for a child's gender, a dummy variable for urban residence, dummy variables for each economic status quintile, and dummy variables for region of residence.[16]

The adjusted deficit in school participation is more than 50 percentage points in Bolivia, Indonesia, and Romania; 24–45 percentage points in Cambodia, Colombia, Jamaica, Mongolia, South Africa, and Zambia; 14–18 percentage points in Burundi, Chad, and Mozambique; and 8 percentage points in India. In all countries, the difference is large and statistically significantly different from zero. In most countries, the unadjusted deficits are of comparable orders of magnitude: the estimated deficits are usually smaller after adjusting for confounding variables, but the effect of the adjustment is not typically large.

The results for the probability that a person has ever attended school are similar to those for current school participation. As the deficit is of a similar order of magnitude, the results imply that a substantial part of the deficit in schooling attainment among people with disabilities comes from the fact that they never attended school at all. Analysis of the Kaplan-Meier grade survival curves (table S.2 in the supplemental appendix, available at http://wber.oxfordjournals.org/) suggests that most of the difference in attainment can be attributed to the decision (or the ability) to enter school. Nevertheless, in seven countries, the deficit at grade 1 widens as children progress through the school system: in Bolivia, Colombia, Indonesia, Jamaica, Romania, South Africa, and Zambia the deficit associated with disability increases by about 7–10

16. Similar results (presented in the supplemental appendix, available at http://wber.oxfordjournals.org/) are found if nearest neighbor matching (using the same set of explanatory variables for matching) is used.

TABLE 4. Percentage of 6- to 17-Year-Olds Reported to Be in School, by Country

	Age 6–11			Age 12–17		
Country	Without disability	With disability	Difference	Without disability	With disability	Difference
Bolivia	0.95 (0.00)	0.38 (0.08)	0.57*** (0.08)	0.83 (0.01)	0.39 (0.11)	0.44*** (0.11)
Burundi[a]	0.38 (0.01)	0.19 (0.06)	0.19*** (0.06)	0.48 (0.02)	0.48 (0.10)	0.00 (0.10)
Cambodia (Socioeconomic Survey)	0.58 (0.01)	0.18 (0.06)	0.40*** (0.06)	0.68 (0.01)	0.31 (0.08)	0.37*** (0.08)
Cambodia	0.67 (0.01)	0.38 (0.06)	0.29*** (0.06)	0.62 (0.01)	0.47 (0.06)	0.15*** (0.06)
Chad	0.36 (0.02)	0.24 (0.12)	0.12 (0.12)	0.43 (0.03)	0.25 (0.09)	0.18* (0.10)
Colombia	0.92 (0.01)	0.56 (0.08)	0.36*** (0.08)	0.74 (0.01)	0.29 (0.06)	0.45*** (0.06)
India	0.70 (0.01)	0.60 (0.02)	0.10*** (0.02)	0.35 (0.00)	0.32 (0.02)	0.04* (0.02)
Indonesia	0.89 (0.00)	0.29 (0.04)	0.59*** (0.04)	0.76 (0.00)	0.18 (0.04)	0.58*** (0.04)
Jamaica	0.99 (0.00)	0.71 (0.09)	0.29*** (0.09)	0.86 (0.01)	0.50 (0.11)	0.36*** (0.11)
Mongolia	0.58 (0.01)	0.41 (0.05)	0.17*** (0.04)	0.73 (0.02)	0.47 (0.05)	0.26*** (0.05)
Mozambique	0.49 (0.01)	0.34 (0.08)	0.15** (0.08)	0.48 (0.01)	0.29 (0.06)	0.19*** (0.06)
Romania	0.79 (0.01)	0.58 (0.11)	0.22** (0.11)	0.84 (0.01)	0.36 (0.07)	0.48*** (0.07)
South Africa	0.96 (0.00)	0.76 (0.04)	0.21*** (0.04)	0.95 (0.00)	0.70 (0.04)	0.25*** (0.04)
Zambia	0.62 (0.01)	0.42 (0.06)	0.20*** (0.06)	0.75 (0.01)	0.58 (0.06)	0.17*** (0.06)

***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.

Note: Numbers in parentheses are robust standard errors.

[a]Age range in Burundi is 6–14.

Source: Author's analysis based on data described in the text.

TABLE 5. Schooling Deficits among 6- to 17-Year-Olds with Disabilities

	Current school participation			Ever-attended school		
Country	Average among 6- to 17-year olds without disabilities (1)	Unadjusted deficit (2)	Deficit adjusted for other factors (3)	Average among 6- to 17-year olds without disabilities (4)	Unadjusted deficit (5)	Deficit adjusted for other factors (6)
Bolivia	0.90	−0.51*** (0.07)	−0.61*** (0.08)	0.98	−0.47*** (0.07)	−0.46*** (0.08)
Burundi[a]	0.41	−0.12** (0.06)	−0.16*** (0.05)	0.45	−0.14** (0.06)	−0.19*** (0.05)
Cambodia (Socioeconomic Survey)	0.63	−0.39*** (0.05)	−0.45*** (0.05)	0.72	−0.46*** (0.05)	−0.56*** (0.06)
Cambodia	0.65	−0.22*** (0.04)	−0.26*** (0.05)	0.74	−0.20*** (0.04)	−0.31*** (0.05)
Chad	0.39	−0.14* (0.08)	−0.14* (0.08)	0.43	−0.12* (0.07)	−0.13 (0.09)
Colombia	0.83	−0.42*** (0.05)	−0.43*** (0.06)	0.95	−0.47*** (0.05)	−0.48*** (0.06)
India	0.54	−0.08*** (0.02)	−0.08*** (0.02)	0.74	−0.05*** (0.02)	−0.07*** (0.02)
Indonesia	0.82	−0.59*** (0.03)	−0.67*** (0.03)	0.94	−0.46*** (0.03)	−0.53*** (0.05)
Jamaica	0.93	−0.33*** (0.08)	−0.28*** (0.09)	1.00	−0.25*** (0.07)	−0.19*** (0.06)
Mongolia	0.65	−0.20*** (0.04)	−0.28*** (0.04)	0.80	−0.17*** (0.04)	−0.37*** (0.05)
Mozambique	0.49	−0.18*** (0.05)	−0.18*** (0.05)	0.61	−0.12** (0.05)	−0.14*** (0.05)
Romania	0.82	−0.39*** (0.06)	−0.53*** (0.07)	0.90	−0.30*** (0.06)	−0.50*** (0.07)
South Africa	0.96	−0.23*** (0.03)	−0.25*** (0.03)	0.85	−0.20*** (0.03)	−0.30*** (0.05)
Zambia	0.68	−0.18*** (0.04)	−0.24*** (0.05)	0.77	−0.16*** (0.04)	−0.23*** (0.05)

***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.

Note: Adjusted differentials correspond to the marginal effect of disability in a probit regression of school participation that includes age, age-squared, and dummy variables for gender, urban residence, economic quintile, and region. Numbers in parentheses are robust standard errors.

[a]Age range in Burundi is 6–14.

Source: Author's analysis based on data described in the text.

percentage points between grades 1 and 8.[17] These countries have relatively high grade 1 completion rates, suggesting that even in countries that are able to get most children into school, special effort may be needed to get children with disabilities into school and to keep them there.

There is substantial heterogeneity across countries in the schooling deficit associated with disability. Part of this variation might be due to differences in the definition of disability. In a survey with a more stringent definition of disability, one might observe a larger deficit, because such a survey would identify individuals who would have to overcome greater obstacles in order to access education. The fact that the two surveys from Cambodia yield schooling deficits that are about 20 percentage points apart suggests that this is a plausible explanation.

Another part of this variation likely relates to overall enrollment. It would not be surprising to observe larger deficits in countries in which enrollment among children without disabilities is high: in these countries there would be more scope to observe a bigger deficit. The schooling deficit does tend to be smaller in countries with the lowest overall enrollment (Burundi, Chad, India, and Mozambique) and larger in countries with higher overall enrollment (Bolivia, Indonesia, and Romania): the correlation between the school participation deficit and the level of participation among youth without a disability is about −0.4 across the 14 datasets (the correlation is similar for the probability of ever having attended school). But the relation is not perfect: in Jamaica and South Africa, for example, where overall school participation is high, the deficit associated with disability is about average for the surveys analyzed here.

Finally, part of the variation in the schooling deficit associated with disability is likely related to differences in the social and policy environment. Countries in which there is greater stigma toward people with disabilities or less effort has been made to ensure their access to schooling will undoubtedly have larger deficits associated with schooling. However, this is but one of many potential explanations of why results might differ across countries.

"Endogeneity" of Disability and Schooling

Disability among 6- to 17-year olds could be partly the result of poverty, which may have a direct effect on the probability of school attendance. Adjusting for economic status in estimating the association between schooling and disability mitigates the potential for such bias. More generally, however, it is possible that other—unobserved—factors affect both the probability of being disabled and the probability of attending school. Indeed, households that disfavor investing in both children's health and education in favor of other types of expenditures are more likely to have infants and children with poor health—who might develop a disability as a result—and low schooling. In this case disability and schooling would be related, but the association would merely reflect

17. The deficit in grade 1 ranges from 15 percentage points in Zambia to 48 percentage points in Bolivia.

parental neglect (see Strauss and Thomas 1995; Haddad, Hoddinott, and Alderman 1997).

One way to address this potential problem is to use a household fixed-effects approach. Such an approach controls for all—observed and unobserved—household-level characteristics common to all children in a household. In such a model the source of identification of the difference in school participation is between children with and children without a disability in the same household. Any generalized household-specific above- or below-average investment in children will have been netted out. Implementing such an approach involves revising the set of control variables used in the "adjusted" models reported in table 5 and replacing all household-, community-, and regional variables with a set of dummy variables each equal to one for each household.[18]

A household fixed-effects specification can be estimated only on the subsample of households that include at least one youth with disabilities and one without (table 6). The results of re-estimating the basic multivariate results without household fixed-effects on the subsample are consistent with those obtained using the full sample, despite the potentially selected nature of this subsample. The household fixed-effects results for current school participation and for ever-attended school are likewise similar to those that exclude fixed-effects for the subsample. In one country (Burundi) the magnitude of the effect increases, in another (Mozambique) it decreases. But in most countries the estimated impact is virtually indistinguishable, suggesting that the association between disability and schooling among 6- to 17-year-olds is not simply a reflection of fixed household attributes, such as parental neglect, but rather a more direct effect of disability on schooling.

Relative Magnitude of School Participation Deficits

How large is the deficit in school participation relative to other sources of inequality? The multivariate models can be used to compare school participation gaps associated with disability, gender, urban or rural residence, and economic status (figure 1).[19]

The deficit associated with disability is clearly large compared with other sources of inequality. In almost all countries it is larger than the deficit associated with being female (which is a "surplus" in some countries). In most countries it is substantially larger than the deficit associated with rural residence, and it is usually larger even than the gap between the poorest and richest quintiles, typically one of the strongest predictors of enrollment.[20] The exceptions are Burundi, Chad, and particularly India, where wealth gaps are larger than all other gaps; Burundi and Mozambique, where rural-urban gaps are larger than

18. In this section, the probit model is replaced by a linear probability model.

19. In each case the deficit is estimated at the means of the other variables.

20. See Filmer (2005) for a comparison of wealth and gender gaps. See Ainsworth and Filmer (2006) for a comparison of the gaps associated with wealth and with orphan status.

TABLE 6. Effect of Disability on Schooling of 6- to 17-Year-Olds in Households with at Least One Child with and One Child without a Disability

	Current school participation		Ever attended school			
Country	Basic multivariate specification (1)	Household fixed-effects specification (2)	Basic multivariate specification (3)	Household fixed-effects specification (4)	Total number of observations (5)	Number ages 6–17 with disabilities (6)
Bolivia	−0.50*** (0.08)	−0.49*** (0.09)	−0.43*** (0.08)	−0.44*** (0.10)	187	61
Burundi[a]	−0.24*** (0.06)	−0.31*** (0.08)	−0.22*** (0.06)	−0.30*** (0.08)	136	61
Cambodia (Socioeconomic Survey)	−0.36*** (0.06)	−0.39*** (0.07)	−0.40*** (0.06)	−0.42*** (0.07)	265	82
Cambodia	−0.21*** (0.04)	−0.22*** (0.05)	−0.23*** (0.04)	−0.23*** (0.04)	649	189
Chad	−0.10 (0.10)	−0.07 (0.11)	−0.07 (0.10)	−0.03 (0.11)	218	52
Colombia	−0.39*** (0.05)	−0.39*** (0.07)	−0.48*** (0.06)	−0.48*** (0.07)	276	98
India	−0.05*** (0.02)	−0.06** (0.02)	−0.04** (0.02)	−0.05** (0.02)	3,574	1,138
Indonesia	−0.51*** (0.04)	−0.51*** (0.06)	−0.43*** (0.04)	−0.44*** (0.06)	545	208
Jamaica	−0.34*** (0.12)	−0.40*** (0.12)	−0.28*** (0.09)	−0.34*** (0.11)	110	42
Mongolia	−0.25*** (0.04)	−0.24*** (0.05)	−0.28*** (0.04)	−0.28*** (0.04)	557	201
Mozambique	−0.23*** (0.06)	−0.13** (0.06)	−0.16** (0.05)	−0.11* (0.06)	370	121
Romania	−0.41*** (0.09)	−0.44*** (0.10)	−0.34*** (0.07)	−0.36*** (0.10)	136	52
South Africa	−0.18*** (0.03)	−0.21*** (0.04)	−0.19*** (0.03)	−0.21*** (0.04)	1,039	361
Zambia	−0.17*** (0.05)	−0.20*** (0.05)	−0.18*** (0.04)	−0.19*** (0.05)	700	253

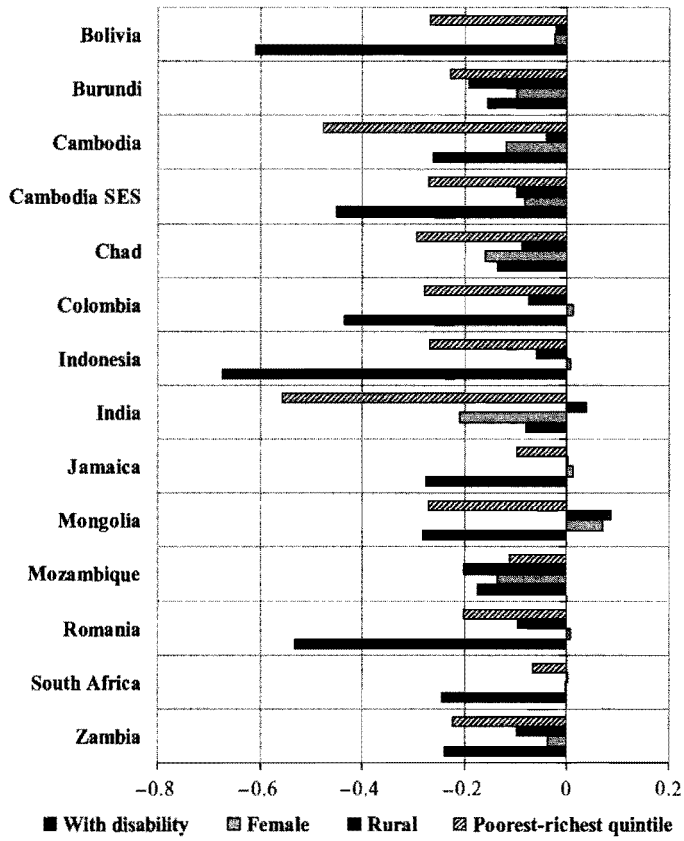
***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.

Note: Coefficients are from linear probability models. Basic specification includes age, age-squared, and dummy variables for gender, urban residence, economic quintile, and region; household fixed-effects specification includes age, age-squared, and a dummy variable for gender. Numbers in parentheses are robust standard errors.

[a]Age range in Burundi is 6–14.

Source: Author's analysis based on data described in the text.

FIGURE 1. Deficits in School Participation Associated with Various Characteristics



Note: Deficits shown are the marginal effects of dummy variables for each characteristic in multivariate probit models for 6- to 17-year-olds, except in Burundi, where the sample covers children ages 6–14.

Source: Author's analysis based on data described in the text.

those for disability; and Mongolia and Zambia, where wealth gaps are only slightly smaller than those for disability. In most countries, however, the gap in school participation between children with and without disabilities is about twice as large as that associated with rural residence or wealth.[21]

21. An interesting additional question would be whether disability interacts with other characteristics in a way that reduces or exacerbates inequalities. A straightforward way to investigate this hypothesis is to estimate the multivariate model of school participation and include interaction terms between disability and each of the other covariates. Given the relatively small number of sample observations, however, the data do not typically yield statistically significant interaction terms—even when the magnitude of the interaction is relatively large—suggesting an inability to estimate these interactions with much precision. Because it is hard to assess whether this is caused solely by statistical power, these results are not reported here. They are reported in the supplemental appendix, available at http://wber.oxfordjournals.org/.

IV. Conclusions

This analysis of data from 14 nationally representative household surveys confirms the many data problems that earlier research has identified as hampering the establishment of a broad empirical base for developing policies targeted to people with disabilities in poor countries. The variation across surveys in the definition of disability makes cross-country comparisons difficult. The small number of people identified as disabled in surveys makes it hard to precisely estimate patterns in the data beyond simple correlations.

Despite these limitations, but keeping them in mind, the data are nevertheless revealing. Consistent with similar surveys the 14 surveys analyzed here identify about 1–2 percent of the population as having a disability. One country with two surveys and varying definitions suggests that the percentage is not always sensitive to the exact definition: different definitions can yield similar prevalence rates, and similar definitions can yield different prevalence rates. In addition, other aspects of the surveys, such as the training of enumerators or the use to which interviewees expect the survey to be put, might affect the overall estimated prevalence rates.

Analysis of these datasets provides little evidence to suggest that children with disabilities are generally more or less likely to live in richer or poorer households. Adults with disabilities do typically live in poorer households, but much of this association appears to come from the fact that they have lower educational attainment. Given this finding, it is particularly worrisome that children with disabilities are almost always much less likely to participate in schooling than are other children. They are also less likely to start school, and in some countries they have lower transition rates. The school participation disability deficit is typically larger than deficits associated with characteristics such as gender, rural residence, or economic status.

This analysis suggests that in developing countries disability is associated with long-run poverty, in the sense that children with disabilities are less likely to acquire the human capital that will allow them to earn higher incomes. In all countries, the schooling gap between children with and without disabilities starts at grade 1, suggesting that efforts are needed to boost enrollments of children with disabilities at the earliest grades in order to increase education attainment for this population. The result that the disability deficit widens from grade to grade in countries that have achieved relatively high enrollment among children without disabilities suggests that special effort may be needed to keep children with disabilities in school.

The results of this analysis should be treated as tentative. Better data are needed. Establishing clear and consistent measures of disability for use in household surveys and national censuses would be a start. A recent review (Mont 2007) suggests that questions that focus on functionality, concentrate on a core set of activities, and allow for variation in the degree of functional limitation (as opposed to the mere presence or absence of a limitation) should

be preferred. To build the quantitative evidence base for empirically grounded policies, these questions must be implemented within samples that are large enough to allow detailed analysis. An important complement to the information that would emerge from the analysis of such data would be evaluations of the impact of alternative interventions that attempt to increase enrollments among children and youth with disabilities.

Funding

The Disability Team of the Human Development Network of the World Bank provided partial funding for this work.

Appendix

Table A-1. Definition of Disabilities in Covered Surveys

Country/year of survey	Name of survey	Definition used in survey	Question from survey instrument
Bolivia 1997	DHS	Mentally retarded, deaf, mute, blind, paralyzed, crippled	Does [X] have any of the following extreme physical impediments?
Burundi 2000	MICS2	Presence of a physical handicap (missing upper or lower limbs, or other body part)	Specific wording not available.
Cambodia 1999	SES	Amputation of one or more limbs, inability to use one or more limbs, blind, deaf, mute, mentally disturbed, permanent disfigurement, other	Does X have a disability? If "yes," what type of disability does X have? [respondent chooses from coded answers]; What was the cause of the disability? [respondent chooses from coded answers].
Cambodia 2000	DHS	Physical impairment	Is there a person who usually lives in your household who has any type of physical impairment? Please give the name of each individual who has a physical impairment. For each individual, then ask: "Has X been physically impaired since birth, or was X's impairment due to an accident?" [respondent chooses from coded answers].

(Continued)

TABLE A-1. Continued

Country/year of survey	Name of survey	Definition used in survey	Question from survey instrument
Chad 2004	DHS	Missing limb, deformed limb, blind, deaf, mute, missing body part, behavioral	Is there a person in your household who . . .is missing a part of the body, for example a hand, an arm, or a leg, has difficulty seeing or is almost or completely blind; has difficulty hearing or is deaf; who has difficulty speaking or is mute; who is missing an extremity such as finger tips, toes, nose or ear; who has behavioral problems.
Colombia 1995	DHS	Blind, deaf, mute, paralysis or missing limb, mental retardation, behavioral problem.	Does [X] have one of the following health problems and how did they acquire the problem? [respondent chooses from coded answers].
India 1992	NFHS (DHS)	Blind, limb impairment of	Does anyone [in household] suffer from: with separate answers for: "Blindness?" with options "yes: partial; yes: complete; "no" and for "Any physical impairment of limbs?" with options "yes: hands"; "yes: legs"; "yes: both".
Indonesia 2003	SUSENAS	Blind, deaf, mute, physical disability, mental disability.	Have a disability? If yes: "Type of disability" [respondent chooses from coded answers]; "Cause of disability" [respondent chooses from coded answers].
Jamaica 1998	LSMS	Physical or mental disability.	Is X physically or mentally disabled?
Mongolia 2000	MICS2	Difficulty seeing, difficulty hearing	Specific wording not available.
Mozambique 1996	SES	Blind, deaf, mute, mental disability paralysis, amputated arm(s), amputated leg(s), other.	Have a disability? If yes: "Type of disability" [respondent chooses from coded answers]; "Cause of disability" [respondent chooses from coded answers].

(Continued)

TABLE A-1. Continued

Country/year of survey	Name of survey	Definition used in survey	Question from survey instrument
Romania 1995	LSMS	Amputation of limb, paralysis of limbs, ankylosis of limbs or column, physical deformation, unilateral or bilateral blindness, deafness, muteness, epilepsy, mental retardation, mental disorder	Do you suffer from a handicap? If yes: "Type of handicap" [respondent chooses from coded answers].
South Africa 2005	General household survey	Sight (blind/severe visual limitation), hearing (deaf, profoundly hard of hearing), communicating (speech impairment), physical (for example needs wheel chair, crutches or prosthesis, limb or hand usage limitation), intellectual (serious difficulties in learning, mental retardation), emotional (behavioral, psychological problems), other	I am now going to ask about disabilities experienced by any persons within the household. Is X limited in his/her daily activities, at home, at work or at school, because of a long-term physical, sensory, hearing, intellectual, or psychological condition, lasting six months or more? If yes, "what difficulty or difficulties does X have?" [respondent chooses from coded answers].
Zambia 2003	LCMS	Blind, partially sighted, deaf, dumb, crippled, mentally retarded, mentally ill, former mental patient	Is X blind, partially sighted, deaf, dumb, crippled, mentally retarded, mentally ill, ex-mental [sic], or has multiple disabilities?

Note: DHS is Demographic and Health Survey; MICS2 is End of Millennium Multiple Indicator Cluster Survey; SES is Socioeconomic Survey; NFHS is National Family Health Survey; SUSENAS is National Socioeconomic Survey; LSMS is Living Standards Measurement Study survey; LCMS is Living Conditions Monitoring Survey.

Source: Author compilation.

REFERENCES

Afzal, Mohammad. 1992. "Disability Prevalence and Correlates in Pakistan: A Demographic Analysis." *Pakistan Development Review* 31(3):217–57.

Ainsworth, Martha, and Deon Filmer. 2006. "Children's Schooling: AIDS, Orphanhood, Poverty, and Gender." *World Development* 34(6):1099–128.

Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser, and John Rust. 2004. "How Large Is the Bias in Self-Reported Disability?" *Journal of Applied Econometrics* 19(6):649–70.

Duggan, Mark, Robert Rosenheck, and Perry Singleton. 2006. *Federal Policy and the Rise in Disability Enrollment: Evidence for the VA's Disability Compensation Program.* NBER Working Paper 12323. Cambridge, Mass.: National Bureau of Economic Research.

Elwan, Ann. 1999. *Poverty and Disability: A Survey of the Literature.* Social Protection Discussion Paper 9932. Washington, D.C.: World Bank.

Figlio, David, N., and Lawrence S Getzler. 2002. *Accountability, Ability and Disability: Gaming the System.* NBER Working Paper 9307. Cambridge, Mass.: National Bureau of Economic Research.

Filmer, Deon. 2005. "Gender and Wealth Disparities in Schooling: Evidence from 44 Countries." *International Journal of Education Research* 43(6):351–69.

Filmer, Deon, and Lant Pritchett. 2001. "Estimating Wealth Effects without Expenditure Data—or Tears: With an Application to Educational Enrollments in States of India." *Demography* 38(1):115–32.

Freedman, Vicki, Linda G. Martin, and Robert Schoeni. 2004. "Disability in America." *Population Bulletin* 59(3):1–32.

Gertler, Paul J., and Jonathan Gruber. 2002. "Insuring Consumption against Illness." *American Economic Review* 92(1):51–70.

Gertler, Paul J., Elaina Rose, and Paul Glewwe. 2000. "Health." In M. Grosh, and P. Glewwe eds., *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*. Washington, D.C.: World Bank.

Haddad, Lawrence, John Hoddinott, and Harold Alderman. 1997. *Intrahousehold Resource Allocation in Developing Countries: Models, Methods, and Policy*. Baltimore, Md.: Johns Hopkins University Press for the International Food Policy Research Institute.

Haveman, Robert, Karen Holden, Barbara Wolfe, Paul Smith, and Kathryn Wilson. 1999. "The Changing Economic Status of U.S. Disabled Men: Trends and Their Determinants, 1982–1991." *Empirical Economics* 24:571–598.

Haveman, Robert, and Barbara Wolfe. 2000. "The Economics of Disability and Disability Policy." In A. J. Culyer, and J. P. Newhouse, eds., *Handbook of Health Economics*. Vol. 1. Amsterdam, New York: Elsevier Science B.V.

Hoogeveen, Johannes G. 2005. "Measuring Welfare for Small but Vulnerable Groups Poverty and Disability in Uganda." *Journal of African Economies* 14(4):603–31.

Jones, Andrew, and Owen O'Donnell. 1995. "Equivalence Scales and the Costs of Disability." *Journal of Public Economics* 56(2):273–89.

LeRoy, Barbara W., Peter Evans, and Marcella Deluca. 2000. "United States and European School-Aged Disability Prevalence: An Investigative Study to Elaborate Differences." Organisation for Economic Co-operation and Development, Paris. (www.oecd.org/dataoecd/9/59/27133749.pdf)

Mont, Daniel. 2007. *Measuring Disability Prevalence*. Social Protection Discussion Paper 706. Washington D.C.: World Bank.

Parsons, Donald O. 1991. "Self-Screening in Targeted Public Transfer Programs." *Journal of Political Economy* 99(4):859–76.

Peters, Susan J. 2003. "Inclusive Education: Achieving Education for All by Including those with Disabilities and Special Needs." Disability Group, Human Development Network, World Bank, Washington, D.C. (http://siteresources.worldbank.org/DISABILITY/Resources/Education/Inclusive_Education_En.pdf)

Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." *Handbook of Development Economics*. Vol. 3. New York: North-Holland Press.

United Nations. 2001. *Guidelines and Principles for the Development of Disability Statistics*. Statistics on Special Population Groups, Series Y, No. 10. New York: Statistics Division.

UNESCO (United Nations Educational, Scientific and Cultural Organization). 2005. *Education for All Monitoring Report 2005: The Quality Imperative*. Paris: UNESCO.

World Bank. 2003. "Education for All: Including Children with Disabilities: Summarized Lessons Learned and Key Policy Findings on the World Bank's Work in Education." *Education Notes*. Human Development Network, World Bank, Washington, D.C.

———. Various years. *Word Development Indicators*. Washington, D.C.: World Bank.

Yeo, Rebecca, and Karen Moore. 2003 "Including Disabled People in Poverty Reduction Work: 'Nothing about Us, without Us.'" *World Development* 31(3):571–90.

Yount, Kathryn M., and Emily Agree. 2005. "Differences in Disability among Older Women and Men in Egypt and Tunisia." *Demography* 42(1):169–87.

Zaidi, Asghar, and Tania Burchardt. 2005. "Comparing Incomes When Needs Differ: Equivalization for the Extra Costs of Disability in the U.K." *Review of Income and Wealth* 51(1):89–114.

The Impact of Decentralized Data Entry on the Quality of Household Survey Data in Developing Countries: Evidence from a Randomized Experiment in Vietnam

Paul Glewwe and Hai-Anh Hoang Dang

Computers were provided to randomly selected districts participating in a household survey in Vietnam to assess the impact on data quality of entering data within a day or two of completing the interview rather than several weeks later in the provincial capital. Provision of computers had no significant effect on the observed distribution of household expenditures and thus no effect on measured poverty. Provision of computers reduced the mean number of errors per household by 5–23 percent, depending on the type of error. Given the already low rate of errors in the survey, however, the goal of increasing the precision of the estimated mean of a typical variable can be achieved at a much lower cost by slightly increasing the sample size. Provision of additional computers did substantially reduce the time interviewers spent adding up and checking the data in the field, with the value of the time saved close to the cost of purchasing desktop computers. JEL Classification: C81, C93, C42, I32, O15

Household survey data are used for many policy and research purposes in developed and developing countries. Yet anyone who works with household survey data soon realizes that they can have many errors and inconsistencies. Statisticians call such errors nonsampling errors. Economists call them measurement errors.

Reducing such errors in household surveys should lead to better research and better policies. Personal computers are a useful tool for this purpose. Data from survey questionnaires can be entered into personal computers using

Paul Glewwe (corresponding author) is a professor in the Department of Applied Economics at the University of Minnesota; his email address is pglewwe@umn.edu. Hai-Anh Hoang Dang is a consultant in the Policy Research Group at the World Bank; his email address is hdang@worldbank.org. The authors would like to thank the staff of Vietnam's General Statistical Office for their cooperation in this research. The authors also benefited from comments and information provided by Graham Kalton, Frank Martin, Rob Swinkels, Phung Duc Tung, and three anonymous reviewers. Glewwe thanks INRA-LEA (Fédération Paris-Jourdan) for hospitality in the fall of 2004, when the first draft of this article was written. A longer version of this article is available at http://wber.oxfordjournals.org/.

THE WORLD BANK ECONOMIC REVIEW, VOL. 22, NO. 1, pp. 165–185 doi:10.1093/wber/lhm023
Advance Access Publication January 28, 2008

software that detects data errors (see, for example, Grosh and Muñoz 1996). Many detected errors can be resolved by returning to the households to gather more accurate information.

Using personal computers to reduce nonsampling errors in developing countries is most effective when data entry takes place near the household, soon after the interview, because survey teams often work only a week or two in one area before moving to another, at which point it difficult for them to return to the previous area to correct errors detected by data-entry programs. Providing personal computers to survey teams could avoid this problem,[1] but doing so can be very expensive for statistical offices in developing countries. Before purchasing more computers, these statistical offices need to know the impact of buying more computers on the errors in the survey data they collect.

This article analyzes a randomized experiment conducted in Vietnam in 2002. Computers for decentralized data entry were randomly provided to some districts and not others. Such experiments are rare in developing countries. The first set of results focuses on per capita expenditures, as the survey's main purpose is to use expenditure data to measure poverty and household welfare. The data show no evidence that providing computers to each district affected measured per capita expenditures.

A second set of results examines errors and inconsistencies in other variables. Decentralized data entry reduces errors by 5–23 percent, depending on the type of error. The use of desktop computers in high population density areas has the largest impact. The gains in data quality from using more personal computers are not cost-effective, however: raising the sample size would reduce measured variance at a much lower cost. Yet providing computers does reduce the time interviewers spend manually checking the data, and the value of the time saved is close to the cost of a desktop computer.

This article is organized as follows. Section I describes the design of the experiment. Section II explains how to test for the impact of decentralized data entry and defines probable errors in the data. Section III reports the results, and section IV discusses the implications for data collection in Vietnam. Section V summarizes the findings and offers recommendations for developing country statistical offices.

I. Design of the Experiment

In 2002 Vietnam's General Statistical Office (GSO) conducted the first Vietnam Household Living Standards Survey (VHLSS). Its main purpose is to monitor poverty and other indicators of household welfare in Vietnam (GSO 2006).

1. The distance could be reduced to zero by providing interviewers with laptop computers, to enable them to enter data during the interview. This is called computer-assisted personal interviewing (CAPI). The decentralized data entry examined here did not use this method.

The 2002 VHLSS collected information from about 75,000 households from all of Vietnam's 61 provinces. Of these households, 45,000 were administered a questionnaire that solicited information on income and basic household characteristics. The other 30,000 were interviewed using a questionnaire that solicited the same information plus data on household expenditures. This study focuses on the 30,000 households interviewed using the second questionnaire. The 45,000 households that completed the first questionnaire are excluded because they were interviewed in the first half of 2002, whereas the experiment was implemented in the last three months of 2002.

The sample for the 2002 VHLSS was drawn as follows. Vietnam's 61 provinces were divided into urban and rural areas, creating 122 strata. Vietnam is divided into about 10,000 communes. The GSO drew a master sample of 3,000 communes, with the number of communes per stratum determined by the square root of the number of households in those strata. Each selected commune was divided into several enumeration areas. One enumeration area was randomly selected from each of the 3,000 communes, with selection probabilities proportional to the number of households per area. From each selected enumeration area, 20 households were randomly selected, yielding a sample of 60,000 households. In early 2002 the sample was increased to 75,000 households by selecting five additional households from each enumeration area. The 30,000 households that completed the second questionnaire are a representative sample of all of Vietnam.

All VHLSS survey teams have two interviewers and one supervisor. Each district has a single team, which conducts all the interviews in its district.

This study estimates the impact on data quality of decentralized data entry using personal computers. More precisely, it examines the impact of providing data-entry computers at the district level, rather than the province level, for the 2002 VHLSS. Providing computers to each district requires 10 times as many computers.[2]

The VHLSS data are entered into computers at Province Statistical Offices in the provincial capital. District-based teams interview all sampled households in an enumeration area and check the data collected manually. This manual checking is very time consuming. Each district then sends the completed household questionnaires to the provincial capital. Data entry is done at Provincial Statistical Offices, using personal computers with a data-entry program that detects more than 100 potential errors and inconsistencies.

Under this system survey teams have few opportunities to return to households to clear up errors or inconsistencies detected by the software when data are entered at provincial capitals. In contrast, the decentralized data-entry system evaluated here allowed survey teams to return quickly to households to correct errors or inconsistencies detected by the software. Because data can be

2. In 2002 Vietnam had 607 districts and 61 provinces. Since then some provinces and districts have been split, increasing the numbers of both.

entered into computers in each district one or two days after an interview, errors and inconsistencies are detected by the software while a survey team is still working in the enumeration area. Thus teams can return to households for which the software detects errors, correct the errors, and enter the corrected data into the computer.

To estimate the impact of decentralized data entry, a classic experimental design was used in 23 provinces during the fourth quarter (October–December) of the 2002 VHLSS. In those provinces all districts included in the fourth quarter of the survey participated in the experiment, a sample of 2,895 households in 202 districts. These districts were divided into "dispersed" and "compact" districts. "Dispersed districts" were districts too dispersed geographically to allow interviewers to return to the District Statistical Office in the district capital within 24 hours of the interview for some or all of the sampled households. Of the 202 districts in the "population" of the 23 provinces, 97 were classified as dispersed. Ten were randomly selected to be treatment districts; the other 87 were controls. The teams in the 10 treatment districts received laptop computers for entering data from household questionnaires at the enumeration area within hours of each interview.

The other 105 districts in the 23 provinces were "compact districts"—districts small enough so that interviewers could return to the District Statistical Office within 24 hours of the interview. Fifteen districts were randomly selected as treatment districts; the other 90 were controls. Each treated district received a desktop computer, kept at the District Statistical Office in the district capital, which permitted data to be entered a day or two after each interview.

Thus 25 of the 202 districts in the 23 provinces were randomly provided laptop or desktop computers to speed data entry and allow interviewers to return to any households for which the software detected unusual or inconsistent data. No additional training was given to interviewers or supervisors in the districts that received computers.

II. Testing the Impact of Decentralized Data Entry

This section describes the statistical methods used to estimate the impact of providing data-entry computers to each district in Vietnam. It also explains how the data were examined to detect nonsampling (measurement) errors.

Statistical Methods Used to Estimate Impact

The impact on data quality of changing a household survey's methodology can be assessed in two ways. First, one can estimate the impact of changes on observed means, variances, and other functions of variables of interest. Second, one can investigate how changes affect the prevalence of data errors. This subsection discusses statistical methods for both approaches.

Consider a variable y. It could be household income or expenditure, the number of errors in a completed questionnaire, a dummy variable indicating a household's poverty status, or the household's "poverty gap." (The poverty gap is the poverty line minus household income; it equals zero if this difference is negative.) The objective is to test whether the observed mean of y varies across the treatment and control groups. If, for example, y measures poverty, more accurate poverty measurement from improved data collection could induce reallocations of government resources to regions or households for which current data underestimate poverty. Examples of such research are Chaudhuri and Ravallion (1994) and Gibson and others (2003).

Of course, changes in the observed mean of y as a result of new data collection methods do not necessarily imply that the new estimate is better; a change could increase bias in the estimated mean. If the observed mean has changed, the new data collection method must be carefully considered to determine whether bias has increased or decreased. The rest of this subsection explains how the observed means of y were checked to determine whether they changed after decentralized data entry was implemented in Vietnam.[3]

If y has a finite mean and variance, then by the Lindberg–Levy Central Limit Theorem its sample mean is asymptotically normally distributed; the asymptotic variance is $\mathrm{Var}[y]/n$ (n is the sample size). If the variance of y is equal across the treatment and control groups, the difference in the sample means of those two groups is also asymptotically normally distributed, with an asymptotic variance $\mathrm{Var}[y](1/n_t + 1/n_c)$, where n_t is the sample size of the treatment group and n_c is the control group. A potentially more precise test compares the difference in the means of the treatment and control groups of these households (which were interviewed in the fourth quarter of the survey) with the same difference for households interviewed in the first three quarters, a difference in differences test. If y has the same variance across all four groups, the double difference $(\bar{y}_{t,\,4} - \bar{y}_{c,\,4}) - (\bar{y}_{t,\,123} - \bar{y}_{c,123})$ is asymptotically normal, with a variance of $\mathrm{Var}[y](1/n_{t,4} + 1/n_{c,4} + 1/n_{t,123} + 1/n_{c,123})$.

Regression analysis can test differences in the mean of y across treatment and control groups. Using only the fourth-quarter data, ordinary least squares regression of y on a constant and a dummy variable indicating the treatment group produces a coefficient on that variable with a t-statistic exactly equal to the test statistic for the difference in means described above. The test using all four quarters of data is replicated by regressing y on a constant, a dummy variable indicating residence (in any quarter) in a district that was treated in the fourth quarter, a dummy indicating households interviewed in the fourth quarter, and an interaction term between both variables. The coefficient on the interaction term tests whether the treatment affected the mean of y.

3. For expositional ease the discussion refers to the variable y; more precisely, it concerns observed measurements of y.

Adding control variables to these regressions may improve the precision of the estimated coefficients. Since the treatment variable is uncorrelated with any variable, adding regressors will not lead to biased estimates of treatment effects.

The first regression uses only the 2,895 households that participated in the experiment in the fourth quarter of the survey. Of these, 2,475 are in the control group and 420 are in the treatment group. They can be used to estimate

$$y = \beta_0 + \beta_1 P + \boldsymbol{\beta}_2' \boldsymbol{x} + \varepsilon \tag{1}$$

where y is the variable of interest, P is the "program" dummy variable that indicates households living in districts that received computers, $\boldsymbol{x}$ is a vector of control variables, and ε is an error that is uncorrelated with P because P was randomized. All estimated standard errors in this article account for clustering at the enumeration area level.

Ignoring the $\boldsymbol{x}$ covariates, β_1 in equation 1 equals $E[\bar{y}_{t,\,4} - \bar{y}_{c,\,4}]$, the difference in the mean of y across the treatment and control districts for the 2,895 households interviewed in the fourth quarter of 2002. That is, $E[\bar{y}_{t,\,4}] = \beta_0 + \beta_1$ and $E[\bar{y}_{c,\,4}] = \beta_0$, so $E[\bar{y}_{t,\,4} - \bar{y}_{c,\,4}] = \beta_1$.

The second regression method adds the households interviewed in the first three quarters (January–September, 2002) in the 202 districts that participated in the experiment in the fourth quarter:

$$y = \beta_0 + \beta_1 Q_4 + \beta_2 P + \beta_3 Q_4 P + \boldsymbol{\beta}_4' \boldsymbol{x} + \varepsilon \tag{2}$$

where Q_4 is a dummy variable indicating households interviewed in the fourth quarter. Ignoring $\boldsymbol{x}$, $\beta_3 = E[(\bar{y}_{t,\,4} - \bar{y}_{c,\,4}) - (\bar{y}_{t,\,123} - \bar{y}_{c,\,123})]$, the difference between the difference of the mean of y across the treatment and control districts in the fourth quarter and the analogous difference in the first three quarters. That is, ignoring covariates, equation 2 implies $E[\bar{y}_{t,\,4}] = \beta_0 + \beta_1 + \beta_2 + \beta_3$, $E[\bar{y}_{c,\,4}] = \beta_0 + \beta_1$, $E[\bar{y}_{t,\,123}] = \beta_0 + \beta_2$ and $E[\bar{y}_{c,\,123}] = \beta_0$, so $E[(\bar{y}_{t,\,4} - \bar{y}_{c,\,4}) - (\bar{y}_{t,\,123} - \bar{y}_{c,\,123})] = (\beta_2 + \beta_3) - \beta_2 = \beta_3$.

Adding a $P\boldsymbol{x}$ term to equation 1 and a $Q_4 P\boldsymbol{x}$ term to equation 2 allows one to estimate whether the program impact varies by household characteristics ($\boldsymbol{x}$).[4] Finally, separate regressions can be estimated for dispersed and compact districts, to see if the impacts vary across these two subexperiments.

4. In theory, one should also add $Q_4\boldsymbol{x}$ and $P\boldsymbol{x}$ as regressors to equation 2, but β_2 should equal zero (the program should have no impact before its implementation), so adding $P\boldsymbol{x}$ is unnecessary. $Q_4\boldsymbol{x}$ was added but later dropped, because it was seldom significant.

Direct Detection of Nonsampling (Measurement) Errors

Examining a variable of interest to see whether changing survey procedures changes its mean or another function of its distribution *indirectly* assesses whether those changes affected errors in household survey data, because that approach does not search for explicit errors. In one sense, such indirect evaluation of changing survey procedures is all that is required, because only changes in the observed distributions of those variables affect policy decisions. Yet it is also instructive to search for measurement errors (or likely errors) directly in variables of interest and even in other variables, because this information can reveal whether particular types of households tend to yield error-ridden data.

This section explains how to analyze measurement errors directly. Almost all variables from household surveys have errors, but only some errors are apparent in the data. Errors can be divided into those that can be detected by analyzing survey data and those that cannot. Examples of detectable errors are an age of 150 years or consumption of 10 kilograms of meat per person per day. Perhaps more common are inconsistencies between two or more variables, such as a five-year-old child who is married. Examples of undetectable errors are an incorrect age of an adult (if it does not contradict other variables, such as date of birth) and errors in income or consumption expenditures (assuming no contradictions with other data).

The distinction between detectable and undetectable errors is not always sharp. Some information may be doubtful but still possible, such as a 15-year-old who reports having finished secondary school or a farmer who reports unusually high fertilizer use per hectare. In practice, one can treat dubious data anomalies that have a small probability of being correct as detectable errors and classify errors that produce unusual (but still plausible) data patterns as undetectable errors. Dividing errors into these two types is subjective, but if the judgments made are explicit, the results should not be misleading.

By definition, undetectable errors cannot be directly analyzed by examining the data; they can be examined only by reinterviewing households soon after the initial interview or by checking other data sources (for example, employer's payroll data).[5]

Fortunately, detectable errors can be directly analyzed. It is useful to divide them into two subtypes: those explicitly checked by the 2002 VHLSS data-entry program and those that were not checked by that program but can be checked by examining the data. This distinction is important, because the data-entry software determines how much can be learned from analyzing each subtype of detectable errors. In particular, some data entry programs force the

5. Undetectable errors can also be analyzed indirectly, by looking at changes in the distributions of the variables of interest after changing survey methods.

data to be consistent with all explicit error checks, so the data provided by the software are "error free"—that is, the data will never have detectable errors of the type checked by the software. Most important, if survey teams cannot return to households to clear up errors detected by the software, data-entry operators or other survey staff must change the data in the office to remove all apparent errors. Such changes could leave in place real errors if the "correction" relies on guesswork, which is often the case. Thus checking only for errors that the software checks yields no information on the impact of survey design changes on nonsampling errors. In contrast, if the software does not force data-entry staff to "fix" errors that it detects and staff are not instructed to "correct" all detected errors, the effect of providing computers to each district may appear in the errors that the program explicitly checks (as well as other detectable errors, as explained below).

Regardless of whether the software forces survey staff to "correct" the errors it finds, one can still detect improvements in data quality by examining errors not explicitly checked by the software. Intuitively, if moving data entry closer to where households are interviewed allows interviewers to return to households to clear up errors detected by the program, the corrected data may reduce errors that are not explicitly checked by the software. For example, the data-entry program may find that the sum of the area of a household's plots of land does not equal the reported total amount. Explicit checking may detect an error in a plot's area, and correcting that error could lead to more accurate crop yields on that plot of land. This article examines both types of detectable errors.

The 2002 VHLSS income and expenditure household questionnaire has nine sections (box 1). The VHLSS data-entry program performed more than 100 data checks. Space constraints preclude describing any but the most common. First, several questions involved summing numerical values from specific questions; many data checks in the program verified those sums. Second, a few checks verified whether the number of people with information from a given questionnaire section matched the people listed in the household roster (Section 1). For example, Section 3 (employment) was to be completed for all household members aged 10 and older; one check is whether Section 3 included information for everyone in the household roster age 10 or older. Third, several data-entry checks involved questions about household participation in certain production activities. If the household reported participation, the data-entry checks looked for the expenditures on or income from those activities; if the household reported no involvement in an activity, the expenditures or income for that activity should be zero (or missing).[6] Fourth, for households reporting land owned or livestock or crops sold or used for other purposes, the sum of

6. The questionnaire uses skip codes, so households reporting no involvement in an activity will have missing values for that activity's income and expenditure variables. In practice, the skip codes were almost always followed, and missing data that should not be missing are rare.

the disaggregated amounts (for example, certain types of land or specific uses of livestock or crops) should equal the total amounts reported for all types or uses. Fifth, many questions had preassigned codes (such as occupation codes or industry codes); the software checked that the values recorded in the data matched one of these codes. The program does not "force" data to be consistent; data-entry personnel can enter data that violate the checks performed by the data-entry program.

BOX 1. STRUCTURE OF 2002 VHLSS INCOME AND EXPENDITURE HOUSEHOLD QUESTIONNAIRE

Section 1: List of Household Members (Roster)
Section 2: Education
Section 3: Employment
Section 4: Health
Section 5: Income and Other Inflows of Money
- Part A. Income from Salary and Wages
- Part B. Agricultural, Forestry, and Fishery Activities
  1. Agricultural, Forest, and Aquaculture Land
  2. Agricultural Production
  3. Income from Livestock
  4. Income on and Expenses for Farm Services
  5. Income from and Expenditure on Tree and Forest Crops
  6. Income from and Expenditure on Aquaculture
- Part C. Nonfarm Businesses and Processing Farm, Forest, and Fishery Products
- Part D. Other Sources of Income

Section 6: Expenditure
- Part A. Expenditure on Food and Drink
- Part B. Expenditure on Nonfood and Other Expenditures
- Part C. Other Expenses (contributions, taxes, and so forth)
- Part D. Other Expenses not Included as Expenditure (savings, insurance, and so on.)

Section 7. Fixed Assets and Durable Goods
Section 8. Housing
Section 9. Participation in Poverty Alleviation Programs

Source: GSO 2006.

The 2002 VHLSS data-entry program undoubtedly detected many errors. The protocol for correcting errors (for control group households) was the following. The supervisor at the provincial capital, where all data were entered, corrected "obvious" errors in the provincial office. For errors without obvious resolutions, the supervisor was to telephone the district office to ask the district

team to investigate, including (if necessary) by revisiting the household. Little is known about what actually happened, but examination of the data for the types of errors detected by the program reveals very few errors (about 0.03 per household). Thus the data-entry program was used to "clean" almost all errors explicitly checked by that program. The extent to which the corrections were valid, as opposed to "educated guesses" by survey teams at the provincial and district levels, is unclear; discussions with survey staff suggest that they made very few revisits to households.

Almost every section of the questionnaire has variables, or combinations of variables, that were not checked by the data-entry program but could have errors or inconsistencies. For example, Section 1 collected data on gender, relationship to the household head, date of birth, age, and marital status. Possible inconsistencies among these variables include age and date of birth; people identified as the head's spouse who report not being married; people reported as children of the head who are no more than 14 years younger (or more than 50 years older) than the head; and the head and his or her spouse reporting the same gender. Inconsistencies may also exist across sections, such as young children with high levels of education, households that report being self-employed farmers but report neither owning nor renting land, and households that report consuming a certain crop grown by the household but not reporting growing that crop in the agriculture section. (The appendix in the long version of this article, available at http://wber.oxfordjournals.org/, presents an extensive list of errors, inconsistencies, and likely errors that were not checked by the 2002 VHLSS software.)

Several control variables (x) were added to the regression analysis of detectable errors. First, to control for unobserved quality in survey personnel, province dummy variables were added (district dummy variables yielded similar, but less precise, results). Second, household size was added; because many questionnaire sections collect information on each household member, larger households provide more information and thus have more opportunities for errors. Third, per capita expenditure was added; wealthier households buy more goods, increasing the amount of information collected and thus increasing opportunities for errors.

III. Empirical Results

This section uses the methodology described above to estimate the impact of decentralized data entry on the measured distribution of per capita expenditures and in the number of errors in the data. Per capita expenditure is of particular interest, because the VHLSS uses that variable to monitor poverty in Vietnam.

Impact on the Mean and Variance of per Capita Expenditures and on Poverty

Vietnam's GSO implements the VHLSS every two years. The results on poverty receive widespread attention both inside and outside Vietnam. This subsection

examines whether introducing computers in each district changed the observed distribution, and functions of the distribution, of per capita expenditures.

If measurement errors were random (uncorrelated with per capita expenditures), with a zero mean, then reductions in measurement errors from providing computers to each district would not affect the mean of observed per capita expenditures but would reduce the observed variance. Since most Vietnamese are not poor, reduced variance would reduce measured poverty (Ravallion 1994). Yet measurement errors may not be random, so decentralized data entry could affect the mean of observed expenditure, leaving the impacts on variance and poverty ambiguous.

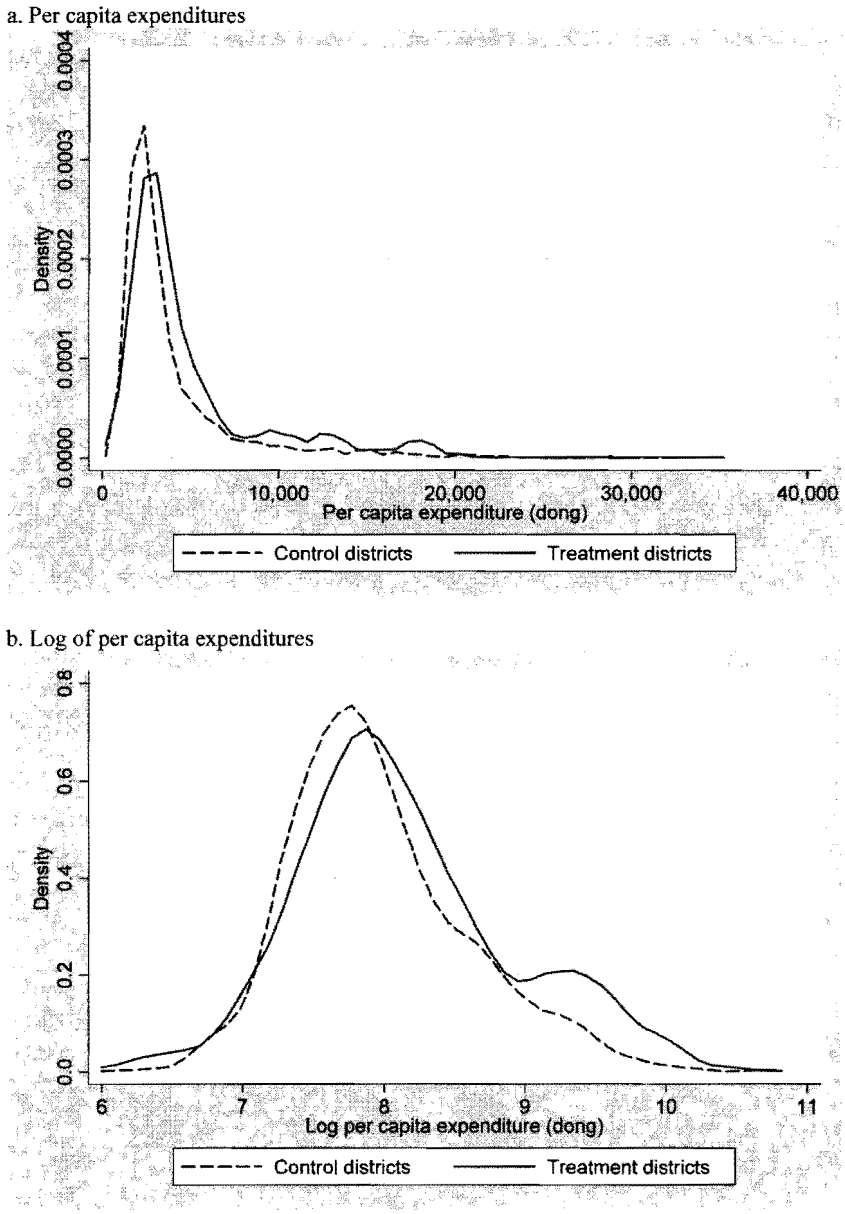
To assess the impact of decentralized data entry on measured per capita expenditures, the density function of that variable can be compared for the treatment and control households in the fourth quarter of the survey. In both panels of figure 1 the density for treatment households lies to the right of the density for control households, but only slightly so. The densities for log per capita expenditures suggest that the variance of the treatment households may be larger than the variance for control households, albeit only slightly. Indeed, Kolomogorov–Smirnov tests fail to reject, even at the 10 percent level, the null hypothesis that the distributions of per capita expenditure for the treatment and control groups are equal.

Next consider parametric tests for particular summary statistics, starting with the variance (table 1). The standard deviation of expenditures per capita is larger for the treatment group than for the control group, indicating that any "corrected" errors tended to regress to the mean (positive errors for low per capita expenditures and negative errors for high values). The Brown and Forsythe (1974) robust test of equality of variances (which is needed, because figure 1 shows that the distribution of expenditures per capita is skewed) reveals that this difference is statistically significant (p-value of 0.04).

This statistical significance is doubtful, however, for several reasons. First, applying that test to log per capita expenditures fails to reject the null hypothesis of equal variances at the 5 percent level (though it does at the 10 percent level). Second, and more important, these tests are biased toward rejecting the null hypothesis, because they ignore the fact that the data are clustered and weighted, which (if properly accounted for) raises the standard error of almost any test statistic. Third, the statistical significance of the difference in variances was also checked by bootstrapping (accounting for clustering); the null hypothesis of equal variances was not rejected, even at the 10 percent level. Thus there is little evidence that providing computers to each district affected the variance of per capita expenditures.

Turn now to the mean of per capita expenditures. By definition, random measurement errors have a zero mean and thus cannot affect the mean. If their mean is not zero, and decentralized data entry reduces measurement errors, the mean of observed expenditures should differ across treatment and control

FIGURE 1. Kernel Density Estimates of per Capita Expenditures



Note: Automatic bandwidths are used for all kernel density estimates.
Source: Authors' analysis based on data described in text.

groups. The second line of table 1 shows that the mean for the treatment group is 3,911 thousand dong, about 7.5 percent higher than the mean for the control group (3,636 thousand dong). The difference (275 thousand dong) is statistically insignificant (using tests that incorporate clustering and sample

TABLE 1. Differences in Distribution of per Capita Expenditures

Statistic	Treatment group	Control group	Difference
Standard deviation (thousands of dong)	4,034.0	3,680.5	353.5
Mean (thousands of dong)	3,910.9	3,636.4	274.6 (416.7)
Poverty rate	0.250	0.269	−0.019 (0.037)
Poverty gap	0.061	0.060	0.001 (0.012)
Squared poverty gap	0.024	0.020	0.004 (0.007)

Note: Numbers in parentheses are standard errors of the differences in means. The difference for the standard deviation is statistically insignificant, as explained in the text. Sample size is 2,895 for all rows.

Source: Authors' analysis based on data described in the text.

weights). Thus provision of computers had little effect on the mean of measurement errors (whether zero or nonzero) in per capita expenditure, leaving unchanged any bias caused by those errors.

Introducing computers at the district level could affect poverty, as measured by per capita expenditures, even if it does not affect mean expenditures. Districts that received computers have a slightly lower poverty rate, but the poverty gap was almost identical and the squared poverty gap (which is sensitive to inequality among the poor) was slightly higher. None of these differences is statistically significant, suggesting that providing computers to each district had no effect on measured poverty.

Double-difference estimates of the impact of decentralized data entry may be more precise and therefore more likely to detect any impacts (table 2). The standard errors of these double-difference estimates (equation 2, without covariates) for mean per capita expenditures and the three poverty indices reveal no increase in precision, and all differences remain insignificant.

TABLE 2. Double-Difference Estimates of Change in per Capita Expenditures

	Rounds 1–3		Round 4		
Statistic	Treatment group	Control group	Treatment group	Control group	Double-difference
Mean (thousands of dong)	3987.6	3135.5	3762.2	3704.2	−794.1 (592.3)
Poverty rate	0.253	0.282	0.259	0.260	0.027 (0.048)
Poverty gap	0.053	0.061	0.061	0.056	0.012 (0.014)
Squared poverty gap	0.016	0.019	0.023	0.018	0.008 (0.007)

Note: Numbers in parentheses are standard errors of the differences in means. Sample size is 11,040 for all rows. Figures for Round 4 are slightly different from those in table 1 because 40 observations were dropped from five districts that were surveyed in Round 4 but not in Rounds 1–3.

Source: Authors' analysis based on data described in the text.

TABLE 3. Regression Estimates of Program Impacts on Error Rates

	Mean errors per household			Regression coefficients on computer-variable	
Source of errors	Quarters 1, 2, and 3 (before experiment)	Quarter 4 control group	Quarter 4 treatment group (with district-level computers)	Level regression (quarter 4 only)	Double-difference regression (all quarters)
All individual sections	0.256	0.229	0.190	−0.183 (0.201)	−0.249 (0.196)
Household sections	0.073	0.077	0.047	−0.479 (0.310)	−0.434 (0.303)
All sections	0.329	0.306	0.236	−0.278* (0.164)	−0.307* (0.162)

*Statistically significant at the 10 percent level.

Note: Results are negative binomial regression estimates. All figures are for the 23 provinces that participated in the experiment. All regressions include a computer dummy variable, province dummy variables, household size, and expenditure per capita. Double-difference regressions add quarter dummy variables. Numbers in parentheses are standard errors, which account for clustered sampling. Sample sizes are 2,895 for the level regressions and 11,040 for double-difference estimates.

Source: Authors' analysis based on data described in the text.

To summarize, both parametric and nonparametric methods find no impact of decentralized data entry in Vietnam on the distribution of per capita expenditures or poverty indices. Point estimates of differences are small, especially in table 1. Providing data-entry computers to each district is thus unlikely to generate results that would change Vietnam's economic policies.

Impact on Detectable Errors

Level and double-difference estimates are presented for the full sample (compact and dispersed districts), with each row representing a separate negative binomial regression (table 3).[7] The fourth and fifth numbers in each row are the negative binomial coefficients that estimate the impact on errors of providing district-level computers; the fourth number is an estimate of β_1 in equation 1, and the fifth is an estimate of β_3 in equation 2. If decentralized data entry reduces errors, these coefficients should be negative. Standard errors are given in parentheses.

The first row examines all errors for individual level sections of the VHLSS questionnaire. The mean number of errors for all these sections is 0.256 for quarters 1–3 and 0.229 for the fourth-quarter control group. Both are larger than the treatment group mean (0.190), but these differences are statistically insignificant in both the level and double-difference regressions. The second row examines all errors in the household-level sections of the VHLSS questionnaire. The mean

7. All regressions were also estimated using ordered probits; the results were very similar. Additional estimates for individual sections of the VHLSS questionnaire are given in the online version of this article, available at http://wber.oxfordjournals.org/.

errors for quarters 1–3 (0.073) and the fourth-quarter control group (0.077) are noticeably higher than for the fourth-quarter treatment group (0.047), but the estimated impact falls just short of significance at the 10 percent level.

The third row sums both individual- and household-level errors. Among households interviewed in quarters 1–3 there are 0.329 errors per household; for the fourth-quarter control group there are 0.306. The figure for districts given computers in the fourth quarter is 0.236. The last two numbers suggest that decentralized data entry reduced aggregate errors by 23 percent, a statistically significant difference (at the 10 percent level) for the level and double-difference estimates. A simple *t*-test of the difference in means across the treatment and control groups in the fourth quarter is also statistically significant (at the 10 percent level).[8]

These results combine compact districts, which received desktop computers, and dispersed districts, which received laptops. The impact of providing computers could vary by type of computer or type of district. Survey teams in dispersed districts travel to the enumeration areas and stay there for several days before returning to the district capitals. The working conditions are difficult, suggesting potential for large reductions in errors from bringing data entry closer to interview sites. On the other hand, using laptops in these areas could entail additional difficulties, such as the lack of electricity for recharging batteries or the damage caused to computers by moving them frequently.

For dispersed districts, in the individual sections of the questionnaire there is no apparent impact of providing computers on total errors: the fourth-quarter control group had slightly fewer errors (0.251) than the treatment group (0.266) (table 4). The regression estimates show a statistically insignificant negative impact of providing laptops on individual-level errors. Yet the mean errors on the household-level sections in the treatment group (0.045) are less than half of the mean in the fourth-quarter control group (0.105), a statistically significant effect (at the 5 percent level). With both types of errors combined, the mean errors for treatment group households (0.311) is 13 percent lower than the error for the fourth-quarter control group (0.356), but it is statistically insignificant.

Providing desktop computers to compact districts reduces errors on individual-level sections by 41 percent, from 0.213 to 0.126 (table 5). This is statistically significant (at the 5 percent level) in the level regression but not in the double-difference regression. The drop in household-level errors is smaller (about 21 percent) and statistically insignificant. Combining both types of errors yields a 37 percent reduction in errors (from 0.266 to 0.168), which is statistically significant at the 5 percent level in the level regression and at the 10 percent level in the double-difference regression.

8. The online version of this article examines errors of consistency between what households report as consumption from own production in the food expenditures section of the VHLSS questionnaire and what they report regarding crops grown and animals raised in the farming section. There was no discernable impact on such errors of providing computers to each district, but these data were very noisy and some apparent errors may not be errors at all.

TABLE 4. Regression Estimates of Program Impacts on Error Rates: Dispersed Districts

	Mean errors per household			Regression coefficients on computer-variable	
Source of errors	Quarters 1, 2, and 3 (before experiment)	Quarter 4 control croup	Quarter 4 treatment group (with district-level computers)	Level regression (quarter 4 only)	Double-difference regression (all quarters)
All individual sections	0.235	0.251	0.266	−0.254 (0.365)	−0.237 (0.330)
Household sections	0.077	0.105	0.045	−0.928** (0.438)	−0.838** (0.415)
All sections	0.313	0.356	0.311	−0.427 (0.305)	−0.383 (0.285)

**Statistically significant at the 5 percent level.

Note: Results are negative binomial regression estimates. In addition to the computer dummy variable, all regressions include province dummy variables, household size, and per capita expenditures. The double-difference regressions add quarter dummy variables. Numbers in parentheses are standard errors, which account for clustered sampling. The sample size was 1,285 for the level regressions and 4,840 for the double-difference estimates.

Source: Authors' analysis based on data described in the text.

The different results in the dispersed and compact districts make sense. Dispersed districts are very rural, and almost all adults are self-employed farmers; most households complete the detailed household-level agricultural section. Because they are less likely to work for wages, seek medical care, or

TABLE 5. Regression Estimates of Program Impacts on Error Rates: Compact Districts

	Mean errors per household			Regression coefficients on computer-variable	
Source of errors	Quarters 1, 2, and 3 (before experiment)	Quarter 4 control group	Quarter 4 treatment group (with district-level computers)	Level regression (quarter 4 only)	Double-difference regression (all quarters)
All individual sections	0.282	0.213	0.126	−0.573** (0.291)	−0.476 (0.303)
Household sections	0.068	0.053	0.042	−0.290 (0.496)	−0.133 (−0.484)
All sections	0.350	0.266	0.168	−0.536** (0.224)	−0.439* (0.230)

*Statistically significant at the 10 percent level.

**Statistically significant at the 5 percent level.

Note: Results are negative binomial regression estimates. In addition to the computer dummy variable, all regressions include province dummy variables, household size, and per capita expenditures. The double-difference regressions add quarter dummy variables. Numbers in parentheses are standard errors, which account for clustered sampling. The sample size was 1,605 for the level regressions and 5,660 for the double-difference estimates.

Source: Authors' analysis based on data described in the text.

send their children to school, they provide less information on the individual sections. Because they provide more data on household-level sections and less data on individual-level sections, they are more likely to have household-level errors and less likely to have individual-level errors than are households in the (more-urbanized) compact districts. This tendency, seen in the first columns of tables 4 and 5, implies that improved data entry will have greater effects on household-level errors in dispersed districts and on individual-level errors in compact districts.

Does the impact of providing computers vary across households? The question can be investigated using regression analysis by adding interaction terms between the program dummy variable (P) and certain household characteristics that may increase errors. One example is household size: more household members giving individual-level information implies more opportunities for errors. Another is per capita household expenditures: better-off households purchase more food and nonfood goods, own more durables, and use more health services, increasing the data collected and thus the possibilities for errors. A third example is the education of the household head (the primary survey respondent): better-educated household heads may make fewer errors answering household questions. Surprisingly, none of these three interaction terms is statistically significant; no observable household characteristics appear to increase or reduce the incidence of errors.

IV. Implications for Data Collection in Vietnam

Providing data-entry computers to each district (instead of each province) has little effect on the measured distribution of per capita expenditures and thus little effect on poverty estimates for Vietnam. Providing computers to districts does reduce detectable errors in the VHLSS for many variables, however. When errors of consistency for reported consumption of home-produced crops and livestock are included, introducing computers reduces errors by only (a statistically insignificant) 5 percent. When such errors are excluded, providing computers to districts reduces errors by (a statistically significant) 23 percent. Errors are reduced in both dispersed districts (which received laptops) and compact districts (which received desktops), but the drop in errors was larger in compact districts (37 percent) than in dispersed districts (13 percent). The impact of computers did not vary by household characteristics.

Does decentralized data entry merit the cost of purchasing computers for all 607 districts in Vietnam? The GSO spends about $500 on desktop computers and about $1,200 on laptops; both require $100 in training costs. About one-third of Vietnam's districts are dispersed and thus require laptops. Assuming that 407 desktop computers and 200 laptops are needed, the total cost is $504,200.

It is harder to value the benefits of purchasing computers. None was found in terms of improved measurement of poverty or per capita expenditures. But it appears that providing computers reduces errors for many other variables.

To assess the value of reducing these errors, consider a random variable x and two alternatives for reducing the standard error of its estimated mean: increasing the sample to reduce the contribution of sampling error to the standard error (of the estimated mean of x) or providing computers to reduce nonsampling errors. The standard error of the mean of x is the standard deviation of x divided by the sample size: $SD(x)/\sqrt{N}$.[9] Assume that the errors have a mean of zero and are uncorrelated with the true value of x.[10] Providing computers reduces $SD(x)/\sqrt{N}$ by reducing random measurement error and thus reducing $SD(x)$; increasing the sample reduces $SD(x)/\sqrt{N}$ by increasing $\sqrt{N}$.

Consider the likely impact on $SD(x)$ of providing computers to each district. Assume that all errors, including those corrected by providing computers, have a zero mean, are uncorrelated with the true x, and have a standard deviation equal to $SD(x)$. The last assumption likely overestimates the true standard deviation of these errors. In particular, the difference between two observations randomly drawn from a normal distribution (the error from replacing the value of x for one household with that of another randomly selected household) is $0.798 \times SD(x)$ (Johnson and others 1994), so what follows is an upper bound of the impact on $SD(x)/\sqrt{N}$ of providing computers to each district. Finally, for simplicity suppose that x is usually measured without error, but for a small fraction of the observations (denoted by k) the observed x is the sum of the true value and the random error with a standard deviation of $SD(x)$. Thus the standard deviation of observed x equals $\sqrt{1+k}$ multiplied by the true standard deviation of x.[11]

Providing desktops to compact districts reduced the mean errors per household from 0.266 to 0.168 (see table 5). This result was based on checking about 40 variables, which implies a reduction in the error rate (k) from about

9. This standard error assumes a simple random sample. Yet almost all household surveys, including the VHLSS, have a multistage clustered sample design. The analysis presented here still holds, because the correct standard error of the estimated mean is $\rho SD(x)/\sqrt{N}$, where ρ is the design effect, which is unchanged when either more computers are added or the sample size is increased using the same multistage clustered design.

10. It is impossible to determine whether measurement errors are biased (have a nonzero mean), but the results presented in section II reveal no evidence that providing computers affected any bias that exists for per capita expenditures. It is also impossible to determine whether measurement errors are correlated with the true values of variables. Yet in most cases correlated errors raise the variance of observed x, so the following analysis is relevant. To see this, let x^* denote the true value of x and let u be the measurement error, so x, the observed value, is $x^* + u$. Assume a linear correlation between u and x^*: $u = \beta(x^* - \mu_{x^*}) + \varepsilon$, where μ_{x^*} is the mean of x^* (to ensure that u has a zero mean) and ε is a random error uncorrelated with x^*. Thus $\mathrm{Var}(x) = \mathrm{Var}(x^* + u) = \mathrm{Var}(x^*(1+\beta) - \beta\mu_{x^*} + \varepsilon) = (1+\beta)^2\mathrm{Var}(x^*) + \mathrm{Var}(\varepsilon)$. Clearly, $\mathrm{Var}(x) > \mathrm{Var}(x^*)$ if the correlation is positive ($\beta > 0$). For negative correlation ($\beta < 0$), $\mathrm{Var}(x) < \mathrm{Var}(x^*)$ only if $(2\beta - \beta^2)\mathrm{Var}(x^*) > \mathrm{Var}(\varepsilon)$. Intuitively, for $\mathrm{Var}(x) < \mathrm{Var}(x^*)$ the impact of the negative correlation must outweigh that of the random component (ε) of u.

11. Given these assumptions, the variance of observed x is the variance of the "true" values plus the variance of the random errors that occur for k percent of the sample. The variance of the random errors that occur for k percent of the sample variance equals the variance of the true values of x, so the variance of the observed values is $1 + k$ times the variance of the true values, and the standard deviation is $(1+k)^{1/2}$ times the true standard deviation.

0.0067 to about 0.0042 for a typical variable. Thus the benefit of providing computers to the 407 compact districts in Vietnam for reducing $SD(x)/\sqrt{N}$ is that $SD(x)$ decreases by a factor of $\sqrt{1.0067/1.0042} - 1$, a reduction of 0.1 percent. This gain is very small given its cost of \$244,200 (407 × \$600). Indeed, a similar gain is obtained by increasing the sample by 0.2 percent, that is, adding just 40 households to the approximately 20,000 households in the compact districts. Increasing the sample costs about \$30 per household (personal communication from GSO), so increasing the size of the sample by 40 households would cost \$1,200—just 0.5 percent of the cost of 407 new computers. Even if the methods used in this article detected only one-fourth of the errors avoided by providing each district with a computer, so that providing computers reduces k from 0.0268 to 0.0168, provision of computers would reduce the standard deviation of observed x by only $\sqrt{1.0268/1.0168} - 1$, or 0.5 percent. This reduction in the standard error of the estimated mean of x can be realized by adding 200 households to the sample, at a cost of about \$6,000, or 2.5 percent of the cost of 407 computers.

Two other factors should be considered when assessing the benefits of purchasing computers for each district in Vietnam. First, the VHLSS is implemented every two years, so each computer can be used for two or three surveys before becoming obsolete; the appropriate comparison is thus 407 computers and 400–600 households (generously assuming that the analysis uncovered only one-fourth of the errors in the data that computers would correct). Second, the computers can be used for two other surveys of similar size. Thus the increased precision of the estimated mean of x from buying computers is equivalent to increasing the sample (summing over three surveys) by 1,200–1,800 households. The cost of increasing the sample to 1,800 is about \$54,000, still only 22 percent of the cost of 407 new desktops.

Yet another benefit of providing computers to each district is that they reduce the time VHLSS interviewers spend manually calculating sums and checking the data. According to Truong (2003), providing computers to each district saves about 3.5 hours of interviewer time for each household (for both dispersed and compact districts).

The VHLSS covers 30,000 households. Given 607 districts in Vietnam, the average district contributes about 50 households. Thus purchasing a computer saves about 175 hours of interviewer time per district. Since the VHLSS takes place every two years and the computers last for two or three surveys, the use of computers saves about 500 interviewer hours. Including other surveys the GSO could implement, using those computers could increase time saved to 1,000–1,500 interviewer hours. The GSO pays a typical interviewer in Vietnam about \$0.20 per hour, so each computer purchased can reduce costs by \$200 to \$300. This is less than the \$600 cost of a desktop computer (including training costs), but wages are rising in Vietnam and computer prices are falling; in a few years it may be cost-effective for the GSO to purchase desktop computers for compact districts in Vietnam.

V. Conclusion

Providing computers in Vietnam had little impact on measured poverty and per capita expenditures.[12] Yet regression analysis shows a statistically significant 23 percent reduction in (detectable) errors for many other variables (excluding errors of consistency between food consumption and production data). This reduction is higher for "compact districts" (41 percent), which were given desktop computers. Is this reduction worth the cost? Given the already low rate of errors in the VHLSS data, the answer appears to be "no." Simple calculations (which assume random errors) suggest that standard errors in estimated means of variables of interest can be reduced less expensively by increasing the sample size slightly.

Yet there is another benefit of providing computers to each district. One reason why the error rate of the VHLSS is low is that interviewers currently spend several hours manually checking each questionnaire. Truong's (2003) study suggests that each computer can save 1,000–1,500 hours of interviewer time, implying a savings per computer of $200–$300. A desktop computer costs $600, which is more than the amount saved, but with wages rising and computer prices falling, purchasing new desktop computers may be cost-effective in a few more years.

These results shed light on the merits of moving data-entry computers closer to where interviewers work. Even so, several caveats must be kept in mind. First, the estimated impacts of providing more computers on observed per capita expenditures and poverty are imprecise. An experiment with a larger sample might reveal policy-relevant impacts on these variables.

Second, estimates of the benefits of reducing errors in other variables assume that those errors are unbiased and uncorrelated with those variables' true values. The conclusion that reducing the standard errors of estimated means by providing more computers is much more expensive than doing so by increasing the sample depends on that assumption (although any change in assumptions about measurement errors must be dramatic to overturn this conclusion).

Third, Vietnam's situation may differ from that of other countries. The experience of the first author in many developing countries suggests that household survey data in other countries have far more errors, which implies greater benefits from introducing computers in other countries, especially those with higher interviewer wages. Moreover, the data analyzed here were already "cleaned" by the data-entry program, which may lead to underestimation of the benefits of computerized data entry (although the detectable errors examined were those that the program did not check).

12. While the per capita expenditure variable is arguably the most important variable, the VHLSS measures many other variables; the authors have not (yet) attempted to investigate other variables as intensively.

This last point suggests that more research of this type is needed. Given the cost of purchasing hundreds of computers, randomized experiments similar to this one should be conducted in any country considering decentralized data entry. For Vietnam this study contributed to the decision not to provide computers to each district. Research is also needed to determine whether providing computers reduces bias; this would require reinterviewing households or comparing other data, such as employer records, with household survey data. Finally, when important policy decisions are made using one or two variables from a household survey, such as per capita expenditures, whenever a major change in data collection methods is made a randomized trial should be conducted with a sample size sufficient to detect changes in the distribution of those variables that have substantial policy implications.

Funding

The funding to provide the computers in this experiment was from the World Bank Office in Hanoi, using funds from the U.K. Department for International Development (DFID).

References

Brown, M., and A. Forythe. 1974. "Robust Test for the Inequality of Variances." *Journal of the American Statistical Association* 69:364–67.

Chaudhuri, Shubham, and Martin Ravallion. 1994. "How Well Do Static Indicators Identify the Chronically Poor?" *Journal of Public Economics* 53(3):367–94.

Gibson, John, Jikun Huang, and Scott Rozelle. 2003. "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey." *Review of Income and Wealth* 49(1):53–68.

GSO (Government Statistical Office). 2006. *Vietnam Household Living Standards Survey (VHLSS), 2002 and 2004: Basic Information*. Hanoi.

Grosh, Margaret, and Juan Muoz. 1996. "A Manual for Planning and Implementing the Living Standards Measurement Study Surveys." Living Standards Measurement Study Working Paper 126, World Bank, Washington, D.C.

Johnson, Norman, Samuel Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions*. Vol. 1, 2nd edn. New York: John Wiley.

Ravallion, Martin. 1994. "Poverty Rankings Using Noisy Data on Living Standards." *Economics Letters* 45(4):481–85.

Truong, Thi Kiem Chuyen. 2003. "Assessment Report on Pilot In-Field Data Entry." Department of Geography, University of Social Sciences and Humanities. Ho Chi Minh City, Vietnam.

Forthcoming papers in

THE WORLD BANK ECONOMIC REVIEW

- **Infrastructure and Public Utilities Privatization in Developing Countries**
  Emmanuelle Auriol and Pierre M. Picard
- **HIV pandemic, medical brain drain and economic development in sub-Saharan Africa**
  Alok Bhargava and Frédéric Docquier
- **World Bank Lending and Financial Sector Development**
  Robert Cull and Laurie Effron
- **Does Aid for Education Educate Children? Evidence from Panel Data**
  Axel Dreher, Peter Nunnenkamp, and Rainer Thiele
- **Comprehensive Wealth and Future Consumption: Accounting for Population Growth**
  Susana Ferreira, Kirk Hamilton, and Jeffrey R. Vincent
- **Comparing the Net Benefits of Incentive Based and Command and Control Environmental Regulations: The Case of Santiago, Chile**
  Raúl O'Ryan and José Miguel Sánchez

THE WORLD BANK
1818 H Street, NW
Washington, DC 20433, USA
World Wide Web: http://www.worldbank.org/
E-mail: wber@worldbank.org

ISBN-13: 978-0-19-954697-8
9 780199 546978