Trade Liberalization and Growth: New Evidence

Romain Wacziarg and Karen Horn Welch

A new data set of on openness indicators and trade liberalization dates allows the 1995 Sachs and Warner study on the relationship between trade openness and economic growth to be extended to the 1990s. New evidence on the time paths of economic growth, physical capital investment, and openness around episodes of trade policy liberalization is also presented. Analysis based on the new data set suggests that over the 1950–98 period, countries that liberalized their trade regimes experienced average annual growth rates that were about 1.5 percentage points higher than before liberalization. Postliberalization investment rates rose 1.5–2.0 percentage points, confirming past findings that liberalization fosters growth in part through its effect on physical capital accumulation. Liberalization raised the average trade to GDP ratio by roughly 5 percentage points, suggesting that trade policy liberalization did indeed raise the actual level of openness of liberalizers. However, these average effects mask large differences across countries. JEL codes: F1, F4, O4

Many developing countries have embarked on programs of external economic liberalization in recent decades. In 1960, just 22 percent of all countries, representing just 21 percent of the global population, had open trade policies, in the sense defined by Sachs and Warner (1995). By 2000, some 73 percent of countries, representing 46 percent of the world's population, were open to international trade (figure 1).[1]

1. The main reason for the discrepancy between the share of countries that are open and the share of the world's population living in open countries is that as of 2000, the world's two largest countries, China and India, remained essentially closed. Sachs-Warner (1995) classify India as open as of 1994. The authors revisited this issue and could not confirm their finding. In fact, in terms of both policy indicators and trade volumes, China appears to be twice as open as India. This issue is discussed later in the article and in an appendix to the working version of this paper (Wacziarg and Welch 2003). For an in-depth comparison of the trade regimes of India and China, see Wacziarg (2003).

FIGURE 1. Openness to Trade, 1960–2000 *Note:* Openness is defined according to the Sachs and Warner (1995) criteria. Sample includes 141 countries.
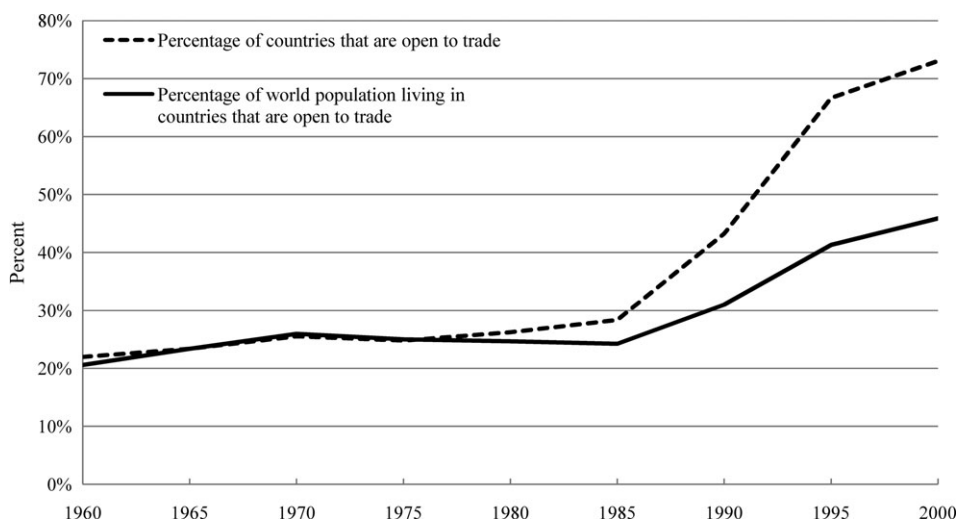
Source: Authors' analysis based on data described in the text.

The effect of this trend toward greater trade policy openness on per capita income growth is the topic of a large body of research. Until recently, a growing academic consensus had emerged that both trade policy openness and higher ratios of trade volumes to gross domestic product (GDP) were positively correlated with growth, even after controlling for a variety of other growth determinants. Attempts to establish a causal link also suggested a positive impact of trade.[2] In a sweeping critical survey of this literature, Rodríguez and Rodrik (2000) argue that these findings are less robust than claimed, because of difficulties in measuring openness, the statistical sensitivity of the specifications, the collinearity of protectionist policies with other bad policies, and other econometric difficulties. Further research on this important topic is called for in view of the doubts their study created about the linkages between trade openness and growth.[3]

Taking over where Rodríguez and Rodrik (2000) left off, the article pursues three goals. The first goal is to update the Sachs-Warner classification by

2. Particularly noteworthy are the contributions of Edwards (1992), Dollar (1992), Ben-David (1993), Sachs and Warner (1995), Ades and Glaeser (1999), and Alesina, Spolaore, and Wacziarg (2000). Among studies trying to establish a causal link running from openness to growth or income levels, see Frankel and Romer (1999), who measure openness by trade volumes, and Wacziarg (2001), who captures openness by using a composite trade policy index.

3. Harrison and Hanson (1999) also criticize the Sachs-Warner classification, in a spirit similar to that of Rodríguez and Rodrik. Their criticisms are revisited in detail later in the article.

presenting a comprehensive cross-country database of trade indicators (tariffs, nontariff barriers, and other measures of trade restrictions) and policy liberalization dates for the 1990s. The second goal is to extend the Sachs-Warner empirical results on outward orientation and growth to the 1990s. The third, and most important, goal is to exploit the timing of liberalization in a within-country setting to identify the changes in growth, investment rates, and openness associated with discrete changes in trade policy.

The availability of almost 50 years of data makes it possible to compare the performance of countries under liberalized and nonliberalized regimes across time. The main empirical analysis presents estimates for the within-country response of per capita income growth, the investment rate, and the ratio of imports plus exports to GDP to trade liberalization, controlling for country and time effects. New evidence is presented on the within-country path of growth in relation to the date of major trade policy changes. Evidence from the large sample is supplemented by a discussion of several developing countries' experiences with trade reform.

The cross-sectional results confirm recent criticisms of the Sachs-Warner findings by showing that these were sensitive to the openness classification used in the 1970–89 period and do not hold for the 1990s. The vast majority of countries in the sample used here are classified as having been open during the 1990s; a simple dichotomous indicator of openness no longer discriminates between slow- and fast-growing countries. The findings here suggest that researchers should exercise caution when using simple dichotomous policy indicators such as the Sachs-Warner dummy variable. However, the dates of trade liberalization—collected by Sachs-Warner from a comprehensive survey of a broad country-specific case literature and updated here to the late 1990s—can be used to estimate the within-country growth and investment effects of trade policy liberalization. In contrast to the cross-sectional findings presented here, the results based on within-country variation suggest that over time the effects of increased policy openness within countries are positive, economically large, and statistically significant.

The article examines a subsample of developing countries for which detailed information was collected on the broader economic and political context of trade reform. It then interprets the large sample results in the context of these country case studies. This effort reveals two lessons. First, the extent to which per capita income growth changed after trade reforms varied widely across countries. While the average effect obtained in the large sample is positive, roughly half of the countries experienced zero or even negative changes in growth following liberalization. Second, generalizations about the factors that may explain these differences are difficult to draw. The institutional environment of countries, the extent of political turmoil, the scope and depth of economic reforms, and the characteristics of concurrent macroeconomic policies all seem to have a role to play, to varying degrees in different countries. While this article paints a picture that is highly favorable to outward-oriented policy

reforms on average, it cautions against one-size-fits-all policies that disregard local circumstances.

The article is organized as follows. Section I presents an updated data set of liberalization dates and policy openness indicators and uses the data to replicate the Sachs-Warner growth regressions. Section II presents within-country evidence on trade liberalization, growth, investment, and trade volumes and discusses the timing of these effects. Section III examines 13 country cases of trade liberalization in order to illustrate the country-specific complexities that underlie the results from the larger sample. The last section provides some concluding remarks.

I. TRADE LIBERALIZATION IN THE 1990s

This section updates the Sachs-Warner classification and results. It also addresses the Rodríguez and Rodrik critique of their study.

The Sachs-Warner Criteria

An update of the Sachs-Warner classification is called for not only because of the problems with their classification of open and closed countries but also because the underlying data—on tariffs, nontariff barriers, exchange rate black market premia, socialist economic systems, and export marketing boards—are of independent interest. This section presents a comprehensive database of these variables for the 1990s. It also presents the results of a painstaking check of the Sachs-Warner classification of openness and updates their data on trade policy openness through 2000.

Sachs-Warner constructed a dummy variable for openness based on five individual dummy variables for specific trade-related policies. A country was classified as closed if it displayed at least one of the following characteristics:

(1) Average tariff rates of 40 percent of more (TAR).
(2) Nontariff barriers covering 40 percent or more of trade (NTB).
(3) A black market exchange rate at least 20 percent lower than the official exchange rate (BMP).
(4) A state monopoly on major exports (XMB).
(5) A socialist economic system (as defined by Kornai 1992) (SOC).

Tariff and nontariff barriers restrict trade directly. A black market premium (BMP) on the exchange rate could have effects equivalent to formal trade restrictions. If, for example, exporters have to purchase foreign inputs using foreign currency obtained on the black market but remit their foreign exchange receipts from exports to the government at the official exchange rate, the BMP acts as a trade restriction. On the basis of Lerner symmetry between import tariffs and export taxes, Sachs-Warner also included the state monopoly on exports criterion as a trade restriction. The socialist regime dummy variable accounts for the trade-limiting aspects of centrally planned economies.

It is important to distinguish the Sachs-Warner dummy variable for openness, which pertains to the 1970s and 1980s, from the Sachs-Warner liberalization dates, which extend from 1950 to 1994 and were compiled independently using a different methodology. While the Sachs-Warner dummy variable was based on the five criteria cited above, the dates of liberalization were obtained from a comprehensive survey of country case studies of liberalization. Where possible, the criteria used to construct the cross-sectional dummy variable for the 1970s and 1980s were used to establish the date of liberalization. Data limitations and lack of consistency in the definitions of the available measures of trade restrictions across time periods, however, prevented Sachs-Warner from using their five criteria to establish the dates of liberalization.[4] The Sachs-Warner methodology was followed as closely as possible in the update presented here.

An Openness Dummy Variable for the 1990s

The sample is based on the 118 countries included in the Sachs-Warner data set.[5] The sample also includes the new data on 23 Eastern European countries and former Soviet republics included in version 6 of the Penn World Tables (Heston, Summers, and Aten 2002). The openness dummy variable *(OPEN90–99)* was based on the five criteria Sachs-Warner use, in order to maintain as much consistency as possible between their data set and the data used here. Data limitations made it impossible to update their dummy variable to the 1990s based on exactly the same data, however.[6]

The main differences between the two data sets include the following:

(1) Because of data availability problems, unweighted tariff data were used here; Sachs-Warner used own import-weighted data. Countries that exceed the *TAR* threshold in the new data set based on unweighted data could conceivably not exceed the threshold based on weighted average data. This is unlikely to be a big problem, however, because the use of unweighted rather weighted tariffs does not result in countries being classified differently in the subsample in which both measures are available.

(2) Nontariff barrier data comparable to those used by Sachs-Warner are hard to obtain. Sachs-Warner used average nontariff barrier data for 1985–88 from the Barro-Lee data set, itself based on data from the United Nations Conference on Trade and Development (UNCTAD). Their data cover only 29 countries for the period 1995–98. Where

---

4. As Sachs-Warner write (p. 24), "Our choice of dating is surely subject to further refinement.... We relied on a wide array of secondary sources, which sometimes contradicted each other." The appendix to their article describes how they compiled their dates of liberalization and identifies the corresponding data sources for each country in their sample. A similar appendix for the updated dates is available in the working paper version of this study (Wacziarg and Welch 2003).

5. Sachs-Warner characterized the openness status of only 111 of these countries.

6. The data sources are detailed in Wacziarg and Welch 2003. The full data set is available in electronic format at www.stanford.edu/~wacziarg/papersum.html. Table 1-A displays the data used to construct the updated openness indicator.

comparable data on nontariff barriers were missing, the countries were classified based only on the other four Sachs-Warner criteria. The limited availability of nontariff barrier data for the 1990s based on a consistent definition required the compilation of an additional nontariff barrier data set, which may be independently useful to researchers. In addition to the 1995–98 average core nontariff barrier data used in the analysis, the data set contains average core nontariff barrier data for 1989–94 and 1999 data for all nontariff barriers.[7]

(3) Sachs-Warner relied on an export marketing index from a World Bank study of African countries (Husain and Faruqee 1994) as the basis for their *XMB* variables and on the Kornai (1992) classification of socialist countries as the basis for their *SOC* dummy variable. In the absence of updated indices from single sources, the same methodology could not be used with the updated data. The *XMB* and *SOC* dummy variables were therefore obtained from a comprehensive review of country case studies. The *XMB* criterion is no longer confined to African countries (as it was in Sachs-Warner), but applies to all countries in the updated data. The definition of an export marketing board was expanded to encompass any form of state monopoly over major exports.[8]

(4) Data on the BMP from Easterly and Sewadeh (2002), the primary source for updating these data, are missing for Belarus, Tajikistan, and Uzbekistan, and only very limited data are available for Armenia, Azerbaijan, Georgia, the Kyrgyz Republic, and Moldova. All are classified as open based on the overall index drawing on limited data. Whenever BMP data were available for former Soviet republics, the data indicate that in 2001 all of these countries except Latvia and Lithuania were closed.

(5) Sachs-Warner deviated in some cases from their self-imposed classification rules. Some adjustments were meant to capture the fact that some countries had undergone changes in trade policy only mid-period, so that a classification based on period averages could be misleading. Other adjustments were made for others' reasons, described in their article. Lacking objective reasons to deviate from stated rules, the updated classification presented here abstains from any such adjustments.

Several features of the new data are worth noting. (The underlying data used to construct the openness status dummy variable for the period 1990–99 are displayed in table A-1.) First, 46 countries that were classified as closed by Sachs-Warner in the 1970–89 period are classified as open in the 1990s.

7. The difference in the definitions reflects the 1999 change in UNCTAD's reporting. Before 1999, UNCTAD collected data on core nontariff barriers, including quotas, licensing, prohibitions, and administered pricing. In 1999, it began reporting all nontariff barriers, which also include technical measures and automatic licensing.

8. Wacziarg and Welch (2003) provide additional details and country-specific sources on export marketing boards and the political transitions from socialism.

Sachs-Warner characterized nine of these countries as closed based on their dates of liberalization. Second, 30 countries were not classified in the Sachs-Warner study, including 23 Eastern European countries and former Soviet republics.[9] Ten of these countries remained closed in the 1990s. Third, of the 111 countries Sachs-Warner classify, 78 were closed and 33 were open in the 1970–89 period. In the 1990s, 32 countries were closed and 79 open. Of the 141 countries classified in the new data set, 42 were closed and 99 open during the 1990s. No country that was classified as open by Sachs-Warner in 1970–89 was classified as closed in the updated data set.

An important and often overlooked drawback of the Sachs-Warner openness dummy variable is that it is based on averages of BMP data over each of two decades (1970–79 and 1980–89), averages of nontariff barriers and tariffs (*TAR*) over the last years of their sample period (1985–88), and end-of-period data for the export marketing board (*XMB*) and socialist (*SOC*) dummy variables. In the new data set, the *XMB* and *SOC* variables are based on their 1999 values rather than beginning-of-period or decade-long data, in order to maintain as much consistency as possible with the Sachs-Warner methodology.[10] Similarly, the nontariff barrier data are available only for 1995–98; decade averages of the tariff data, which are available, are therefore used. As a result, some countries classified as closed could conceivably have become open late in the decade, and some countries classified as open could have been closed over most of the period. Decade dummy variables thus provide only a rough characterization of a country's outward orientation, especially in a decade in which many countries actively engaged in liberalization. A better approach is to rely more on liberalization dates, as is done below.

Trade Liberalization Dates since 1994

In principle, the liberalization date is the date after which all of the Sachs-Warner openness criteria are continuously met (data limitations often imposed reliance on country case studies of trade policy). The choice of liberalization dates was based on primary-source data on annual tariffs, nontariff barriers, and BMPs. A variety of secondary sources was also used, particularly to identify when export marketing boards were abolished and multiparty governance systems replaced Communist Party rule. Because of data limitations, the European Bank for Reconstruction and Development (EBRD 1994) classification and standards of

9. The other seven countries are Cape Verde, Iceland, Lesotho, Liberia, Malta, Panama, and Swaziland. Because of lack of data, Sachs-Warner did not classify these and four other countries (Comoros, Fiji, Seychelles, and Suriname). The new data set did not allow for the determination of the openness status of these four countries in the 1990s.

10. Sachs-Warner's *XMB* indicators are based on data from 1991; the *SOC* indicators are based on data from 1987. Using 1999 data is thus consistent with their approach, however questionable that approach may be. Most countries that abolished export marketing boards in the 1990s did so during the first half of the decade. However, relying on end-of-period *SOC* data means that some Eastern European countries and former Soviet republics are classified as open.

openness were used for several transition economies, just as they are in Sachs-Warner. Table A-2 presents the dates of trade liberalization.[11]

Despite the clear criteria stated above, Sachs-Warner's dates of liberalization could not conform to their five formal criteria for openness, because comparable data were lacking for many time periods. Hence, there is much scope for disagreement with the Sachs-Warner classification, especially in light of new data published since their study. Systematic review of the Sachs-Warner dates since 1990 raised questions about the liberalization status or dates for several countries.[12] Sixteen countries labeled as closed at the end of the Sachs-Warner sample period (1994) liberalized between 1995 and 2001 (table 1).[13] The dates of liberalization cited by Sachs-Warner differ in five countries (Côte d'Ivoire, the Dominican Republic, Mauritania, Niger, and Trinidad and Tobago).

Thirty-five countries remained closed as of 2001, including five that were not classified in the Sachs-Warner study and four (Belarus, Croatia, Estonia, and India) for which the authors disagree with Sachs-Warner's assessment (table 2). Of 141 countries in the sample, 18 liberalized between 1995 and 2001 and 35 remained closed as of 2001.

The Rodríguez and Rodrik Critique

Rodríguez and Rodrik (2000) find that the *BMP* and *XMB* variables played a major role in the classification of countries as open or closed. They state that a dummy variable for openness based on the *BMP* and *XMB* criteria alone leads to the classification of countries as open or closed that is much closer to that generated by OPEN (the Sachs-Warner dummy variable) than one based on the *SOC, TAR*, and *NTB* dummy variables alone. They show that the *BMP* and *XMB* criteria generate a dummy variable that differs from the Sachs-Warner dummy variable in only six cases, while the *TAR, NTB*, and *SOC* criteria used jointly generated a dummy variable that differs from the Sachs-Warner dummy variable in 31 cases. Hence, they argue that the Sachs-Warner dummy variable for 1970–89 largely reflected the *BMP* and *XMB* criteria. Moreover, they argue that the *XMB* criterion affected only the African countries (many of which were classified as closed based on this criterion alone) and therefore amounted to an Africa dummy variable.[14]

---

11. The working paper version of this study (Wacziarg and Welch 2003) provides detailed country summaries of liberalization episodes, along with an explanation of the dates chosen.

12. Wacziarg and Wallack (2004) systematically checked the Sachs-Warner liberalization dates before 1990 in a subset of their sample, uncovering little disagreement.

13. Table 1 also presents data for Cape Verde and Panama, which were not classified in the Sachs-Warner study.

14. Sachs-Warner based the XMB criterion entirely on the Husain and Faruqee's (1994) study of African countries that had been involved in a World Bank or International Monetary Fund structural adjustment program between 1987 and 1991. Rodríguez and Rodrik (2000) noted that Sachs-Warner classify all but one of the Sub-Saharan African countries as closed based on the XMB criterion, which is not applied to any other region. This study gathered and used XMB data for countries other than African ones.

TABLE 1. Liberalization Dates of Countries That Differ from or Were Not Included in Sachs-Warner List

Country	Date of liberalization
Cape Verde	1991
Dominican Republic	1992[a]
Trinidad and Tobago	1992[a]
Côte d'Ivoire	1994[a]
Niger	1994[a]
Armenia	1995
Azerbaijan	1995
Egypt, Arab Rep. of	1995
Mauritania	1995[a]
Mozambique	1995
Tanzania	1995
Bangladesh	1996
Ethiopia	1996
Georgia	1996
Madagascar	1996
Panama	1996
Tajikistan	1996
Venezuela, R.B. de	1996
Burkina Faso	1998
Burundi	1999
Pakistan	2001
Serbia and Montenegro	2001
Sierra Leone	2001

[a]Year differs from that in Sachs and Warner (1995) (see text for explanation).
Source: Authors' analysis based on data described in the text.

TABLE 2. Countries that Remained Closed as of 2001

Algeria	India[a]	Russian Federation
Angola	Iran, Islamic Rep. of	Rwanda
Belarus[a]	Iraq	Senegal
Central African Republic	Kazakhstan	Somalia
Chad	Lesotho[b]	Swaziland[b]
China	Liberia[b]	Syrian Arab Republic
Congo, Dem. Rep. of	Malawi	Togo
Congo, Rep. of	Malta[b]	Turkmenistan
Croatia[a]	Myanmar	Ukraine
Estonia[a]	Nigeria	Uzbekistan
Gabon	Papua New Guinea	Zimbabwe
Haiti		

[a]Disagreement with Sachs and Warner (1995) (see text for explanation).
[b]Not classified in Sachs and Warner (1995).
Source: Authors' analysis based on data described in the text.

To what extent are the updated Sachs-Warner data subject to the Rodríguez and Rodrik critique? *BMP* was the sole criterion on the basis of which 26 of

42 countries were classified as closed in the 1990s; *XMB* was the sole criterion on which nine countries were classified as closed. Three countries were classified as closed based on both the *BMP* and *XMB* criteria, leaving just four countries (Bangladesh, China, India, and Pakistan) classified as closed based on the other three criteria. Bangladesh was classified as closed based on both the *TAR* and *BMP* criteria. China was classified as closed based on the *BMP* and *SOC* criteria. India was classified as closed because of its tariff and nontariff barriers. Pakistan was classified as closed because of tariffs.

The *BMP* and *XMB* criteria generated a dummy variable that differs from the 1990–99 updated Sachs-Warner dummy variable in only two cases, while the *TAR*, *NTB*, and *SOC* criteria used jointly generate a dummy variable that differs from the updated Sachs-Warner dummy variable in 38 cases.[15] The openness status dummy variable for 1990–99 is thus subject to the same criticisms Rodríguez and Rodrik lodged against the Sachs-Warner classification for the 1970–89 openness dummy variable.

The Rodríguez and Rodrik critique is valid in terms of country status based on the *OPEN90–99* dummy variable. It is less valid for the liberalization dates. As most countries were classified as closed based on the *XMB* and *BMP* criteria, not surprisingly, when they open up these variables change. The *XMB* and *BMP* variables determined the year of liberalization in many countries that opened up during the 1990s. The exceptions tend to be Eastern European countries and former Soviet republics, which opened based on the *SOC* criterion (general reforms related to liberalization). The *TAR* criterion was not a decisive factor in assigning a liberalization date for any country; *NTB* was the determining factor only in Panama. However, policy changes that reduced the *BMP* or removed *XMB*s were generally accompanied by changes in the levels of other types of trade barriers, such as tariff and nontariff barriers, that had initial values below the Sachs-Warner thresholds of 40 percent. Hence, liberalization dates do not simply capture changes in the *BMP* and *XMB* variables, but they also reflect broader liberalization. Given that the dates of liberalization in the new data set were cross-checked against a case study literature of outward-oriented reforms in developing countries, it is likely that they reflect important shifts in trade policy.[16]

Updating the Sachs-Warner Results

The Sachs-Warner study attracted considerable attention in part because their estimated effect of the cross-sectional dummy variable for openness in explaining annual growth between 1970 and 1989 was very large (about 2 percentage

15. Among the countries in which the TAR, NTB, and SOC dummy variables and the updated Sachs-Warner dummy variable disagree, 20 are in Africa and 10 are Eastern European countries or former Soviet republics. These countries were classified as closed based on either the XMB criterion or the BMP criterion, or both.

16. Wacziarg and Wallack (2004) show that the Sachs-Warner liberalization dates are good indicators of the timing of major trade policy changes by thoroughly checking these dates against the case study literature of trade liberalization in 25 developing countries.

points). The updated data on trade policy openness make it possible to extend the Sachs-Warner regressions through the late 1990s. As this is not the main focus of this article, these results are reported only briefly.

As a consistency check, the Sachs-Warner regression was first replicated for 1970–89 (column 1 in table 3 replicates column 7 in Sachs-Warner's table 11). The only difference is that the new calculations are based on a newer release of the Penn World Tables data (version 6 instead of version 5). The openness dummy variable for 1970–89 enters highly significantly, with a magnitude of 1.98 percentage points of annual growth. This result is consistent with the results in Sachs-Warner, who find a coefficient of 2.2. In contrast, the updated Sachs-Warner dummy variable enters insignificantly in the same specification for the 1990s (column 2 of table 3).

The cross-sectional effect of openness on growth was estimated by constructing openness indicators based on the dates of liberalization. The openness status for 1980, for example, takes on a value of 1 if a country had liberalized by 1980 and a value of 0 otherwise. Subsequent growth (after 1980) can then be regressed on this variable and other controls. Dummy variables were constructed for each decade (1970, 1980, and 1989) in this fashion. An advantage of this method over the period-specific dummy variables is that the period-specific dummy variables are based partly on information from the end of the period (*TAR*, *NTB*, *XMB*, and *SOC*) and partly on period averages (*BMP*). Constructing openness indicators based on the dates of liberalization instead isolates only the countries that were open at the beginning of a period.

The econometric specification is identical to that in Sachs-Warner; it restricts the time span of each regression to a single decade. The effect of the liberalization status in the 1970s is weaker and smaller than in the 1980s but positive and significant at the 90 percent level. The Sachs-Warner results were likely driven by the strong effect of liberalization on growth in the 1980s (columns 3 and 4 of table 3). This effect is positive but statistically indistinguishable from zero in the 1990s when countries are grouped according to their liberalization status as of 1989. These results suggest that the Sachs-Warner cross-sectional findings are highly sensitive to the decade under consideration and that the updated openness indicator can no longer effectively distinguish fast-growing from slow-growing countries.[17]

II. Within-Country Liberalization Dynamics

This section argues that better use can be made of data on the dates of liberalization. With almost 50 years of data on growth and openness, it is possible to

---

17. Wacziarg and Welch (2003), who conduct many more replications of the initial Sachs-Warner cross-sectional findings, conclude that no matter how the liberalization dummy variable was defined, the results for the 1990s show an insignificant effect of the updated dummy variable on growth. This result is in sharp contrast with the results for the 1970–89 period.

TABLE 3. Replication of Sachs-Warner Cross-sectional Regressions

Variable	(1) Growth 1970–89	(2) Growth 1989–98	(3) Growth 1970–80	(4) Growth 1980–89	(5) Growth 1989–98
Real GDP per capita (t)	−1.5929	−1.150	−1.292	−1.397	−1.261
	(4.89)	(1.95)	(2.83)	(3.84)	(2.13)
Sachs-Warner openness dummy variable(1970–89 or 1990–98 periods)	1.9845	0.136			
	(3.87)	(0.21)			
Openness status based on liberalization dates (t)			1.387	2.574	0.521
			(1.86)	(4.17)	(0.84)
Secondary-school enrollment rate (t)	0.8059	4.689	0.169	1.822	4.872
	(0.68)	(2.43)	(0.10)	(1.40)	(2.52)
Primary-school enrollment rate (t)	1.4003	1.381	2.455	−0.139	1.616
	(1.65)	(0.86)	(2.01)	(0.11)	(0.99)
Government Consumption to GDP ratio ($t, t + X$)	−0.0844	−0.063	−0.005	−0.065	−0.059
	(3.02)	(1.32)	(0.19)	(2.51)	(1.26)
Number of revolutions per year ($t, t + X$)	−0.4359	−0.986	−1.238	−0.211	−1.030
	(0.58)	(1.08)	(1.12)	(0.21)	(1.13)
Number of assassinations per capita per year ($t, t + X$)	0.0296	0.483	0.276	0.188	0.473
	(0.13)	(1.56)	(0.94)	(0.54)	(1.54)
Deviation of the price level of investment (t), as in Sachs-Warner	−0.1709	−0.734	−0.476	0.350	−0.721
	(0.53)	(1.24)	(0.99)	(0.87)	(1.23)
Gross domestic investment/ real GDP ($t, t + X$)	0.0757	0.051	0.076	0.103	0.040
	(2.64)	(1.01)	(2.02)	(2.30)	(0.76)
Extreme political repression (from Sachs-Warner)	−0.6974	0.165	−0.907	−0.780	0.224
	(1.66)	(0.28)	(1.47)	(1.51)	(0.38)
Population density ($t − 10$)	0.0006	0.0009	0.001	0.001	0.001
	(0.90)	(1.40)	(0.60)	(0.87)	(1.49)
Intercept	12.2482	7.752	9.334	10.635	8.288
	(4.87)	(1.81)	(2.84)	(3.86)	(1.92)
Adjusted R^2	0.546	0.211	0.35	0.53	0.32
Number of observations	91	89	99	97	89

Note: Numbers in parentheses are *t*-statistics. The beginning date of each period (1970 in columns 1 and 3, 1980 in column 4, and 1989 in columns 2 and 5) is denoted by *t*. (*t, t + X*) denotes the average computed between dates *t* and *t + X* (*X* = 20 in column 1 and 10 in columns 2–5). The dependent variable is defined as the real annual per capita growth rate of GDP in the relevant period.

Source: Authors' analysis based on data described in the text. Growth, income, and investment data are from Heston, Summers and Aten (2002).

assess the within-country effects of discrete changes in trade policy openness.[18] This section compares the means of economic growth and other variables of interest, such as physical capital investment rates and trade volumes, under liberalized and nonliberalized regimes.

Liberalization and Growth

Fixed-effects regressions of growth on a binary liberalization indicator, defined by the dates of liberalization, were run to assess the within-country effect of growth on liberalization. The regressions amount to difference regressions in growth or difference-in-difference regressions in log income:

$$(1) \qquad\qquad \log y_{it} - \log y_{it-1} = \alpha_i + \beta LIB_{it} + \varepsilon_{it}$$

where y_{it} is per capita income in country i at time t and $LIB_{it} = 1$ if t is greater than the year of liberalization and no reversals of the trade policy reforms have occurred, and 0 otherwise. The sample was not restricted to countries that underwent reforms. The residual term is modeled as $\varepsilon_{it} = \nu_i + \eta_t + \mu_{it}$ and in all regressions, the ν_i terms are treated as country fixed effects and η_t terms as fixed effects.

Over the sample period 1950–98, 31.7 percent of country-year observations occur in a liberalized regime ($LIB_{it} = 1$) (table 4). The conditional mean of annual growth of per capita GDP given that a country is liberalized is 2.71 percent, while the mean is 1.18 percent in a nonliberalized regime, a difference of 1.53 percentage points of annual growth. These simple conditional means are based on both cross-sectional and within-country variation.

Panel (1) of table 5 displays country and time fixed-effects regressions of growth on the liberalization indicator, in order to isolate within-country variation. The regression for 1950–98 indicates a within-country difference in growth between a liberalized and a nonliberalized regime of 1.42 percentage points (column 1). This coefficient is estimated with a high level of statistical precision (the *t*-statistic exceeds 5).[19] The estimated within-country

18. Sachs and Warner provide some within-country evidence on liberalization and growth for a sample of 37 reformers, presenting estimates for one fixed-effects regression of growth on dummy variables for three time periods around liberalization episodes. They show that average growth was depressed by 0.88 percentage points in the three years before liberalization, rose 1.09 percentage points a year in the three years following liberalization, and rose 1.33 percentage points a year thereafter relative to growth in the three years before liberalization. These limited results are of the same order of magnitude as the more detailed research presented here, which investigates the robustness of these estimates, extends them in time (the sample period spans 1950–98 rather than 1966–93) and space (the sample includes up to 133 countries rather than 37), and presents new evidence on investment and openness.

19. This effect was estimated allowing for first-order autocorrelation of the residuals, using the Baltagi-Wu fixed-effects method. The coefficient on liberalization was 1.32, with a *t*-statistic of 4.14, in line with the fixed-effects results reported here. The simpler fixed-effects estimates, with *t*-statistics based on robust standard errors, are reported here because of concerns over the small T properties of the Baltagi-Wu estimator, particularly when the sample is restricted to specific decades.

TABLE 4. Summary Statistics for Variables Used in Fixed-Effects Regressions

Variable	Number of observations	Mean	Standard deviation	Minimum	Maximum
Liberalization	7,191	0.317	0.465	0.0	1.0
Investment rate (percent)	5,078	15.291	9.128	−3.590	52.880
Openness ratio	5,078	60.505	42.880	3.110	473.860
Growth (annual percent)	4,936	1.784	6.153	−48.732	43.754
Per capita GDP (purchasing power parity US$)	5,072	5,739.380	5,826.636	276.000	39,129.000

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

TABLE 5. Fixed-Effects Regressions of Growth, Investment, and Openness on Liberalization Status, 1950–98

Item	(1) 1950–98	(2) 1950–70	(3) 1970–90	(4) 1990–98
Dependent variable: Growth				
Liberalization	1.417	0.611	1.787	2.547
	(5.05)	(1.29)	(3.11)	(2.39)
Number of observations	4,936	1,728	2,312	1,116
Number of countries	133	108	112	133
Adjusted R^2	0.05	0.03	0.04	0.04
Dependent variable: Investment rate				
Liberalization	1.937	2.545	1.237	0.762
	(9.06)	(7.57)	(2.91)	(2.16)
Number of observations	5,078	1,844	2,321	1,140
Number of countries	136	110	117	136
Adjusted R^2	0.04	0.10	0.11	0.02
Dependent variable: Openness				
Liberalization	5.531	2.302	4.097	−1.803
	(7.42)	(1.89)	(3.74)	(0.83)
Number of observations	5,078	1,844	2,321	1,140
Number of countries	136	110	117	136
Adjusted R^2	0.22	0.02	0.14	0.08

Note: Numbers in parentheses are robust *t*-statistics. Regressions are based on the specifications in equations (1)–(3). All regressions include time and country fixed-effects (estimates not reported).

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

effect increases over time, reaching its maximum in the 1990s (column 2–4). These results stand in sharp contrast to the cross-sectional results: countries that liberalized in the 1990s experienced a larger postliberalization increase in growth than countries that liberalized in any other decade. Indeed, the estimated difference in growth in the 1990s is roughly 2.55 percentage points.

Liberalization and Investment

The empirical literature on trade and growth suggests that the effects of liberalization on economic growth are mediated largely by the rate of physical capital investment. Several researchers, including Levine and Renelt (1992), Baldwin and Seghezza (1996), and Wacziarg (2001), suggest that the investment rate is an important channel linking trade and growth. This finding is based largely on cross-country findings. Fixed-effects regressions of investment rates on the liberalization indicator were run in order to investigate this issue in a within-country context:

$$(2) \qquad \frac{I_{it}}{Y_{it}} = \eta_i + \phi LIB_{it} + \omega_{it}$$

where I_{it} is physical capital investment and Y_{it} is GDP in country i at time t, and ω_{it} captures country and year effects.

Panel (2) of table 5 reports the estimates of such regressions. The within-country evidence confirms past cross-country findings. For the period 1950–98, countries with liberalized regimes experienced average rates of physical capital investment that were 1.94 percentage points higher than those of countries with nonliberalized regimes. This represents 20 percent of this variable's standard deviation in the pooled sample. The effect is largest in the initial period of the sample (1950–70).

Fixed-effects regressions of growth on the investment rate were run in order to get a rough notion of how much of the effect of trade policy openness on growth can be attributed to the investment channel. The coefficient on investment in the baseline 1950–98 regression was 0.15 percentage points, with a *t*-statistic of 8.05.[20] The effect of liberalization on investment in the corresponding regression was 1.94 percentage points. Multiplying the two yields an estimate of the effect of liberalization on growth through investment of roughly 0.29 percentage points, about 21 percent of the total effect of liberalization on growth. The analysis provides suggestive evidence that investment constitutes an important channel through which trade-centered liberalization affects growth within countries.

Liberalization and Openness

Is trade policy liberalization followed by a break in the volume of trade, as measured by the ratio of imports plus exports to GDP? If this is the case, it suggests that liberalization did increase the level of openness of the economy. Determining this effect is important, because announced reforms may be poorly implemented or counteracted by alternative trade barriers. If trade liberalization is associated

---

20. The full results are presented in the working paper version of this study (Wacziarg and Welch 2003).

with increases in trade volumes, one could be more confident that it actually raised the level of exposure of the reforming country to the world economy.[21]

This issue is examined by running the following regression:

(3)
$$\frac{X_{it} + M_{it}}{Y_{it}} = v_i + \delta LIB_{it} + w_{it}$$

where X_{it} denotes exports and M_{it} denotes imports. The results suggest that liberalization raises openness by 5.53 percentage points of GDP for the full sample period (Panel (3) of table 5). This effect is indistinguishable from zero in the 1990–98 time period, however, perhaps because more time is needed to observe the effects of recent liberalizations on trade volumes. In most periods, however, trade liberalization is associated with sustained and large increases in the effective level of exposure of the typical reforming country to the world economy.

Timing of Effects

The simple average difference between growth in nonliberalized and liberalized regimes may mask important timing issues. It provides no information on how soon the effects occur or whether they cease to be felt a few years after reform. This subsection examines the time path of growth, investment, and openness for an average country before and after liberalization.

Average annual growth rates, investment rates, and openness ratios are displayed in figures 2 through 4 for 20 years before and 20 years after liberalization in a sample of 81 countries that underwent permanent liberalizations (that is, liberalizations that were not reversed as of 2000). As several countries had varying numbers of years of data before and after their liberalization, the average at each point in time is based on different samples of countries.[22]

Several observations can be made about the figures. First, despite not controlling for any fixed effects, the increase in growth following liberalization is remarkably similar to that shown in table 5: growth before trade-centered reforms averages 1.5 percent and rises to roughly 3 percent postreform (figure 2). Second, there does not seem to be a strong time pattern: the effects appear to be immediate and do not die out after a few years. Third, the few years immediately preceding liberalization are low-growth years: reforms are often preceded by downturns or crises.

The investment rate seems to take off during the 10 years following liberalization and remain high thereafter (figure 3). The plotted effect seems larger than that uncovered in the fixed-effects regressions. Openness follows a more or less

21. Even absent effects on actual openness, liberalization could still have effects on growth and investment, through pro-competitive effects or technological transfers, for example.

22. The figures did not look different when the sample was restricted to countries with continuously available data. The availability of data forced a reduction in the time span to eight years before and after liberalizations and in the country coverage to 39 countries. These figures are available in the working paper version of this study (Wacziarg and Welch 2003).

FIGURE 2. Sample Means for Growth before and after Liberalization
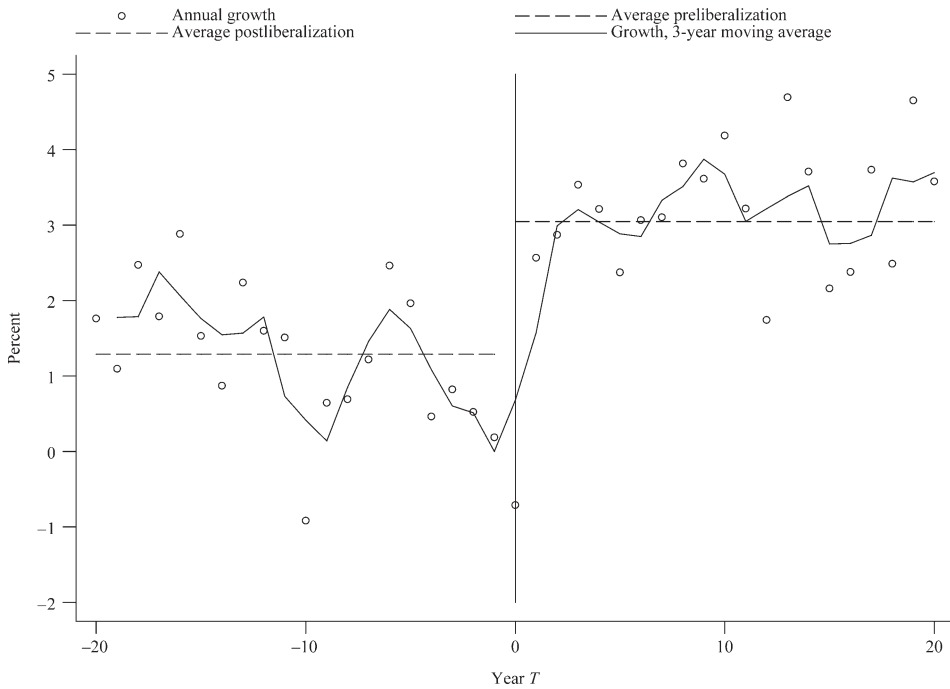
Source: Authors' analysis based on data described in the text.

linear upward trend, without an apparent break at the date of liberalization (figure 4). More formal tests based on fixed effects did reveal an effect attributable to liberalization, even after controlling for time fixed effects, however.

Dummy variables for four (nonoverlapping) periods surrounding the reforms were defined in order to further examine the timing of the growth, investment, and openness responses to liberalization. Fixed-effects regressions were then run on growth, investment, and openness. The specification is as follows:

$$(4) \qquad \log y_{it} - \log y_{it-1} = \alpha_i + \beta_1 D_{1it} + \beta_2 D_{2it} + \beta_3 D_{3it} + \beta_4 D_{4it} + \varepsilon_{it}$$

where $D_{1it} = 1$ if $T - 3 \leq t \leq T - 1$ and zero otherwise; $D_{2it} = 1$ if $T \leq t \leq T + 2$; $D_{3it} = 1$ if $T + 3 \leq t \leq T + 6$, and $D_{4it} = 1$ if $t > T + 6$; and T denotes the date of liberalization. The coefficients on these dummy variables capture the average difference in growth between these years and the period preceding three years before liberalization (the baseline period). The corresponding specifications for the investment rate and openness ratio were also run (table 6).[23]

23. Countries that experienced policy reversals or multiple liberalizations, for which definitions of the dummy variables are not straightforward, had to be dropped. Dropping these variables reduced the size of the sample for the growth regression from 133 to 118 countries.

FIGURE 3. Sample Means for Investment before and after Liberalization
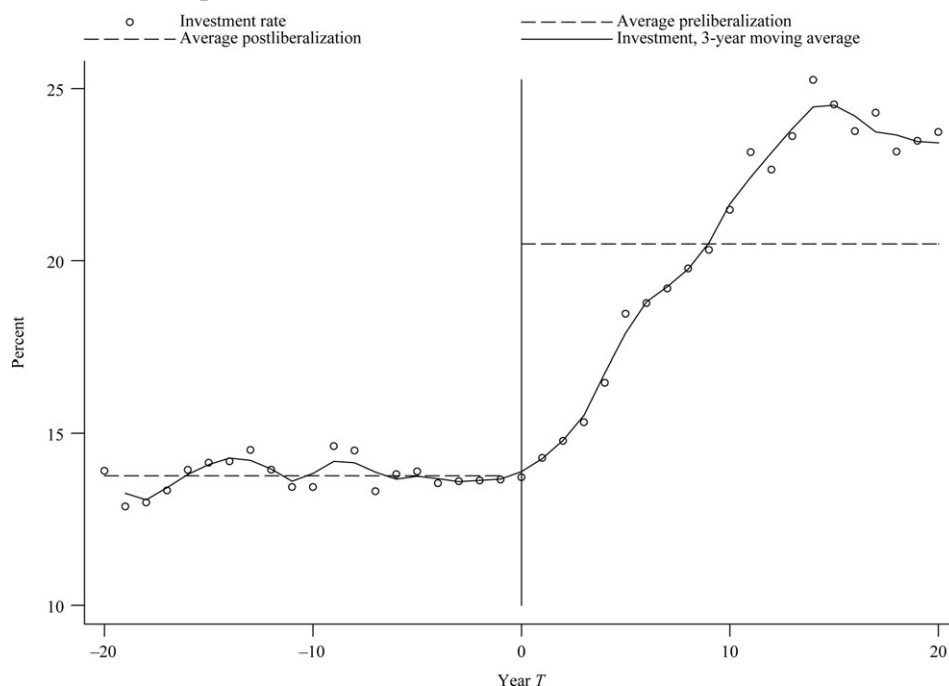
Source: Authors' analysis based on data described in the text.

The results are consistent with the observations made about figures 2–4. Countries that liberalize often do so following periods of economic turmoil: growth is depressed by 0.55 percentage points in the three years before liberalization relative to the preceding years. Tornell (1998) shows that 60 percent of episodes of economic reform, including trade reform, occur in the aftermath of a domestic political or economic crisis. Measuring growth differences relative to "early prereform" outcomes prevents falsely attributing to reforms growth differences that stem from depressed economic circumstances in the years immediately preceding the reforms. In the three years following liberalization, growth rises slightly (by 0.30 percentage points), but the effect is statistically indistinguishable from zero. Sustained growth differences become apparent three years after reform, with annual increases in growth of 1.44 points in period $T + 3$ to $T + 6$ and of 1.0 percentage point after that relative to the baseline period. The typical timing pattern revealed by these regressions shows growth to be slightly depressed before liberalization and to increase 1.0–1.5 percentage points three years after reforms. A similar pattern applies to investment and openness. These estimates reflect sample averages and may mask interesting country-specific differences, as discussed below.

FIGURE 4. Sample Means for Openness before and after Liberalization
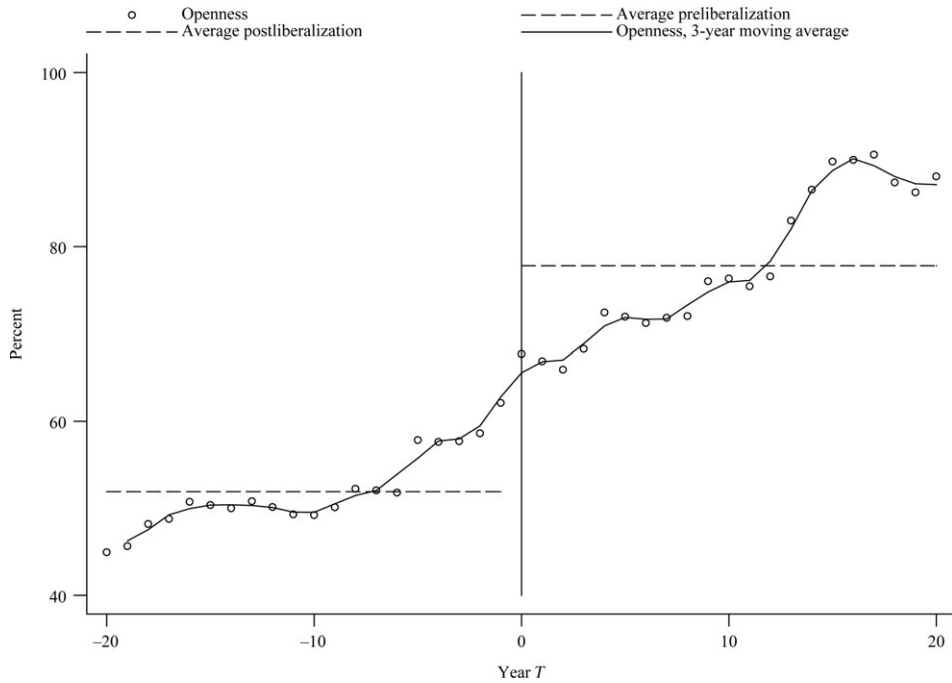
Source: Authors' analysis based on data described in the text.

TABLE 6. Fixed-Effect Regressions: Timing of the Effects of Liberalization on Growth, Investment, and Openness

Item	(1) Growth	(2) Investment	(3) Openness
D_1	−0.555	−1.040	−1.979
	(1.14)	(2.88)	(1.32)
D_2	0.300	−0.160	0.795
	(0.61)	(0.41)	(0.63)
D_3	1.438	1.197	3.606
	(3.27)	(2.98)	(2.21)
D_4	1.015	2.129	13.371
	(2.30)	(5.47)	(9.17)
Number of observations	4,230	4,357	4,357
Number of countries	118	121	121
Adjusted R^2	0.04	0.08	0.26

Note: Number in parentheses are robust-statistics. Regressions are based on the specification in equation (4). All regressions include time and country fixed-effects (estimates not reported). Definition of dummy variables, where T represents the date of liberalization, is as follows: $D_1 = 1$ if $T − 3 \leq t \leq T − 1$ and zero otherwise. $D_2 = 1$ if $T \leq t \leq T + 2$ and zero otherwise. $D_3 = 1$ if $T + 3 \leq t \leq T + 6$ and zero otherwise. $D_4 = 1$ if $t > T + 6$ and zero otherwise.

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

Concurrent Policies

It is difficult to attribute differences in growth purely to trade liberalization. Countries carrying out trade reforms often simultaneously adopt policies favoring domestic deregulation, privatization, and other microeconomic reforms and macroeconomic adjustments, making it difficult to interpret the coefficient on liberalization in a within-country growth regression as the total effect of trade liberalization per se.[24] A more realistic interpretation of these estimates is that they capture the impact of trade-centered reforms more broadly. In what follows, we describe our efforts to address this important concern.

SCOPE OF REFORMS. The working paper version of this study (Wacziarg and Welch 2003) distinguishes countries that carried out overall reforms from those that carried out external sector reforms in relative isolation from other domestic reforms. Wacziarg and Wallack (2004) examine 22 episodes of trade liberalization, most of them in developing countries in the 1980s. Fourteen of these episodes were accompanied by market-oriented domestic reforms; eight occurred in relative isolation from major shifts in domestic policy. The distinction between pure trade reforms and overall reforms was based largely on whether the countries implemented a substantial program of privatization and deregulation at the same time as trade reforms.

Isolating the sample of countries that were part of the Wacziarg and Wallack (2004) study and examining whether the within-country effects of liberalization on growth differed between trade reformers and overall reformers reveal several noteworthy findings. First, even though the sample was restricted to 22 countries, the estimates were remarkably similar to those obtained for the full sample of 133 countries. Second, the estimates of the impact of trade liberalization in countries that carried out trade reforms in isolation were similar to the corresponding estimates for countries that also reformed their domestic sectors, despite the crude nature of the distinction between overall reformers and pure trade reformers. While the interpretation of these suggestive results requires caution, a plausible conclusion is that the effect of trade-centered reforms is in large part attributable to an external reform component. This issue is further addressed below in the context of individual country experiences.

OTHER EXTERNAL REFORMS. Trade reforms are sometimes associated with other types of external reforms, such as capital market liberalization. To the extent such reforms are adopted simultaneously, estimates may capture the impact of these financial reforms rather than trade reforms. This argument is frequently invoked to criticize the type of estimates presented above.

---

24. An analogous point is often made in a cross-country context. Rodríguez and Rodrik (2000) and other observers suggest that "bad" government policies tend to go together, making it difficult to disentangle the effects of protectionist trade policy from those of poor macroeconomic management, poor governance, or poor institutions in general.

This issue is investigated by looking at data on the timing of financial reforms. Bekaert, Harvey, and Lunblad (2001) examine the impact of capital market liberalization on economic growth in a panel context, using both cross-sectional and within-country time variation. Using data from Bekaert and Harvey (2000) on the dates of official regulatory reforms pertaining to financial markets, they find robust positive effects of financial liberalization.[25] Their dates are compared with the dates of trade liberalization in the data set used here.

Bekaert and Harvey (2000) characterize the date of official financial liberalization for 40 of the 106 countries in the sample that had liberalized by 2001.[26] Of these, only two (Brazil and Turkey) have exactly the same year of official financial regulatory reform and trade liberalization. Only nine countries implemented financial sector reforms within three years before and after the date of trade liberalization, and just 17 did so within five years before and after. Many countries that enacted trade reforms never enacted financial liberalization, so the numbers cited above overstate the extent of coincidence between financial and trade liberalization dates. There is thus little evidence that trade reform and financial market liberalization occur concurrently and that the estimates may confound the effects of the two types of reform.

III. Country Case Studies

The econometric results presented above summarize the effect of trade liberalization on growth and other variables for a sample of very diverse countries. Fixed-effects regressions allow all time-invariant country characteristics to be controlled for. The estimated coefficients on liberalization are not country-specific, however; they represent average responses. The reaction of individual countries to reforms is likely to vary, especially as the depth and scope of reforms differed across countries.

Much can be learned from the considerable heterogeneity in the response of growth to trade reform. This section examines specific cases of reform in countries representative of the broader sample for which enough data on growth, investment, and openness are available before and after reforms. The goal is to get a sense of the subtleties of reform in specific cases and to illustrate the economic mechanisms that give rise to the average estimated effects. The time paths of growth, investment, and openness are first examined for a subsample of 24 developing countries for which data are available for at least eight years before and after liberalization. A more detailed discussion then focuses on 13 of these countries.

25. Henry (1999, 2000) uses data on economic and political reforms for a smaller set of 18 developing countries.

26. Details of the comparison between the Bekaert and Harvey (2000) dates and the dates presented here are available on request.

The average difference in growth, investment rates, and openness ratios between the pre- and postliberalization periods is shown for 24 countries (table 7). The countries were chosen from the sample of 39 countries for which at least eight years of data are available on either side of the date of liberalization, restricting the sample to emerging markets, the main focus of this study. The data reveal positive growth differences in 13 of the 24 countries and negative differences in six of them; the remaining five countries exhibit an effect close to zero. Postliberalization growth effects appear large in Mauritius, Indonesia, Uruguay, Republic of Korea, Chile, Taiwan (China), and Uganda. Among countries that experienced positive differences, the magnitude of the growth increase ranged from 0.83 percentage points of per capita income growth in Poland to 3.62 points in Mauritius. The range of growth decline was of a similar magnitude.

Before and after comparisons of investment rates and openness also reveal large variations across countries. The postliberalization surge in investment rates was particularly strong in the Republic of Korea, Taiwan (China), Indonesia, Jordan, and Guinea-Bissau. About half of the 24 countries exhibited zero or negative differences in investment rates.

TABLE 7. Mean Growth, Investment, and Openness Changes in 24 Countries

Country	Growth difference	Investment difference	Openness difference	Year of liberalization	Sample period
Mauritius	3.62	0.34	35.90	1968	1951–98
Indonesia	3.32	9.80	25.96	1970	1961–98
Uruguay	3.08	−1.01	11.22	1990	1951–98
Korea, Rep. of	3.02	18.44	43.40	1968	1954–98
Chile	2.80	−1.12	26.33	1976	1952–98
Taiwan	2.29	9.91	55.77	1963	1952–98
Uganda	2.24	1.63	−6.60	1988	1951–98
Ghana	1.99	−3.91	9.13	1985	1956–98
Guinea	1.85	−2.74	7.28	1986	1960–98
Guyana	1.80	−7.49	84.49	1988	1951–98
Benin	1.74	1.64	8.72	1990	1960–98
Mali	1.19	0.86	15.68	1988	1961–98
Poland	0.83	−4.30	3.35	1990	1971–98
Paraguay	0.42	2.01	49.71	1989	1952–98
Cyprus	0.34	−4.05	29.13	1960	1951–96
Colombia	0.18	0.48	5.91	1986	1951–98
Tunisia	−0.30	−5.58	31.94	1989	1962–98
Philippines	−0.40	1.03	39.54	1988	1951–98
Israel	−0.96	−6.10	21.42	1985	1951–98
Botswana	−1.99	3.98	22.27	1979	1961–98
Mexico	−2.16	−4.59	17.56	1986	1951–98
Hungary	−2.41	−1.19	−4.17	1990	1971–98
Guinea-Bissau	−2.95	5.59	9.89	1987	1961–98
Jordan	−4.28	5.75	40.61	1965	1955–98

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

Closer examination of postliberalization changes in growth, investment, and openness for a restricted sample of developing countries thus reveals considerable heterogeneity in their experiences with reform. The following case studies develop hypotheses that could account for these differences.

From the sample of 24 developing countries for which there are at least eight years of data on either side of liberalization, a subsample of 13 countries was selected to study in greater detail. A set of countries was chosen that was small enough to allow their preexisting conditions, overall policy environment, and macroeconomic circumstances to be examined while maintaining a geographically diverse sample reflecting the range of country-specific growth effects identified above. The goal was to uncover patterns that could explain cross-country differences in individual countries' responses to liberalization and suggest directions for future research.

The subsample was selected to include a geographically diverse set of countries that experienced growth effects of liberalization in roughly the same proportions as the 24 countries discussed above. It includes 13 countries, seven of which experienced higher mean growth rates following liberalization (Indonesia, Republic of Korea, Chile, Taiwan (China), Uganda, Ghana, and Poland). The growth difference was negative in four countries (Israel, Botswana, Mexico, and Hungary). In two countries, Colombia and the Philippines, liberalization was associated with roughly zero difference in their mean growth rates. Table A-3 describes all countries' concurrent reforms, macroeconomic environment, and political context.

Examination of these case studies suggests that the packaging and timing of reforms are important factors in explaining differences in postliberalization growth patterns. Countries that followed through by deepening trade reforms over time did better than countries that did not. Neither active governmental disengagement from industrial policy nor broad-based reforms were necessary conditions for success. Countries that counteracted short-lived programs of external liberalization with domestic interventions and countries that adopted tight macroeconomic policies, faced unfavorable terms of trade shocks, or suffered from political instability did not perform as well as other countries.

Sustained Reforms

In the majority of countries that experienced higher growth following liberalization, trade reforms were not strictly limited to the period of liberalization; these countries continued to deepen trade reform after liberalization. Chile, for example, which liberalized in 1976, recovered from the Latin American debt crisis and continued to grow during the late 1980s. During this period, it decreased tariffs and implemented several bilateral free trade agreements. Both Korea and Taiwan (China), which liberalized in the 1960s, continued to lower tariffs and remove nontariff barriers, particularly during the mid-1980s and

1990s. Indonesia sustained the initial reforms of 1970 with reductions in export duties in 1976 and additional trade-centered liberalization throughout the 1980s. In Uganda, the 1988 liberalization was followed by a second wave of external reforms in 1993–94.

Scope of Reforms

Whether trade reforms were part of a package of other domestic reforms or occurred in relative isolation does not seem to help predict the effect on growth. Among countries that implemented broad-based reforms, and in which postliberalization growth increased, Chile and Poland stand out as prototypical success stories of reform. Both implemented broad-based domestic reforms, of which trade liberalization was only a part. In Colombia, Hungary, and Mexico, which Wacziarg and Wallack (2004) classify as broad-based refor-mers, average growth following liberalization actually fell. Political instability is probably at the heart of Colombia's lack of increased growth. In Hungary, the decline may have occurred because the domestic portion of the reform program (banking sector reforms, privatizations) was in large measure delayed until 1995. To the extent that external and domestic reforms are complemen-tary, the full effects for Hungary may not be apparent in the growth data, which extend only to 1998.

The case of Mexico is more complex. The privatization program began before trade liberalization, in 1984, with the sale of small- and medium-sized businesses, and continued after 1986, with the sale of larger enterprises, such as the national telephone company, parts of the banking industry, and the national airline. While Mexico maintained large government oligopolies that prevented broad industrial restructuring and resource reallocation, one can hardly argue that its entry into the General Agreement on Tariffs and Trade (GATT) in 1986 and the concurrent reduction in external barriers occurred in isolation from other domestic reforms.

The flip side of this coin is a country like Ghana, which, according to Wacziarg and Wallack (2004), implemented trade reforms in relative isolation (privatiza-tion, for instance, did not begin until the early to mid-1990s). It experienced a 2 percentage point increase in mean growth after the 1985 liberalization.

Other interesting cases are the success stories of Southeast Asia, where many economies, including Korea and Taiwan (China), implemented policies aimed at increasing foreign direct investment (FDI) at the same time or after external liberalization. Indonesia, Korea, and Taiwan (China) pursued growth strategies with widespread government involvement in the economy. In Indonesia, government involvement increased during the 1970s, after external liberali-zation began. Both Korea and Taiwan (China) adopted activist industrial pol-icies, with the government involved in "picking winners." That the growth performance of these countries was unprecedented the 1998 Asian crisis shows that government disengagement from the economy is not a necessary condition for successful trade reforms. What all these countries shared was an

outward-oriented development model in which increasing exports was a central pillar of the growth strategy.

One cannot point to the breadth of reform as an unambiguous criterion explaining differences in the growth response to liberalization. The picture that emerges is far from simple. The set of economies that experienced higher growth following liberalization includes both those that maintained heavy government involvement in the economy (Indonesia, Korea, and Taiwan [China]) and those that actively reduced the role of government (Chile and Poland). The set of countries that experienced negative or zero growth differentials after liberalization includes Colombia, Hungary, and Mexico, countries that actively disengaged the government from domestic economic activity at the time of trade reforms.

Counteractive Policies

Some of the 13 countries in the sample implemented policies that actively counteracted the effects of trade reform and as a result did not experience increases in growth rates.[27] In Israel, social pacts based on broad coalitions of labor, government, and industry set the patterns for prices, wages, and the exchange rate in ways that mitigated the effects of trade openness on domestic producers. In the Philippines, trade liberalization was accompanied by a large increase in the share of state-owned enterprises in the economy, including a doubling of the share in GDP of financial transfers from the government to state-owned enterprises between 1987 and 1989. Such interventions, designed partly to protect domestic producers in the face of increased import competition, may have precluded the realization of gains from trade.

Macroeconomic Factors

Countries that did not experience growth increases after liberalization often suffered from mitigating circumstances, associated in particular with restrictive macroeconomic policies or terms of trade shocks. In Hungary and Mexico, two countries in which growth fell following liberalization, trade reform was followed by tight monetary policies involving high interest rates, which depressed growth. In Mexico, currency overvaluation undid the effects of trade liberalization in the late 1980s and early 1990s.

In Botswana, terms of trade considerations account for the absence of a postliberalization growth surge. Volatility on world diamond markets increased shortly after Botswana implemented trade reforms in 1979. The weak diamond marked caused a recession in 1981–82 that resulted in a postliberalization growth rate that was about 2 percentage points lower than the preliberalization rate. Thus, terms of trade considerations are essential in accounting for the absence of a postliberalization growth surge in Botswana.

27. Wacziarg and Wallack (2004) discuss some of these cases in greater detail.

Political Instability

Several countries suffered from severe forms of political instability, preventing realization of the gains from trade liberalization. A prime example is Colombia, where instability persisted throughout the 1990s. Other examples include Israel and the Philippines. In contrast, economies that seem to have experienced higher growth following reform also witnessed periods of relative political stability. Taiwan (China) is a case in point, as are Chile, Indonesia, and Korea, where liberalization coincided roughly with the rise to power of authoritarian regimes, resulting in a degree of lasting political stability following periods of political unrest.

IV. CONCLUSION

This article presents an updated data set of trade policy indicators and liberalization dates. It revisits the evidence on the cross-country effects of Sachs-Warner's simple dichotomous indicator of outward orientation on economic growth, confirming the pitfalls of this indicator first identified by Rodríguez and Rodrik (2000). It shows that the Sachs-Warner dichotomous indicator effectively separates fast-growing from slow-growing countries in the 1980s and to a lesser extent in the 1970s, but fails to do so in the 1990s. Simple dichotomous indicators of outward orientation are too crude to capture the complexities of trade policy.

Instead, liberalization dates that capture episodes of discrete shifts in trade policy can be useful for estimating within-country growth responses The Sachs-Warner dates of liberalization were painstakingly checked and updated, based on quantitative data and a thorough review of country-specific case studies of reform. The new and robust evidence indicates that these dates of liberalization mark breaks in growth, investment, and openness within countries. Over the 1950–98 period, countries that liberalized their trade regimes experienced average annual growth rates that were about 1.5 percentage points higher than before liberalization. Postliberalization investment rates rose 1.5–2.0 percentage points, confirming past findings that liberalization fosters growth in part through its effect on physical capital accumulation. Liberalization raised the average trade to GDP ratio by roughly 5 percentage points, after controlling for year effects, suggesting that trade policy liberalization did indeed raise the actual level of openness of liberalizers. Trade-centered reforms thus have significant effects on economic growth within countries.

These within-country estimates represent the average effect of liberalization on growth, investment, and openness; they mask differences in the individual responses of countries to trade liberalization. Restricting the sample to 13 countries sheds light on the sources of these differences. Countries that experienced positive effects tended to deepen trade reforms. But active industrial policies, such as those implemented in Southeast Asia, did not preclude growth

gains from trade liberalization, and broad-based reforms appear to be neither a necessary nor a sufficient condition for reaping these gains. Countries that experienced negative or no effects on growth tended to have suffered from political instability, adopted contractionary macroeconomic policies in the aftermath of reforms, or undertaken efforts to counteract trade reform by shielding domestic sectors from necessary adjustments. Future research should seek to clarify the factors accounting for heterogeneity in the growth effects of trade reform.

APPENDIX

TABLE A-1. Trade Policy Variables for Economies in Sample, 1990s

Economy	OPEN90–99 (1 = open)[a]	Average tariff, 1990–99 (percent)[b]	Core nontariff barrier coverage rate (percent), 1995–98[c]	Average BMP, 1990–99 (percent)[d]	Export Marketing Board (1 = country has exporting board)[e]	Socialist (1 = country is socialist)[e]
Albania	1	15.90	—	7.53	0	0
Algeria	0	23.97	—	177.91	0	0
Angola	0	—	—	23.62	0	0
Argentina	1	12.54	2.1	9.30	0	0
Armenia	1	—	—	0	0	0
Australia	1	7.91	—	0	0	0
Austria	1	6.91	—	0	0	0
Azerbaijan	1	—	—	0	0	0
Bangladesh	0	43.70	—	83.27	0	0
Barbados	1	15.58	—	2.31	0	0
Belarus	0	12.63	—	—	1	0
Belgium	1	6.91	—	0	0	0
Benin	1	28.61	1.0	1.93	0	0
Bolivia	1	10.34	—	1.49	0	0
Botswana	1	20.55	—	7.82	0	0
Brazil	1	17.32	21.6	13.76	0	0
Bulgaria	1	17.37	—	7.44	0	0
Burkina Faso	1	29.13	—	1.98	0	0
Burundi	0	7.40	—	29.55	0	0
Cameroon	1	18.43	—	1.98	0	0
Canada	1	6.81	—	0	0	0
Cape Verde	1	22.05	—	0	0	0
Central African Republic	0	12.80	—	1.55	1	0
Chad	0	15.80	—	1.98	1	0
Chile	1	11.33	5.2	9.84	0	0
China	0	31.06	—	35.89	0	1
Colombia	1	14.30	10.3	8.87	0	0
Congo, Dem. Rep. of	0	25.47	—	34.67	1	0
Congo, Rep. of	0	17.97	—	1.98	1	0
Costa Rica	1	10.60	6.20	5.37	0	0
Côte d'Ivoire	1	22.00	30.90	1.98	0	0
Croatia	0	—	—	37.76	0	0
Cyprus	1	10.64	21.60	2.16	0	0

(Continued)

TABLE A-1. Continued

Economy	OPEN90–99 (1 = open)[a]	Average tariff, 1990–99 (percent)[b]	Core nontariff barrier coverage rate (percent), 1995–98[c]	Average BMP, 1990–99 (percent)[d]	Export Marketing Board (1 = country has exporting board)[e]	Socialist (1 = country is socialist)[e]
Czech Republic	1	6.08	—	0.22	0	0
Denmark	1	6.91	—	0	0	0
Dominican Republic	1	16.70	6.20	16.31	0	0
Ecuador	1	11.29	—	9.34	0	0
Egypt, Arab Rep. of	1	30.23	—	12.45	0	0
El Salvador	1	9.38	5.20	13.59	0	0
Estonia	0	1.12	—	25.09	0	0
Ethiopia	0	22.55	—	111.43	0	0
Finland	1	6.91	—	0	0	0
France	1	6.91	—	0	0	0
Gabon	0	19.87	—	1.98	1	0
Gambia, The	1	13.55	—	4.69	0	0
Georgia	1	—	—	0	0	0
Germany	1	6.91	—	0	0	0
Ghana	1	14.93	—	2.96	0	0
Greece	1	6.91	—	1.24	0	0
Guatemala	1	10.27	—	6.03	0	0
Guinea	1	—	—	3.99	0	0
Guinea-Bissau	1	—	—	0	0	0
Guyana	0	13.70	—	28.23	0	0
Haiti	0	10.00	—	81.12	0	0
Honduras	1	8.90	—	9.21	0	0
Hong Kong (China)	1	—	2.10	-0.02	0	0
Hungary	1	12.11	—	5.40	0	0
Iceland	1	3.98	—	1.24	0	0
India	0	48.65	93.80	7.45	0	0
Indonesia	1	16.27	31.30	7.10	0	0
Iran, Islamic Rep.	0	—	—	1,199.31	0	0
Iraq	0	—	—	138,935.90	0	0
Ireland	1	3.98	—	2.50	0	0
Israel	1	7.80	—	2.09	0	0
Italy	1	6.91	—	0	0	0
Jamaica	1	14.68	—	15.46	0	0
Japan	1	5.98	—	-0.35	0	0
Jordan	1	15.83	—	3.37	0	0
Kazakhstan	0	—	—	55.34	0	0
Kenya	1	27.47	—	15.94	0	0
Korea, Rep. of	1	11.28	25.00	0.03	0	0
Kyrgyz Republic	1	—	—	—	0	0
Latvia	1	5.73	—	7.29	0	0
Lesotho	1	17.40	—	3.49	0	0
Liberia	0	—	—	2,306.86	0	0
Lithuania	1	4.33	—	7.45	0	0
Luxembourg	1	6.91	—	0.38	0	0
Macedonia, FYR	1	—	—	18.45	0	0
Madagascar	1	7.13	—	5.93	0	0
Malawi	0	19.80	—	28.83	0	0
Malaysia	1	11.70	19.60	1.35	0	0

(Continued)

TABLE A-1. Continued

Economy	OPEN90−99 (1 = open)[a]	Average tariff, 1990−99 (percent)[b]	Core nontariff barrier coverage rate (percent), 1995−98[c]	Average BMP, 1990−99 (percent)[d]	Export Marketing Board (1 = country has exporting board)[e]	Socialist (1 = country is socialist)[e]
Mali	1	15.66	—	1.98	0	0
Malta	1	7.23	—	1.20	0	0
Mauritania	1	28.23	—	1.55	0	0
Mauritius	1	27.00	16.70	5.25	0	0
Mexico	1	12.53	13.40	2.24	0	0
Moldova	1	—	—	0	0	0
Morocco	1	23.75	13.40	3.54	0	0
Mozambique	1	16.25	—	6.87	0	0
Myanmar	0	5.70	—	2,280.77	0	0
Nepal	0	15.28	—	24.23	0	0
Netherlands	1	6.91	—	0	0	0
New Zealand	1	6.35	—	2.50	0	0
Nicaragua	1	9.90	—	9.98	0	0
Niger	1	18.30	—	1.87	0	0
Nigeria	0	29.74	11.50	151.32	0	0
Norway	1	4.87	—	0	0	0
Pakistan	0	54.73	—	9.74	0	0
Panama	1	10.67	—	0	0	0
Papua New Guinea	0	16.67	—	16.57	1	0
Paraguay	1	10.91	0.00	11.83	0	0
Peru	1	16.80	—	8.75	0	0
Philippines	1	19.09	—	4.36	0	0
Poland	1	12.46	—	2.42	0	0
Portugal	1	6.91	—	2.04	0	0
Romania	0	13.50	—	104.30	0	0
Russian Federation	0	11.24	—	50,979.69	1	0
Rwanda	0	38.40	—	50.78	0	0
Senegal	0	13.05	—	1.98	1	0
Sierra Leone	0	30.25	—	61.47	0	0
Singapore	1	0.32	2.10	0.80	0	0
Slovak Republic	1	7.35	—	5.34	0	0
Slovenia	1	10.60	—	10.06	0	0
Somalia	0	—	—	246.55	0	0
South Africa	1	9.05	8.30	3.46	0	0
Spain	1	6.91	—	1.71	0	0
Sri Lanka	1	24.34	22.70	7.84	0	0
Swaziland	1	15.10	—	7.62	0	0
Sweden	1	6.91	—	0.00	0	0
Switzerland	1	1.38	—	0.00	0	0
Syrian Arab Republic	0	16.00	—	279.97	0	0
Taiwan (China)	1	9.85	—	0.95	0	0
Tajikistan	1	—	—	—	0	0
Tanzania	0	25.12	—	22.17	0	0
Thailand	1	29.54	17.50	1.80	0	0
Togo	0	15.25	—	1.98	1	0
Trinidad and Tobago	1	14.86	—	13.22	0	0
Tunisia	1	28.25	—	3.67	0	0
Turkey	1	15.28	19.80	1.15	0	0

(Continued)

TABLE A-1. Continued

Economy	OPEN90–99 (1 = open)[a]	Average tariff, 1990–99 (percent)[b]	Core nontariff barrier coverage rate (percent), 1995–98[c]	Average BMP, 1990–99 (percent)[d]	Export Marketing Board (1 = country has exporting board)[e]	Socialist (1 = country is socialist)[e]
Turkmenistan	0	—	—	42.86	1	0
Uganda	1	14.37	3.10	19.33	0	0
Ukraine	0	9.73	—	9.02	1	0
United Kingdom	1	6.91	—	0.00	0	0
United States	1	5.96	—	0.00	0	0
Uruguay	1	14.00	0.00	9.88	0	0
Uzbekistan	0	—	—	Dual exchange rate	0	0
Venezuela	1	14.31	17.70	4.13	0	0
Yemen, Re. of	1	20.00	—	8.34	0	0
Serbia and Montenegro	0	—	—	106.44	0	0
Zambia	0	18.43	1.00	62.55	0	0
Zimbabwe	0	20.43	—	132.81	0	0

— not available.

[a]Based on application of Sachs and Warner (1995) criteria; see Wacziarg and Welch (2003) for details.

[b]Unweighted average tariff, 1990–99, based on data from UNCTAD (2001), World Bank (2000), and WTO (various years).

[c]Coverage rate of core nontariff barriers (quotas, licensing, prohibitions, and administered pricing) on capital good and intermediates, based on data from Michalopoulos (1999).

[d]Figures represent [(parallel exchange rate/official exchange rate) − 1] *100, based on data from Easterly and Sewadeh (2002).

[e]Based on literature reviews; see Wacziarg and Welch (2003) for details.

Source: Author compilation.

TABLE A-2. Liberalization and Openness Dates for Countries in Sample

Economy	Period of temporary liberalization (where applicable)	Year uninterrupted openness began[a]	
		Sachs and Warner (1995)	Wacziarg and Welch (2003)
Albania		1992	1992
Algeria		Closed	Closed
Angola		Closed	Closed
Argentina		1991	1991
Armenia		Closed	1995
Australia		1964	1964
Austria		1960	1960
Azerbaijan		Closed	1995
Bangladesh		Closed	1996
Barbados		1966	1966
Belarus		1994	Closed

TABLE A-2. Continued

Economy	Period of temporary liberalization (where applicable)	Year uninterrupted openness began[a]	
		Sachs and Warner (1995)	Wacziarg and Welch (2003)
Belgium		1959	1959
Benin		1990	1990
Bolivia	1956–79	1985	1985
Botswana		1979	1979
Brazil		1991	1991
Bulgaria		1991	1991
Burkina Faso		Closed	1998
Burundi		Closed	1999
Cameroon		1993	1993
Canada		1952	1952
Cape Verde		n.a.	1991
Central African Republic		Closed	Closed
Chad		Closed	Closed
Chile		1976	1976
China		Closed	Closed
Colombia		1986	1986
Congo, Dem. Rep. of		Closed	Closed
Congo, Rep. of		Closed	Closed
Costa Rica	1952–61	1986	1986
Côte d'Ivoire		Closed	1994
Croatia		1993	Closed
Cyprus		1960	1960
Czech Republic		1991	1991
Denmark		1959	1959
Dominican Republic		Closed	1992
Ecuador	1950–82	1991	1991
Egypt, Arab Rep.		Closed	1995
El Salvador	1950–61	1989	1989
Estonia		1992	Closed
Ethiopia		Closed	1996
Finland		1960	1960
France		1959	1959
Gabon		Closed	Closed
Gambia, The		1985	1985
Georgia		Closed	1996
Germany		1959	1959
Ghana		1985	1985
Greece		1959	1959
Guatemala	1950–61	1988	1988
Guinea		1986	1986
Guinea-Bissau		1987	1987
Guyana		1988	1988
Haiti		Closed	Closed
Honduras	1950–61	1991	1991

(Continued)

TABLE A-2. Continued

| Economy | Period of temporary liberalization (where applicable) | Year uninterrupted openness began[a] | |
		Sachs and Warner (1995)	Wacziarg and Welch (2003)
Hong Kong (, China)		Always open	Always open
Hungary		1990	1990
Iceland		n.a.	n.a.
India		1994	Closed
Indonesia		1970	1970
Iran, Islamic Rep. of		Closed	Closed
Iraq		Closed	Closed
Ireland		1966	1966
Israel		1985	1985
Italy		1959	1959
Jamaica	1962–73	1989	1989
Japan		1964	1964
Jordan		1965	1965
Kazakhstan		Closed	Closed
Kenya	1963–67	1993	1993
Korea, Rep. of		1968	1968
Kyrgyz Republic		1994	1994
Latvia		1993	1993
Lesotho		n.a.	Closed
Liberia		n.a.	Closed
Lithuania		1993	1993
Luxembourg		1959	1959
Macedonia, FYR		1994	1994
Madagascar		Closed	1996
Malawi		Closed	Closed
Malaysia		1963	1963
Mali		1988	1988
Malta		n.a.	Closed
Mauritania		1992	1995
Mauritius		1968	1968
Mexico		1986	1986
Moldova		1994	1994
Morocco	1956–64	1984	1984
Mozambique		Closed	1995
Myanmar		Closed	Closed
Nepal		1991	1991
Netherlands		1959	1959
New Zealand		1986	1986
Nicaragua	1950–60	1991	1991
Niger		Closed	1994
Nigeria		Closed	Closed
Norway		Always open	Always open
Pakistan		Closed	2001
Panama		n.a.	1996

(Continued)

TABLE A-2. Continued

Economy	Period of temporary liberalization (where applicable)	Year uninterrupted openness began[a]	
		Sachs and Warner (1995)	Wacziarg and Welch (2003)
Papua New Guinea		Closed	Closed
Paraguay		1989	1989
Peru	1948–67	1991	1991
Philippines		1988	1988
Poland		1990	1990
Portugal		Always open	Always open
Romania		1992	1992
Russian Federation		Closed	Closed
Rwanda		Closed	Closed
Senegal		Closed	Closed
Serbia and Montenegro		Closed	2001
Sierra Leone		Closed	2001
Singapore		1965	1965
Slovak Republic		1991	1991
Slovenia		1991	1991
Somalia		Closed	Closed
South Africa		1991	1991
Spain		1959	1959
Sri Lanka	1950–56; 1977–3	1991	1991
Swaziland		n.a.	Closed
Sweden		1960	1960
Switzerland		Always open	Always open
Syrian Arab Republic	1950–65	Closed	Closed
Taiwan (China)		1963	1963
Tajikistan		Closed	1996
Tanzania		Closed.	1995
Thailand		Always open	Always open
Togo		Closed.	Closed
Trinidad and Tobago		Closed.	1992
Tunisia		1989	1989
Turkey	1950–53	1989	1989
Turkmenistan		Closed	Closed
Uganda		1988	1988
Ukraine		Closed	Closed
United Kingdom		Always open	Always open
United States		Always open	Always open
Uruguay		1990	1990
Uzbekistan		Closed	Closed
Venezuela, R.B. de	1950–59; 1989–93	Closed	1996
Yemen, Rep.		Always open	Always open
Zambia		1993	1993
Zimbabwe		Closed	Closed

Note: n.a. means not classified. Closed denotes countries closed as of 1994 in the Sachs-Warner column and closed as of 2000 in the Wacziarg-Welch column.

[a]Based on the Sachs and Warner (1995) criteria and broader literature review. See Wacziarg and Welch (2003) for details.

TABLE A-3 Trade Liberalization and Concurrent Events in Subsample of 13 Countries

Country (year of liberalization)	Sample period	Concurrent events
Countries that experienced negative or zero growth after liberalization		
Botswana (1979)	1961–98	Since gaining independence in 1966, Botswana has had one of the fastest growth rates in the world (IMF 2002), growing at an annual rate of 7.7 percent between 1965 and 1998 (Rodrik 2003). Income per capita (in purchasing parity power–adjusted terms in 1998) was four times the African average.
		Botswana experienced a mean growth difference of −1.99 percent in the years before and after liberalization. This differs that despite trade liberalization and an export-oriented economy, government intervention has been high in Botswana, where the public sector accounts for a large share of the economy.
		Botswana's economy expanded when diamond mining began in 1971. The recession of 1981/82 was partly a result of a weak world diamond market. The late 1980s were a period of new mining activity and strong demand that supported overall economic growth. During the early to mid-1990s, recessionary conditions on the world diamond market led to a severe economic slump in Botswana. Because diamond exports account for 70 percent of Botswana's export earnings and more than a third of its GDP, volatile diamond prices have had a significant impact on the country's overall economic growth. Despite volatility, growth remained positive throughout most of the sample period, however (EIU various years; IMF 2002).
Colombia (1986)	1951–98	In December 1990, Colombia was unable to repay its debt principal payments; it was unable to refinance its debt until April 1991. In the wake of this crisis, Colombia pursued a variety of market-oriented reforms in addition to further trade liberalization. Price controls were lifted, a financial sector reform was implemented, the exchange control system was liberalized, the regulatory framework was modernized, and some industries were privatized (Wacziarg and Wallack 2004).
		Throughout the 1990s, substantial trade reforms were implemented, including bilateral trade agreements with other Latin American countries in 1993/94 (Henry 1999). The rise in civil unrest beginning in 1992 and the political instability that persisted throughout the 1990s likely limiting postliberalization economic growth.

(Continued)

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
Hungary (1990)	1971–98	Hungary experienced a period of declining growth and poor economic conditions between 1971 and 1991. In 1988–89, the leader of the majority party changed after 30 years, and a period of political uncertainty ensued. In 1989, the new government implemented a stabilization program that included higher taxes, tighter monetary policy, and the devaluation of the currency (World Bank 1995).
		In 1990/91, Hungary implemented an IMF restructuring program. In 1995, it implemented structural reforms, including currency devaluation, a new exchange rate mechanism, a tight wage policy in the public sector, and fiscal measures to enhance revenues and cut expenditures. Hungary accelerated privatization efforts, restructuring enterprises (including major commercial banks) and implementing financial sector and public finance reforms in the mid-1990s. Significant improvements were also made in the legal and regulatory framework of the financial sector (Wacziarg and Wallack 2004).
		Economic recovery began in 1992/93. During the mid-1990s, Hungary adhered to the IMF plan and experienced gradual stabilization and recovery. Growth did not return to the levels seen before liberalization, however. Persisting high levels of debt and current account deficits may have limited the gains from trade liberalization. In addition, in 1993 the government tightened monetary policy and increased interest rates, which likely dampened the economic recovery (World Bank 1995). While structural reforms were implemented in 1995, the full effects may not have been evident before the end of the study period (1998).
Israel (1985)	1951–98	The wars of 1967 and 1973 limited economic growth. In 1977, both tariff and currency barriers were relaxed; in 1979, the government approved a five-year plan to reduce inflation and customs rates. In January 1980, tariffs were further reduced on imports from the European Economic Community.
		Israel invaded Lebanon in June 1982; roughly a year later it entered a deep economic crisis, characterized by triple-digit inflation, a widening trade gap, rapidly mounting foreign debt, and significant real exchange rate appreciation. In July 1985, the government implemented an emergency economic stabilization plan in order to stop hyperinflation; it also signed a free trade agreement with the United States (Henry 1999). Inflation dropped significantly in late 1985 and 1986, and the IMF announced its support of Israeli reform efforts. In 1986, Israel fixed the exchange rate to a trade-weighted currency basket. In January 1987, it devalued the currency 19 percent and implemented other changes affecting the tax system and money markets.

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		Despite devaluations of the currency in 1988 and 1989, the interest rate increased, because of currency volatility. In 1991, Israel implemented a crawling band exchange rate system. The shekel was devalued by 6 percent in order to boost the economy, which was suffering as a result of the Gulf War. In November 1995, a free trade area treaty affirming Israel's special trade status with the European Union was signed (Henry 1999).
		Despite trade reforms implemented throughout the period, Israel's heterodox stabilization program may have offset the effects of trade liberalization. Social pacts based on broad coalitions of labor, government, and industry set the patterns for prices, wages, and the exchange rate (Wacziarg and Wallack 2004). In addition, inflation, currency volatility, and high interest rates in the late 1980s and early 1990s reduced Israel's competitiveness and the gains from trade.
Mexico (1986)	1951−98	The 1940s−1960s was a period of political and social stability and relatively rapid economic growth in Mexico (Tornell 2002). In the early 1970s, expansionary fiscal and monetary policy led to increasing levels of debt, escalating prices, and an overvaluation of the exchange rate. By 1976, inflation was increasing and private investment decreasing. In August 1976, the government was forced to devalue the peso and decrease government expenditure (Gonzalez 1994).
		The discovery of oil in 1977 stimulated the economy between 1978 and 1982: in 1981, oil accounted for three-fourths of Mexico's exports. Government spending, financed by international borrowing, increased, however, resulting in the overvaluation of the peso. By mid-1981, the international price of oil had fallen; by 1982, Mexico declared itself unable to service its debt. The government devalued the peso by 30 percent in February 1982 and implemented a two-tiered foreign exchange system in August 1982. Mexico experienced a severe recession during the Latin American debt crisis of 1982−83 (Gonzalez 1994).

(Continued)

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		In 1984, Mexico pursued a policy of privatization and liberalization in order to attract FDI (Henry 1999). In 1985, it implemented a program of stabilization and structural adjustment, including trade liberalization. It joined GATT in 1986 and significantly reduced import restrictions and tariff barriers. A debt-rescheduling agreement was signed in August 1985. In July 1986, an IMF agreement was implemented, facilitating additional debt restructuring. Further trade liberalization measures were implemented in August 1987 (Henry 1999). Mexico also pursued a privatization program during the 1980s, which continued into the 1990s, with the privatization of the banking industry (Henry 1999; Wacziarg and Wallack 2004).
		An economic and fiscal crisis occurred again in 1994–95. It was accompanied by a period of political unrest, including the Chiapas uprising and the assassination of several PRI figures (Henry 1999). In December 1994, Mexico devalued the peso and implemented a floating exchange rate regime. In 1995, it received a bailout, which prevented it from defaulting on its debt and granted it continued access to international capital markets (Tornell 2002).
		Despite the economic recovery and trade liberalization that occurred in the late 1980s, Mexico never recovered to its precrisis levels of growth. The persisting macroeconomic instability and lack of additional structural reforms appear to have been key factors in limiting the gains from trade liberalization by preventing economic restructuring and reallocation of resources. According to the IMF (1999), further banking sector reforms and continued economic restructuring were necessary to sustain economic growth. The macroeconomic environment was hindered by the volatile price of oil, uncertainty regarding debt negotiations, and speculative attacks on the peso. As the government decreased expenditure under the structural adjustment program, domestic demand fell. During the late 1980s and early 1990s, the currency became overvalued again, effectively offsetting trade liberalization measures. Nontrade barriers to competition also existed, in the form of government monopolies and oligopolies, which limited restructuring. The high interest rate, aimed at preventing speculative attacks and attracting foreign capital, limited domestic demand and restructuring (Gonzalez 1994).The economy improved between 1995 and 1998 as a result of the implementation of structural reforms and the success of the floating exchange rate in mitigating the effect of external shocks (IMF 1999).

(Continued)

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
Philippines (1988)	1951–98	During the 1960s the Marcos regime increased trade barriers, which remained in effect until the 1980s. During the 1983–86 economic crisis, the inflation rate increased significantly. The currency was devalued by 50 percent in 1984, and expansionary monetary policy limited capital inflow and economic growth. The Philippines secured debt rescheduling agreements between 1985 and 1988, and the IMF approved a stabilization plan in 1989.
		In 1986 (the end of the Marcos era) the Philippines implemented trade liberalization measures, including the lifting of import restrictions. Despite these reforms government investment in state enterprises roughly doubled during the sample period, as did state enterprises' percentage of total economic activity, employment, and net financial flows (Wacziarg and Wallack 2004)
		The Philippines implemented capital market liberalization, including reform of the foreign exchange rate, in 1992. The IMF approved the country's economic performance and rescheduled additional debt. Further trade reforms, including the removal of quantitative restrictions, were also implemented during the early 1990s (Henry 1999).
		Despite further trade liberalization measures, the Philippines has not witnessed the increased economic growth experienced in other countries following liberalization, possibly because of limited structural reforms and the high level of government involvement in state enterprises. Pritchett (2003) cites the institutional uncertainty that arose from political instability in the Philippines following liberalization as a factor that may have limited investment and economic growth.

Countries that experienced positive growth increased after liberalization

Country (year of liberalization)	Sample period	Concurrent events
Chile (1976)	1952–98	When Salvatore Allende assumed power in 1970, he nationalized Chile's copper mines, banks, and other enterprises. Government expenditure increased dramatically: the country's budget deficit rose from 2.7 percent of GDP to 25.0 percent between 1970 and 1973. The black market currency premium exceeded 600 percent in 1972; inflation exceeded 100 percent in 1973 (Easterly and Sewadeh 2002; Stallings and Brock 1993).

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		In 1973, Augusto Pinochet overthrew President Allende in a military coup. Between 1975 and 1982, structural changes to liberalize the financial system were implemented. Quantitative restrictions were eliminated in 1973; tariffs were significantly reduced between 1973 and 1979, when they were set at a uniform rate of 10 percent. In 1979, the exchange rate was fixed to the U.S. dollar, capital controls reduced, the tax system simplified, and privatization pursued (Stallings and Brock 1993).
		Trade reform in the early years of the pro-market Pinochet administration was accompanied by privatization, elimination of the fiscal deficit, and the lifting of price and interest rate controls. Liberalization of the labor market also facilitated overall economic restructuring (Wacziarg and Wallack 2004).
		In 1980/81, Chile privatized its social security system and implemented banking reforms. It experienced an economic crisis during the Latin American debt crisis, during which it was unable to access credit markets and the government assumed control of troubled banks. In 1982, GDP fell 14 percent and inflation doubled.
		Between 1982 and 1985, the peso was devalued, tariff rates were raised to 35 percent, and privatization efforts were reversed (Stallings and Brock 1993). In 1985, the peso was gradually depreciated with a crawling peg, tariffs were reduced to 15 percent, and privatization resumed. During the mid- to late-1980s, Chile decreased tariffs, rescheduled its debt, and reprivatized the banking sector. During the 1990s, it signed free trade agreements with Colombia and Mexico and engaged in substantial capital market liberalization (Henry 1999).
Ghana (1985)	1956–98	Upon gaining independence in 1957, Ghana pursued a strategy of import substitution. It implemented a series of restrictive trade policies, including tariffs, nontariff barriers, and exchange rate controls. It also established a variety of state enterprises. By 1966, Ghana's currency was overvalued and a cycle of political instability (including military coups) began. Rising inflation followed by currency devaluations ensued during the late 1960s and 1970s (Leith and Lofchie 1993).

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		Ghana experienced another economic crisis in 1982, during which inflation increased and foreign exchange reserves dropped to very low levels. In 1983, the government launched a four-year economic recovery program that included restructuring the country's physical infrastructure and economic institutions and reducing inflation through prudent monetary, fiscal, and trade policies. The 1985 trade liberalization program was part of the Rawlings administration's World Bank- and IMF-supported economic recovery program. Multiple exchange rates implemented to promote exports were later replaced with unified rates and subjected to a series of devaluations. Public sector employment (including in state enterprises) was cut and distortions in wages reduced (Wacziarg and Wallack 2004). Ghana continued to implement trade and capital market reforms through the late 1980s and 1990s.
Indonesia (1970)	1961–98	Indonesia suffered an economic crisis during the early 1960s, during which budget deficits rose and annual inflation reached 640 percent. Under pressure from the army, in March 1966 President Sukarno transferred some power to Suharto; in March 1967 Suharto was named president. A five-year development plan to stabilize the economy and promote growth was implemented that successfully stabilized the economy. Capital market liberalization occurred in 1970. In February 1976, the government reduced the 10 percent export duty on a wide range of commodities. During the 1970s, government intervention increased despite the implementation of trade liberalization reforms. The government increased its control of state-owned banks and other enterprises. Oil revenue was significant during the 1970s; economic growth weakened in the early 1980s, partly as a result of falling oil prices. However, the impact was mitigated by the country's rapid adjustment and a relatively low debt burden (Temple 2003). In June 1983, the government announced a series of bank liberalization reforms, followed by further reforms in 1988 when credit subsidies were removed (Temple 2003). Devaluations occurred in 1983 and 1986.

(Continued)

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		During 1984/85, Indonesia entered into bilateral trading agreements with the Soviet Union, the United States, and several other countries. In May 1986, the government announced new measures to attract foreign investment. The oil market crashed in the second quarter of 1986. During 1986, further trade and investment liberalization measures were implemented, with the gradual removal of qualitative restrictions and nontariff barriers.
		The government implemented a large-scale privatization program during the late 1980s and early 1990s. Between 1991 and 1995, it implemented banking reforms to strengthen the system; to stimulate lending, it later weakened these regulations. During 1991–95, capital market reforms aimed at improving stock exchange were implemented (EIU various years; Henry 1999).
Korea, Rep. of (1968)	1954–98	Political turmoil in Korea in the late 1950s forced President Syngman Rhee's resignation in 1960. A military coup followed in 1961, along with continued political unrest. Inflation increased and foreign exchange reserves decreased significantly before Korea stabilized and started its slow transition to democratic rule in 1964 (Haggard, Cooper, and Moon 1993).
		Korea transitioned from a policy of import substitution to export-oriented growth during the mid-1960s. Tariffs and nontariff barriers were reduced, and the government created export-processing zones and adopted other mechanisms for increasing FDI (Sakurai 1995). The currency was devalued, the tax system and interest rates reformed, and capital markets liberalized. In 1965, the Export Development Committee was established; in 1966, quantitative restrictions were eliminated. Liberalization was not universal, however; certain sectors remained protected, and government involvement in the economy remained pervasive.
		The assassination of President Chung-Hee Park in November 1979 unleashed a year of political and economic crisis. In 1980, significant banking reforms were announced; in 1981, a five-year economic plan of structural adjustment was initiated. Economic growth was dampened during the Asia financial crisis of 1982–84.

(Continued)

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
		Capital and banking sector reforms were implemented in 1984. Further trade reforms, including reductions in tariffs and nontariff barriers, were implemented in the mid- to late-1980s. Banking and capital market reforms were deepened in 1991 in an effort to attract FDI. In 1993, a five-year plan for reform and further financial system regulation was adopted (Henry 1999).
Poland	1971–98	Poland's economy collapsed during the 1970s. In August 1980, the Solidarity movement began, and a period of political unrest ensued. Martial law remained in effect through December 1982. In 1986, Poland was accepted into the IMF and began to pursue debt restructuring. In 1989, hyperinflation impeded economic growth, causing Poland's debt to reach 74 percent of GDP (de Menil 2003).
		In 1990, the government implemented a swift and comprehensive set of market reforms, including trade liberalization, in order to stabilize the economy. The Balcerowicz Plan included removal of price controls, reduction of government expenditure and investment, devaluation of the exchange rate, and removal of subsidies for energy (Wacziarg and Wallack 2004). Trade liberalization measures included the liberalization and elimination of exchange controls and the abolition of state trading monopolies and nearly all quotas and tariffs. The currency was devalued by more than 50 percent in January 1990 and then gradually depreciated based on a crawling peg until 1995 (de Menil 2003).
		Despite a deep recession in 1991, Poland persisted in its liberalization program, implementing a new IMF plan in 1993 that included tax reform and continued privatization. During the mid-1990s, Poland continued to implement reforms, including currency reform, privatizations, and policies, to promote FDI. It applied for EU membership in 1994 and became a member of the Organisation for Economic Co-operation and Development in 1996.
		According to de Menil (2003), productivity gains appear to have been the primary factors in Poland's growth during the 1990s. He believes that comprehensive structural reforms facilitated economic transformation, the reallocation of resources, and the rapid adoption of Western principles of management and standards of efficiency.

TABLE A-3 Continued

Country (year of liberalization)	Sample period	Concurrent events
Taiwan (China) (1963)	1952–98	Most of the economies in this subsample implemented trade liberalization in the wake of economic and often political crises. In contrast, Taiwan had a stable economic environment and relatively low tariff rates at the time of liberalization. Trade liberalization in the early 1960s involved further tariff reductions as well as incentives, such as the creation of export-promotion zones, to attract FDI (Sakurai 1995).
		Between 1985 and 1987, Taiwan further reduced tariffs and nontariff barriers. In 1985, it implemented polices to promote FDI and liberalize the foreign exchange market. In 1987, it tightened capital controls. In 1988, it implemented capital market reform measures along with additional trade reform measures. In 1989–92, it implemented banking reforms and privatization measures (Henry 1999).
Uganda (1988)	1951–98	Between independence (in 1962) and 1980, Uganda experienced economic devastation, as a result of mismanagement and war. Capital was destroyed, and manufacturing operated at extremely low capacity. In 1981, Uganda implemented an IMF reform program that included floating the currency, removing price controls, and imposing fiscal austerity. The reform program was initially successful, but success was not sustained and the IMF withdrew its support in 1984, a year that marked the beginning of a period of economic collapse and civil war.
		In 1985, Uganda implemented policies to promote FDI and liberalize the foreign exchange market. A new economic recovery program was launched in 1987. Political unrest led to a tightening of the capital market, however. In 1988, further trade and capital market liberalization measures resumed, followed by banking reforms and privatization between 1989 and 1992. In 1993–94, further trade and capital market liberalization measures were implemented, including the liberalization of the interest rate. A variety of currency regimes was implemented between 1988 and 1992; the currency was pegged to the U.S. dollar and a composite of other currencies before a flexible exchange rate system was adopted in 1996 (Amvouna 1998).

Source: Author compilation.

REFERENCES

Ades, Alberto F., and Edward L. Glaeser. 1999. "Evidence on Growth, Increasing Returns and the Extent of the Market." *Quarterly Journal of Economics* 114(3):1025–45.

Alesina, Alberto, Enrico Spolaore, and Romain Wacziarg. 2000. "Economic Integration and Political Disintegration." *American Economic Review* 90(5):1276–96.

Amvouna, Anatolie Marie. 1998. "Determinants of Trade and Growth Performance in Africa: A Cross-Country Analysis of Fixed versus Floating Exchange Rate Regimes." African Economic Policy Discussion Paper *16*. USAID Equity and Growth through Economic Research (EAGER) Project. U.S. Agency for International Development, Washington, D.C.

Baldwin, Richard, and Elena Seghezza. 1996. *Testing for Trade-Induced, Investment-Led Growth*. NBER Working Paper 5416. Cambridge, Mass.: National Bureau of Economic Research.

Bekaert, Geert, and Campbell R Harvey. 2000. "Foreign Speculators and Emerging Equity Markets." *Journal of Finance* 55(2):565–614.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad. 2001. *Does Financial Liberalization Spur Growth?* NBER Working Paper 8245. Cambridge, Mass.: National Bureau of Economic Research.

Ben-David, Dan. 1993. "Equalizing Exchange: Trade Liberalization and Income Convergence." *Quarterly Journal of Economics* 108(3):653–79.

de Menil, Georges. 2003. "History, Policy, and Performance in Two Transition Economies: Poland and Romania." In Dani Rodrik, ed., *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton, N.J.: Princeton University Press.

Dollar, David. 1992. "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs 1976–1985." *Economic Development and Cultural Change* 40(3):523–44.

Easterly, William, and Mirvat Sewadeh. 2002. Global Development Network Growth Database. World Bank, Washington, D.C.

EBRD (European Bank for Reconstruction and Development). 1994. *Transition Report: Economic Transition in Eastern Europe and the Former Soviet Union*. London: European Bank for Reconstruction and Development.

Edwards, Sebastian. 1992. "Trade Orientation, Distortions and Growth in Developing Countries." *Journal of Development Economics* 39(1):31–57.

EIU (Economist Intelligence Unit). Various years. *Country Profiles and Quarterly Economic Reports*. London.

Frankel, Jeffrey A., and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89(3):379–99.

González, Diana Alarcón. 1994. *Changes in the Distribution of Income in México and Trade Liberalization*. Tijuana, México: El Colegio de la Frontera Norte Press.

Haggard, Stephen, Richard N. Cooper, and Chung-in Moon. 1993. "Policy Reform in Korea in Political and Economic Interactions." In Robert H. Bates, and Anne O. Krueger, eds., *Economic Policy Reform: Evidence from Eight Countries*. Cambridge, Mass.: Blackwell.

Harrison, Ann, and Gordon Hanson. 1999. "Who Gains from Trade Reform? Some Remaining Puzzles." *Journal of Development Economics* 59(1):125–54.

Henry, Peter Blair. 1999. "Chronological Listing of Major Policy Events in Emerging Markets." Working paper. Stanford, Calif.: Stanford University, Graduate School of Business. (https://faculty-gsb.stanford.edu/henry/Homepage/PDF/Reform%20List.pdf)

———. 2000. "Stock Market Liberalization, Economic Reform, and Emerging Market Equity Prices." *Journal of Finance* 55(2):529–64.

Heston, Alan, Robert Summers, and Bettina Aten. 2002. *Penn World Table Version 6.0*. Center for International Comparisons (CICUP), University of Pennsylvania, Philadelphia.

Husain, Ishrat, and Rashid Faruqee. 1994. *Adjustment in Africa: Lessons from Country Case Studies*. Washington, D.C.: World Bank.

IMF (International Monetary Fund). 1999. Press Release 99/28. 7 July.

———. 2002. "IMF Concludes 2002 Article IV Consultation with Botswana." Public Information Notice, 1 November. Washington, D.C.

Kornai, Janos. 2002. *The Socialist System: The Political Economy of Communism*. Oxford: Clarendon Press.

Leith, J. Clark, and Michael F. Lofchie. 1993. "The Political Economy of Structural Adjustment in Ghana." In Robert H. Bates, and Anne O. Krueger, eds., *Political and Economic Interactions in Economic Policy Reform: Evidence from Eight Countries*. Cambridge, Mass.: Blackwell.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 84(4):942–63.

Michalopoulos, Constantine. 1999. "Trade Policy and Market Access Issues for Developing Countries." Policy Research Working Paper 2214. World Bank, Washington, D.C.

Pritchett, Lant. 2003. "A Toy Collection, a Socialist Star, and a Democratic Dud: Growth Theory, Vietnam, and the Philippines." In Dani Rodrik, ed., *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton, N.J.: Princeton University Press.

Rodríguez, Francisco, and Dani Rodrik. 2000. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In Ben Bernanke, and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2000*. Cambridge, Mass.: MIT Press.

Rodrik, Dani. 2003. *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton, N.J.: Princeton University Press.

Sachs, Jeffrey D., and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 1:1–118.

Sakurai, Makoto. 1995. "East Asian Agricultural Trade and Protectionism." In Toshihiko Kawagoe, and Sueo Sekiguchi, eds., *East Asian Economies: Transformation and Challenges*. ASEAN Economic Research Unit, Institute of Southeast Asian Studies, Singapore.

Stallings, Barbara, and Philip Brock. 1993. "The Political Economy of Economic Adjustment: Chile." In Robert H. Bates, and Anne O. Krueger, eds., *Political and Economic Interactions in Economic Policy Reform: Evidence from Eight Countries*. Cambridge, Mass.: Blackwell.

Temple, Jonathan. 2003. "Growing into Trouble: Indonesia after 1966." In Dani Rodrik, ed., *In Search of Prosperity: Analytic Narratives on Economic Growth*. Princeton, N.J.: Princeton University Press.

Tornell, Aaron. 1998. *Reform from Within*. NBER Working Paper 6497. Cambridge, Mass.: National Bureau of Economic Research.

———. 2002. "Economic Crises and Reform in Mexico." In Stephen Haber, ed., *Crony Capitalism and Economic Growth in Latin America: Theory and Evidence*. Stanford, Calif.: Hoover Institution Press.

UNCTAD (United Nations Conference on Trade and Development). 2001. Trade Analysis and Information System (Trains). CD-ROM. Version 8.0. Geneva: United Nations Conference on Trade and Development.

Wacziarg, Romain. 2001. "Measuring the Dynamic Gains from Trade." *World Bank Economic Review* 15(3):393–429.

———. 2003. "India in the World Trading System." Working paper. Stanford, Calif: Graduate School of Business, Stanford University.

Wacziarg, Romain, and Jessica Seddon Wallack. 2004. "Trade Liberalization and Intersectoral Labor Movements." *Journal of International Economics* 64(2):411–39.

Wacziarg, Romain, and Karen Horn Welch. 2003. *Trade Liberalization and Growth: New Evidence*. NBER Working Paper 10152. Cambridge, Mass.: National Bureau of Economic Research.

World Bank. 1995. *Hungary: Structure Reforms for Sustainable Growth*. World Bank Country Study. Washington, D.C.

———. 2000. International Trade and Development: Data on Import Tariff and NTBs. Washington, D.C. (www1.worldbank.org/wbiep/trade/TR_data.html).

WTO (World Trade Organization). 1996–2002. *Trade Policy Reviews*. Various countries. Geneva.

Comprehensive Wealth and Future Consumption: Accounting for Population Growth

Susana Ferreira, Kirk Hamilton, and Jeffrey R. Vincent

Economic theory predicts that the current change in national wealth, broadly defined to include natural and human capital as well as produced capital ("genuine savings"), determines whether the present value of future changes in consumption is positive or negative. Theoretical research has focused on the effects of population growth on this relation, but no rigorous empirical investigation has been conducted. Panel data for 64 developing countries during 1970–82 are used to test the effects of three adjustments for population growth, including one that controls for omitted wealth. Although the adjustments have substantial impacts on estimates of genuine savings, they lead to only limited improvements in the relation between those estimates and subsequent consumption changes. Even without adjustments for population growth, adjustments for natural resource depletion improve the relation significantly. Policymakers and economists can interpret published estimates of genuine savings as signals of future consumption paths if and only if the estimates include adjustments for natural resource depletion. But better estimates of capital stocks are needed before it can be confidently said that adjustments for population growth significantly improve the accuracy of those signals. JEL codes: O40, Q01, C33.

A well-established body of economic theory indicates that the current change in a country's wealth—broadly defined to include natural and human capital as well as produced capital—determines whether the country's stream of future consumption will lie above or below its current consumption level.[1] This comprehensive measure of the net change in national wealth is commonly referred

1. Seminal articles include Weitzman (1976), Hartwick (1977), Dasgupta and Heal (1979), and Dixit, Hammond, and Hoel (1980). An edited volume by Perrings and Vincent (2003) contains a selection of theoretical and empirical studies. Dasgupta (2001) and Weitzman (2003) provide comprehensive theoretical treatises.

to as "genuine savings" (Hamilton and Clemens 1999). An implication is that current levels of consumption are unsustainable in countries in which genuine saving rates are negative. An example is a country in which natural resource depletion is outstripping investment in more reproducible forms of capital.

This theory—and evidence of a lack of sustained economic growth in many resource-rich developing countries—prompted the World Bank, the United Nations, and other international and national organizations to launch initiatives aimed at creating more-comprehensive national wealth accounts (United Nations 1993;World Bank 1997, 2006).[2] The World Bank's (various years) widely used *World Development Indicators* now includes data on net national saving rates adjusted for depreciation of produced capital and depletion of natural resources, in addition to conventional savings measures, which account only for gross investment in produced capital.

The proposition that changes in capital stocks have an impact on future economic performance is neither new nor surprising. One of the few consistent findings in the extensive cross-country growth literature is that gross investment in produced capital is significantly correlated with growth in gross domestic product (GDP). An early review by Levine and Renelt (1992), which characterized most of the conclusions from initial growth studies as "fragile," reported a robust, positive correlation between gross investment and GDP growth. Sala-i-Martin (1997) subsequently argued that Levine and Renelt's extreme-bounds analysis constituted an overly stringent test of significance. Sala-i-Martin based his analysis on complete distributions of the regression coefficients, but he, too, found that gross investment was significantly correlated with GDP growth.[3]

The genuine savings literature differs from the growth literature in two principal ways. One is that it focuses on changes in consumption rather than GDP growth, which makes it more welfare oriented. The other is that it emphasizes net savings measures that are broader than ordinary gross investment. It draws particular attention to changes in natural resource stocks.

Ferreira and Vincent (2005) analyze the performance of the broader net savings measures published in the *World Development Indicators* (World Bank various years). Using panel data for 1970–2001 they test whether measures adjusted for the depreciation of produced capital and the depletion of natural capital were correlated with subsequent changes in countries' consumption levels. They find that these measures are correlated in non–Organisation for

2. Evidence of a "resource curse" comes from both case studies (for example, Gelb and Associates 1988) and cross-country econometric studies (for example, Sachs and Warner 1995).

3. Since Burnside and Dollar's (2000) study, the empirical growth literature has focused largely on the effectiveness of foreign aid, but it has continued to provide evidence that gross investment contributes positively to GDP growth. These results refer to long-run growth; Easterly (1999) shows that there is no evidence of a link between investment and growth in the short run. Easterly and Levine (2001) emphasize the role of total factor productivity growth, rather than factor accumulation, in explaining differences in income and growth across countries.

Economic Co-operation and Development (OECD) countries: adjusted net savings in a given year had a positive and significant impact on consumption changes during the subsequent decade. They find no evidence of a similar relation in OECD countries, a result they attribute to the fact that the *World Development Indicators* measures refer purely to factor accumulation and exclude technical change.

The effects of population growth have been a recent focus of theoretical research on genuine savings (Dasgupta 2001; Dasgupta and Mäler 2001; Hamilton 2002; Arrow, Dasgupta, and Mäler 2003; Asheim 2004). This research indicates that population growth can reduce per capita consumption possibilities by spreading existing capital stocks more thinly across a larger number of people and that empirical genuine savings estimates should account for this wealth-diluting (capital-widening) effect. Using data for a small sample of developing countries and regions, Dasgupta (2003) shows that this effect can be large enough to flip genuine savings estimates from positive to negative. Estimates that ignore this effect run the risk of exaggerating countries' future consumption possibilities.

This article investigates whether current per capita savings is correlated with future changes in per capita consumption. It uses the formulas linking saving and future consumption derived from the underlying growth theory. It takes wealth dilution and several other effects of population growth into account, including the effects of changes in the population growth rate over time.

The sample is a panel that includes annual data on 64 developing countries for 1970–82, with additional data for 1983–2003 used to construct some of the variables. The analysis is limited to developing countries in view of Ferreira and Vincent's (2005) results and because developing countries are the countries in which population growth rates are highest and thus affect genuine savings estimates the most. As by-products, additional evidence is generated suggesting that adjustments for natural resource depletion improve the performance of genuine savings measures. The analysis also generates econometric estimates of national wealth components not included in produced and natural capital.

The article is organized as follows. Section I summarizes pertinent economic theory, which leads into the specification of the econometric models in section II. Section III describes the data sources, the construction of key variables, and estimation issues. Section IV presents the results. The last section discusses the implications of the results for the construction and use of estimates of cross-country genuine savings.

I. Theoretical Underpinnings

The test of savings measures used here builds on the result in Dasgupta (2001) that states that current genuine savings equals the present value of future changes in consumption. Hamilton and Hartwick (2005) derive this result as a corollary to a more general result that links growth in savings to growth in

consumption. They show that the following relation holds in a competitive economy:[4]

$$(1) \qquad \int_t^\infty \frac{dC(v)}{dv} e^{-\int_t^v \rho(\tau)d\tau} dv = G(t)$$

where C is total (not per capita) consumption, G is total genuine savings, and ρ is the interest rate (the consumption discount rate, not the utility discount rate). This relation rests on less restrictive assumptions than the theoretical relation that provides the basis for Ferreira and Vincent's (2005) econometric test. Following Weitzman (1976), they examine the relation between current genuine savings and the difference between average future consumption and current consumption, not the present value of future consumption changes.

Dasgupta (2001) shows that per capita genuine savings mirrors the change in social welfare ("dynamic average utilitarianism") in an economy with a growing population under three conditions: the population grows at a constant rate, per capita consumption is independent of population size, and production has constant returns to scale. Under these assumptions the relation between current genuine per capita savings and future changes in per capita consumption is given by the following equation:

$$(2) \qquad \int_t^\infty \frac{dc(v)}{dv} e^{\int_t^v -(\rho(\tau)-\gamma)d\tau} dv = g(t).$$

In this equation c is per capita consumption (C/N, where N is population, assumed to equal the labor force); g is genuine per capita savings (G/N); and γ is the constant population growth rate.

The inclusion of population growth leads to three differences between equations (1) and (2). Two are obvious: consumption and genuine savings are expressed in per capita terms in equation (2), and the discounting factor on per capita consumption is reduced by the population growth rate. The use of per capita consumption and genuine savings is the only population adjustment Ferreira and Vincent (2005) make.

The third difference is the wealth-dilution effect of population growth, embodied in $g(t)$. Hamilton and Hartwick (2005) derive this effect as follows. Production in their model is based on the constant returns to scale technology $F = F(K, R, N)$, where R is the current extraction of an exhaustible resource. Production yields a homogeneous good, which can be either invested in produced

---

4. See Dixit, Hammond, and Hoel (1980) for the definition of a competitive economy. Hamilton and Hartwick (2005) analyze an optimal Dasgupta–Heal economy, but their result requires only that the economy be competitive.

capital or consumed. Per capita production is $f(k, r) \equiv F/N = F_K k + F_R r$, where $k = K/N$ and $r = R/N$. Per capita wealth is $w = k + F_R s$, where $s = S/N$ and S is the resource stock. Genuine per capita savings is given by $g = \dot{k} + F_R \dot{s}$, or

$$(3) \qquad g = \frac{\dot{K}}{N} - F_R r - \gamma w.$$

The first two terms on the right side are per capita investment in produced capital and per capita depletion of natural capital. The third term, $-\gamma w$, captures the wealth-dilution effect. It represents the reduction in per capita wealth that results from sharing existing wealth with the population increment.

The assumption of constant population growth in this model is restrictive. Population growth rates vary widely, even over short periods. The population growth rate in the sample used here, for example, had a within-country standard deviation equivalent to 10 percent of the mean. In supplemental appendix S.1 (available at http://wber.oxfordjournals.org) the following modified version of equation (2) is shown to hold when the population growth rate varies over time:

$$(4) \qquad \int_t^\infty \left(\frac{dc(v)}{dv} + \frac{d\gamma(v)}{dv} w(v) \right) e^{\int_t^v -(\rho(\tau)-\gamma(\tau))d\tau} dv = g(t).$$

The extra term on the left side accounts for the change in the population growth rate. The wealth-dilution term in $g(t)$ is now $-\gamma(t)w$, because γ is no longer constant. Proposition 6 in Asheim (2004) implies that the sign of genuine per capita savings in this model indicates whether social welfare is increasing or decreasing under two different measures of social welfare: the present value of total utility (population times per capita utility) or the present value of per capita utility.[5]

All of the preceding expressions are in continuous time for an infinite period. Empirical work requires expressions in discrete time over a finite interval. When the population growth rate is changing over time, the discrete-time approximation to equation (4) is given by the following equation:

$$(5) \qquad \sum_{v=t+1}^{t+T} \left(\frac{(C_{iv+1}/N_{iv+1} - C_{iv}/N_{iv}) + (\gamma_{iv+1} - \gamma_{iv})(W_{iv}/N_{iv})}{\Pi_{j=t+1}^{v}(1 + \rho_{ij} - \gamma_{ij})} \right) = g_{it}$$

where i denotes the country and ρ denotes the interest rate. The continuous-time result is interpreted as implying that savings at time t equals the present

5. Dasgupta (2001) shows that the ethics of maximizing per capita utility are questionable: the optimal allocation of a fixed quantity of a good across two groups of unequal size will lead to a smaller per capita allocation to the larger group if the utility function exhibits decreasing marginal returns.

value of future changes in consumption (beginning with period $t + 2$ minus period $t + 1$) plus the population growth-wealth interaction term. This can be written more compactly as follows:

$$(6) \qquad PV\Delta C_{it} + PV(\Delta \gamma_{it} w_{it}) = g_{it}.$$

The first term on the left side stands for the present value of future changes in per capita consumption; the second term stands for the present value of per capita wealth, weighted by the change in the population growth rate. This equation provides the basis for the empirical tests.

II. ECONOMETRIC MODELS

If population growth rates are constant over time, equation (6) implies that one should test the relation between per capita current genuine savings and per capita future consumption changes by estimating the following equation:

$$(7) \qquad PV\Delta C_{it} = \beta_0 + \beta_1 g_{it} + \varepsilon_{it}.$$

If population growth rates vary over time, one should instead estimate the following equation:

$$(8) \qquad PV\Delta C_{it} + PV(\Delta \gamma_{it} w_{it}) = \beta_0 + \beta_1 g_{it} + \varepsilon_{it}.$$

This equation differs from equation (7) in three ways: it includes $PV(\Delta \gamma_{it} w_{it})$ in the dependent variable, it includes the wealth-dilution term in g_{it}, and it includes adjustments for time-varying population growth rates in the discount rates in $PV\Delta C_{it}$, $PV(\Delta \gamma_{it} w_{it})$ and the wealth-dilution term. Both equations were estimated. Given that population growth rates varied over time in the countries in the sample, the expectation was that the results for equation (8) would be stronger than those for equation (7) in ways defined below.

Strictly interpreted, equation (6) implies the joint hypotheses $\beta_0 = 0$ and $\beta_1 = 1$: there is a one-to-one relation between genuine savings and consumption changes. A weaker hypothesis is simply $\beta_1 > 0$: genuine savings and consumption changes are positively correlated. The theory refers to a situation in which genuine savings include all changes in wealth. Any empirical estimates of genuine savings will inevitably be incomplete, which could bias the estimate of β_1 away from 1. This bias was expected to be smaller, and the estimates of β_1 therefore closer to 1, for more comprehensive savings measures. Equations (7) and (8) were estimated sequentially to test this hypothesis. Initially, g was set equal to gross national savings. It was then adjusted sequentially, for the depreciation of produced capital, the depletion of natural capital, and the dilution of produced and natural capital. As discussed in the next section, these

adjustments are crude. The resulting measurement error could weaken the convergence of the estimates of β_1 toward 1.

One might expect country fixed effects to be added to equations (7) and (8) (changing the intercepts from β_0 to β_{0i}) to control for omitted wealth components that are more or less constant over time. In fact, in the presence of time-invariant total omitted wealth X, equation (3) becomes

$$(9) \qquad \tilde{g} = \frac{\dot{K}}{N} - F_R r - \gamma\left(w + \frac{X}{N}\right) = g - X\frac{\gamma}{N}$$

where $\tilde{g}$ is true (unobserved) genuine savings.[6] As long as population N is changing over time, an ordinary fixed effect will not solve the problem of the omission of $X\gamma/N$ from the genuine savings variable in equations (7) and (8). The problem occurs even if the population growth rate is constant (and nonzero). To solve this problem one must include the ratio of the population growth rate to the total population as an additional explanatory variable. Equations (7) and (8) thus become

$$(10) \qquad PV\Delta C_{it} = \beta_0 + \beta_1 g_{it} + \beta_{2i}\frac{\gamma}{N_{it}} + \varepsilon_{it}$$

and

$$(11) \qquad PV\Delta C_{it} + PV(\Delta\gamma_{it} w_{it}) = \beta_0 + \beta_1 g_{it} + \beta_{2i}\frac{\gamma_{it}}{N_{it}} + \varepsilon_{it}.$$

From equation (9) one would expect β_{2i} to be negative if significant wealth were omitted and zero otherwise. The absolute value of this coefficient provides an estimate of X_i, the total omitted wealth for country i. These two equations and, for comparison, fixed-effects versions of equations (7) and (8) are estimated.[7]

III. Data and Estimation Issues

This section describes the data sources, the construction of key variables, and estimation issues.

Data

Data for constructing the variables in equations (7), (8), (10), and (11) were obtained from the *World Development Indicators* (World Bank various years).

---

6. The authors are indebted to a referee for pointing this out.

7. Fixed effects were used instead of random effects, because the sample is not random: it includes the population of developing countries. An *F*-test was used to test the null hypothesis that the fixed effects were equal to zero. The hypothesis was rejected in all models.

All monetary variables are expressed in constant 2000 U.S. dollars.[8] The final sample includes 64 non–OECD countries. Complete series for 1970–2003 were available for most countries, with missing values occurring mainly in 2003. Hence the panel is unbalanced but reasonably complete.

The time horizon, T, was set equal to 20 years in constructing the present value of future changes in per capita consumption. Although Ferreira and Vincent (2005) use 10 years in their benchmark model, they show that the econometric relation between genuine savings and their consumption measure improves when they extend the time horizon to 20 years. This improvement is not surprising: "green" accounting theory refers to an infinite time horizon. The 20-year time horizon reduced the sample in the econometric analysis to 1970–82: a test was run to determine whether genuine savings in year t was positively correlated with the actual present value of consumption changes during the period $t + 1$ to $t + 21$.

Each equation was estimated sequentially for four comprehensive savings measures. In increasing order of completeness, the four measures are as follows:

1. *Gross savings*. This measure implicitly includes both gross investment in produced capital within the country's borders and the current change in holdings of foreign assets.
2. *Net savings*. This measure equals gross savings minus depreciation of produced capital.
3. *Green savings*. This measure was constructed by subtracting estimates of the current depletion of subsoil assets and forest resources from net savings.
4. *Population-adjusted savings*. This measure equals green savings minus the wealth-dilution term.

Summary statistics for the 64 countries reveal that the means of $PV\Delta C$ and $PV\Delta C + PV(\Delta \gamma w)$ are the same order of magnitude as the savings measures, suggesting that an empirical relation between current savings and future consumption changes is plausible (table 1).[9] The means of the savings measures decrease sharply with the progressive adjustments for depreciation of produced capital (net savings), depletion of natural capital (green savings), and wealth dilution (population-adjusted savings). In line with Dasgupta's (2003) findings, the mean of population-adjusted savings is negative, suggesting that on average

---

8. A reviewer pointed out that exchange rates in 2000 may not be representative, because countries were still recovering from the effects of the Asian financial crisis. Choosing a different base year would alter the relative size of the different variables in the model across countries (because of different base year exchange rates). With country fixed effects, however, it is within-country variation over time rather than variation across countries that determines the econometric results. The results are therefore invariant to the choice of base year.

9. Supplemental appendix S.2 provides detail on data sources and the procedures followed in constructing the variables. Supplemental appendix S.3 provides detail on individual countries.

TABLE 1. Descriptive Statistics for Key Variables

Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
$PV\Delta C$	799	111.1	356.9	−2,201.3	1,414.9
$PV\Delta C + PV(\Delta \gamma w)$	790	50.2	379.3	−2,067.4	1,284.1
Gross savings	794	240.8	363.5	−86.8	2,516.3
Net savings	794	144.7	254.1	−221.3	2,124.0
Green savings	794	53.6	198.2	−1,591.3	1,489.1
Population-adjusted savings	793	−148.5	289.4	−2,321.1	1,028.2
Population growth rate (percent)	858	2.37	0.81	−0.13	4.37
Total population (millions)	858	41.5	136	0.120	1,010.0

Note: All variables except population growth rate and total population are expressed in per capita terms in 2000 U.S. dollars. $PV\Delta C$ and population-adjusted savings are computed using country-specific constant interest and population growth rates, as in equation (7); $PV\Delta C + PV(\Delta \gamma w)$ is computed using country-specific constant interest rates but time-varying population growth rates, as in equation (8). The sample of 64 countries covers the period 1970–82.

Source: Authors' analysis based on data from World Bank (various years).

the countries in the sample dissaved after accounting for the spreading of existing wealth across growing populations. The adjustment for wealth dilution also increases the variability of the savings measure, as indicated by the larger standard deviation for population-adjusted savings than for green savings.

Estimation Issues

There is a risk of endogeneity when population-adjusted savings is the savings measure. The estimates of natural capital in the wealth-dilution term are constructed using data on future resource rents. It is possible that a positive consumption shock in period $t + s$, which is reflected in $PV\Delta C_{it}$ on the left side of the regression equations, might induce a country to extract more resources to pay for the additional consumption. Because resource rents would also increase in period $t + s$, the dependent and explanatory variables would be simultaneously determined. This risk is greater when the adjustments for time-varying population growth rates are included, because future resource rents that appear in the current wealth-dilution term also appear in the $PV(\Delta \gamma_{it} w_{it})$ term on the left side of equations (8) and (11). Although this risk is reduced by the facts that current wealth was not used in constructing $PV(\Delta \gamma_{it} w_{it})$ (see equation (5)) and that future per capita wealth in $PV(\Delta \gamma_{it} w_{it})$ is weighted by the discount rate and the change in the population growth rate, it does not necessarily become negligible.

Equations that involved population-adjusted savings using instrumental variables were estimated using the generalized two-stage least squares fixed-effects estimator of Balestra and Varadharajan-Krishnakumar (1987) in

order to reduce the bias that could result. The set of instruments included lagged values of green savings, produced capital, the percentage of the population of working age, and the population growth rate, and a time trend. These variables were correlated with contemporaneous savings in the first-stage regressions (and were thus relevant instruments); lagging them ensured that they were exogenous. Standard errors were corrected in all models for serial correlation, and time dummy variables were included to control for unobserved global factors that affected consumption-savings decisions across all countries in a given year.

The risk of spurious correlation must be considered given that time-series data were used. Formal testing for stationarity and cointegration is not appropriate for the data, because time-series tests have little power in samples that cover as short a period as those used here (Banerjee 1999). There is ample evidence that consumption is nonstationary but cointegrated with income, so that savings is stationary (see, for example, Davidson and others 1978; Hamilton 1994). Although consumption is nonstationary, the dependent variable is constructed as the present value of future changes in per capita consumption. Examination of the autocorrelograms for these series confirms their stationarity. Ferreira and Vincent (2005) examine autocorrelograms for longer time series of similar variables, because the time horizon in their study was only 10 years. They, too, find that the series are stationary.

IV. Results

Estimates of β_1 were derived for the four model specifications and the four savings measures. Comparison of estimates in the last row (population-adjusted) of table 2 with those in the second from last row (green) indicates the impact of accounting for wealth dilution; comparison of estimates in column 3 with those in column 1 (and column 4 with column 2) indicates the impact of accounting for changes in the population growth rate through adjustment to the dependent variable; and comparison of the estimates in column 2 with those in column 1 (and column 4 with column 3) indicates the impact of controlling for time-invariant omitted wealth by including the ratio of the population growth rate to total population as an additional explanatory variable. For reference, column 5 shows results from a model with the dependent variable used by Ferreira and Vincent (2005) (that is, the difference between average future consumption and current consumption). That model includes no adjustments for population other than the variables expressed in per capita terms. A 20-year time horizon and country-specific (but time-invariant) interest rates were used instead of the 10-year horizon and the fixed 3.5 percent rate used by Ferreira and Vincent to make the results more directly comparable to those in the other columns.

Consider first the results in columns 1–4. The most striking result is the adjustment for natural resource depletion. The hypothesis $\beta_1 > 0$ is supported

TABLE 2. Estimates of β_1 for Four Model Specifications and Four Savings Measures

Variable	1 Equation (7)	2 Equation (10)	3 Equation (8)	4 Equation (11)	5 Ferreira and Vincent (2005)
Dependent variable	$PV\Delta C$	$PV\Delta C$	$PV\Delta C + PV(\Delta\gamma\nu)$	$PV\Delta C + PV(\Delta\gamma\nu)$	$\overline{C} - C$
Time dummy variables?	Yes	Yes	Yes	Yes	Yes
Country fixed effects?	Yes	No	Yes	No	Yes
Control for omitted wealth?	No	Yes	No	Yes	No
Savings measure					
Gross	−0.642*	−0.084	−0.764*	−0.106	−0.597**
	(0.365)	(0.255)	(0.415)	(0.258)	(0.268)
Net	−0.610*	−0.200	−0.729*	−0.234	−0.533**
	(0.364)	(0.316)	(0.412)	(0.324)	(0.274)
Green	0.425**	0.405**	0.558**	0.504**	0.801**
	(0.203)	(0.178)	(0.274)	(0.197)	(0.362)
Population-adjusted	0.413**	0.392**	0.560**	0.496***	0.788***
	(0.163)	(0.165)	(0.213)	(0.182)	(0.287)

***Significant at the 1 percent level; **significant at the 5 percent level; *significant at the 10 percent level.

Note: Numbers in parentheses are robust standard errors corrected for serial correlation.Two-stage least squares estimates are shown for population-adjusted savings. Fixed-effects estimates are shown for equations (7) and (8). Pooled ordinary least squares estimates are shown for equations (10) and (11). The sample of 64 countries and 788 observations covers the period 1970–82.

Source: Authors' analysis based on data from World Bank (various years).

only for the two savings measures that include this adjustment, green savings and population-adjusted savings. The estimates for gross savings and net savings are negative; although none is significantly different from zero at the 5 percent significance level, the estimates for equations (7) and (8) are significant at the 10 percent level and have very large negative values. The expected positive correlation between current savings and future consumption changes occurs only when the savings measure is expanded to include natural capital.

The sign change between net and green savings is consistent with the expectation that estimates of β_1 should be closer to 1 for savings measures that are more comprehensive. The estimates on green and population-adjusted savings remain significantly below 1 in all four equations, however. In addition, the estimates for population-adjusted savings are virtually identical to those for green savings rather than being closer to 1. One would expect the coefficient on population-adjusted savings to be biased upward if the instrumental variables did not completely purge that saving measure of endogeneity. In fact, the adjustment for wealth dilution does not substantially improve the empirical relation between current savings and future consumption changes.

The adjustment to the dependent variable increases the absolute value of the estimates, with the estimates for green and population-adjusted savings in columns 3 and 4 about one-fifth to one-third larger than the corresponding estimates in columns 1 and 2 and reaching values of 0.5 or more. Although these increases are not statistically significant, they are nonetheless substantial. There is thus some evidence that the adjustment moves the estimates closer to 1.

The control for omitted wealth affects only the coefficients on gross and net savings, which rise toward zero and become less significant (compare column 2 with column 1 and column 4 with column 3). It makes sense that omitted variables bias should be greater for these measures than for the more comprehensive ones. Although the coefficients on green and population-adjusted savings barely change, the estimates of β_{2i}, the country-specific coefficients on the ratio of population growth rate to total population, are significantly different from zero for many countries when they are the savings measures.

Consider the results for population-adjusted savings in equation (11), the model that includes all three population adjustments. Estimation of this model generated estimates of β_{2i} for 62 countries (estimates for China and India were not possible, because of the low estimates of the omitted-wealth control variables for these countries).[10] Fifty-one of the estimates—four-fifths of the total—were nonpositive (either significantly negative or not significantly different from zero), as theory predicts they should be. (Recall that a negative β_{2i} indicates a positive amount of omitted wealth.) Forty-five were negative, with half of those (22) significantly different from zero. There is thus econometric

10. For China and India, figures in hundredths (population growth) are divided by figures in billions (population). The resulting variables are close enough to zero to make the matrix of explanatory variables singular when they are included in it.

evidence that the estimates of produced and natural capital failed to account for significant amounts of time-invariant national wealth in about half of the countries.

The plausibility of the 51 nonpositive estimates of β_{2i} can be gauged by comparing them with the difference between the present value of future consumption flows, capitalized over a 20-year time horizon, and the sum of the values of produced and natural capital. The present value of consumption is a broad measure of a country's total wealth (Hamilton and Hartwick 2005; World Bank 2006); this procedure identifies the residual amount of the consumption stream that must be generated by some form of capital other than produced and natural capital. These variables need to be in total, not per capita, terms to be compared properly, because $-\beta_{2i}$ provides an estimate of total, not per capita, omitted wealth. A simple ordinary least squares regression was run with the 51 nonpositive estimates as the dependent variable and the corresponding residual amounts (means for 1970–82) as the explanatory variable. Although the fit was not tight ($R^2 = 0.101$), the slope parameter (-0.151) was significantly different from zero ($P = 0.023$). The magnitude of the slope parameter implies that the omitted wealth components determined by the estimation of equation (11) accounted for about one-sixth of the wealth omitted from the estimates of produced and natural capital.

Comparison of the results in column 5 with those in column 1 indicates that differences between the definition of the dependent variable in the study by Ferreira and Vincent (2005) and this study have the greatest impact when the savings measures incorporate adjustments for natural capital. The more restrictive theoretical basis of the model used by Ferreira and Vincent, which draws on Weitzman (1976) rather than Dasgupta (2001), causes it to exaggerate the correlation between current savings and future consumption: the coefficients on green and population-adjusted savings in column 5 are not significantly different from 1, whereas those in column 1 are.

V. Discussion

The econometric results indicate that there is a positive correlation in developing countries between current per capita savings and the present discounted value of changes in future per capita consumption when the measure of savings is expanded to incorporate natural resource wealth. This result holds when additional adjustments for wealth dilution linked to population growth and the effects of changing population growth rates are taken into account. Conventional savings measures are negatively related with the present value of changes in consumption; adding adjustments for natural capital reverses the relation and makes it positive, as theory predicts it should be. The improved performance of savings measures after making this adjustment is consistent with the results from the more restrictive model used by Ferreira and Vincent (2005). The results presented here imply that policymakers and economists

should interpret the net national savings rates for developing countries published in the *World Development Indicators* (World Bank various years) as signals of future consumption paths if and only if the rates include this adjustment for natural capital.

The three population-related adjustments evaluated lead to only minor improvements in the relation between current saving and changes in future consumption. The coefficients on green and population-adjusted savings increase by a third when the dependent variable is adjusted for changes in the population growth rate over time, but the increases are not statistically significant. Controlling for time-invariant omitted wealth by including the ratio of population growth rate to total population improves the relation in the sense that the coefficients on gross and net savings are no longer significantly less than zero; it does not affect the coefficients on green and population-adjusted savings. This suggests that the most important unobserved time-invariant components of total wealth are related to natural wealth and that the estimates of natural wealth, crude though they are, account for them surprisingly well.

The lack of significant impact of the adjustment for wealth dilution is surprising. The adjustment has a substantial impact on the savings estimates, as shown in table 1 and supplemental appendix S.3. Measurement error may well be to blame. The estimates of the stocks of produced and natural capital are crude, especially for produced capital (Pritchett 2000). Measurement error could also explain the lack of significance of the increases in the coefficients on green and population-adjusted savings when the dependent variable is adjusted for changes in the population growth rate over time, as this adjustment involves the capital stock estimates. Adjusting genuine savings for wealth dilution is justified theoretically, and the estimates presented here and those of Dasgupta (2003) and Hamilton and Atkinson (2006) indicate that it has a potentially large impact on estimates of genuine savings. Better estimates of capital stocks are needed, however, before it can confidently be stated that this adjustment significantly improves the performance of genuine savings as an indicator of future consumption changes.

The data provided in national accounts in developing countries are of questionable quality. The analysis suggests three priorities for producing better data: strengthening basic national accounts data, including data on gross savings and depreciation; updating and refining estimates of natural resource extraction and harvest costs, as well as constructing time series of resource lifetimes; and extending the coverage of natural resource data, particularly for agricultural soils, fisheries, and diamonds.

<div align="center">References</div>

Arrow, Kenneth J., Partha S. Dasgupta, and Karl-Göran Mäler. 2003. "The Genuine Savings Criterion and the Value of Population." *Economic Theory* 21(2):217–25.

Asheim, Geir B. 2004. "Green National Accounting with a Changing Population." *Economic Theory* 23(3):601–19.

Balestra, Pietro, and Jayalakshmi Varadharajan-Krishnakumar. 1987. "Full Information Estimations of a System of Simultaneous Equations with Error Component Structure." *Econometric Theory* 3(2):223–46.

Banerjee, Anindya. 1999. "Panel Data, Unit Roots, and Cointegration: An Overview" *Oxford Bulletin of Economics and Statistics* 6(0):607–29.

Burnside, Craig, and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):47–68.

Dasgupta, Partha S. 2001. *Human Well-being and the Natural Environment.* New York: Oxford University Press.

———. 2003. "Population, Poverty, and the Human Environment." In Karl-Göran Mäler, and Jeffrey R. Vincent eds., *Environmental Degradation and Institutional Responses,* vol. 1 *of Handbook of Environmental Economics.* Amsterdam: North-Holland.

Dasgupta, Partha S., and Geoffrey M. Heal. 1979. *Economic Theory and Exhaustible Resources.* Cambridge: Cambridge University Press.

Dasgupta, Partha S., and Karl-Göran Mäler. 2001. "Wealth as a Criterion for Sustainable Development." Discussion Paper 139. Beijer International Institute of Ecological Economics, Stockholm.

Davidson, James E. H., David F. Hendry, Frank Srba, and Stephen Yeo. 1978. "Econometric Modelling of the Aggregate Time-Series Relationship between Consumers' Expenditure and Income in the United Kingdom." *Economic Journal* 88(352):661–92.

Dixit, Avinash, Peter Hammond, and Michael Hoel. 1980. "On Hartwick's Rule for Regular Maximin Paths of Capital Accumulation and Resource Depletion." *Review of Economic Studies* 47(3): 551–56.

Easterly, W. 1999. "The Ghost of Financing Gap: Testing The Growth Model Used in the International Financial Institutions." *Journal of Development Economics* 60(2):423–38.

Easterly, W., and R. Levine. 2001. "It's Not Factor Accumulation: Stylized Facts and Growth Models." *World Bank Economic Review* 15(2):177–219.

Ferreira, Susana, and Jeffrey R. Vincent. 2005. "Genuine Savings: Leading Indicator of Sustainable Development?" *Economic Development and Cultural Change* 53(3):737–54.

Gelb, Alan, and Associates. 1988. *Oil Windfalls: Blessing or Curse?* New York: Oxford University Press.

Hamilton, James D. 1994. *Time Series Analysis.* Princeton, N.J.: Princeton University Press.

Hamilton, Kirk. 2002. "Sustaining per Capita Welfare with Growing Population: Theory and Measurement." *Paper presented at the Second World Congress of Environmental and Resource Economists*, June 27–30, Monterey, California.

Hamilton, Kirk, and Giles Atkinson. 2006. *Wealth, Welfare and Sustainability: Advances in Measuring Sustainable Development.* Cheltenham, UK: Edward Elgar.

Hamilton, Kirk, and Michael Clemens. 1999. "Genuine Savings Rates in Developing Countries." *World Bank Economic Review* 13(2):333–56.

Hamilton, Kirk, and John M. Hartwick. 2005. "Investing Exhaustible Resource Rents and the Path of Consumption." *Canadian Journal of Economics* 38(2):615–21.

Hartwick, John M. 1977. "Intergenerational Equity and the Investing of Rents from Exhaustible Resources." *American Economic Review* 67(5):972–74.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4):942–63.

Perrings, Charles, and Jeffrey R. Vincent eds. 2003. *Natural Resource Accounting and Economic Development: Theory and Practice.* Cheltenham, UK: Edward Elgar.

Pritchett, Lant. 2000. "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) Is *Not* Capital." *Journal of Economic Growth* 5(4):361–84.

Sachs, Jeffrey D., and Andrew M. Warner. 1995. *Natural Resource Abundance and Economic Growth.* NBER Working Paper 5398. Cambridge, Mass.: National Bureau of Economic Research.

Sala-I-Martin, X. 1997. "I Just Ran Two Million Regressions." *American Economic Review* 87(2): 178–83.

United Nations. 1993. *Integrated Environmental and Economic Accounting*. Studies in Methods, Series F No. 61. New York: United Nations.

Weitzman, Martin L. 1976. "On the Welfare Significance of National Product in a Dynamic Economy." *Quarterly Journal of Economics* 90(1):156–62.

———. 2003. *Income, Wealth, and the Maximum Principle*. Cambridge, Mass.: Harvard University Press.

World Bank. 1997. *Expanding the Measure of Wealth: Indicators of Sustainable Development*. ESD Studies and Monographs Series 17. World Bank, Washington, D.C.

———. 2006. *Where Is the Wealth of Nations? Measuring Capital for the 21st Century*. Washington, D.C.: World Bank.

World Bank. Various years. *World Development Indicators*. Washington, D.C.: World Bank.

Comparison of Net Benefits of Incentive-Based and Command and Control Environmental Regulations: The Case of Santiago, Chile

Raúl O'Ryan and José Miguel Sánchez

The ambient permit system proposed in the literature for cost-effective pollution reduction is difficult to implement and may result in lower net benefits than using another instrument. The article develops a model for comparing the environmental net benefits of three policy instruments for Santiago, Chile, when the policy problem is to meet a given ambient quality standard. Two market-based instruments—the ambient permit system and a simpler emission permit system—are examined along with an emission standard, a command and control instrument usually favored by regulators. Both emission permit system and emission standard are costlier than the ambient permit system, sometimes in large part because they improve ambient emission concentrations beyond the required target in much of the city, but the ambient permit system requires a lower degree of control to comply with the standard. The somewhat costlier emission permit system and emission standard provide much higher net benefits than the ambient permit system when the health benefits of their "excessive" air quality improvements are taken into account. These benefits are different from the fact that an ambient permit system is administratively costlier to implement. JEL code: Q25

Theory suggests that when a regulator wants to obtain a cost-effective (or minimum cost) solution for improving environmental quality in a given airshed or watershed, tradable permits or pollution taxes are the appropriate instrument. For the simple case of a uniformly distributed pollutant, the solution is a unique emission tax or an emission permit system that allows one-for-one emission trades among sources in different locations. This simplifies implementation, requiring only the total allocated emission permits that allow reaching

the required air quality target. A unique price for each emission permit would result independent of the location of the emitting source.

However, when the pollutant is not uniformly distributed—as is the case for particulates and many other local pollutants—the optimal system requires that pollution permits be issued not for the amount of emissions at the source, as in an emission permit system, but for the deposition at each receptor point in the airshed, through an ambient permit system. The required overall air quality must be obtained when measured by depositions at each receptor point. As a result, different prices for each unit of concentration reduction emerge at each receptor location. The design and implementation of the instrument become quite complex, requiring multiple interactions among sources that are not based on one-for-one emission trades. To ease implementation, a simple—but not optimal—approximation is to define different trading zones within which sources can trade on a one-for-one basis. Any trading between zones, if allowed at all, must be based on transfer coefficients that consider how pollutants disperse. An example is the Regional Clean Air Management Program in Southern California (RECLAIM), which defines two different zones. Emission permits have been issued for each zone, but trading between these zones is not allowed.[1]

Simulation studies for both developed and developing economies of the static efficiency gains from the use of incentive-based instruments, in particular of an ambient permit system, rather than of command and control instruments or an emission permit system, conclude that the cost reductions produced by an ambient permit system are significant in some cases and not very large in others (Atkinson and Lewis 1974; Hahn and Noll 1982; Seskin, Anderson, and Reid 1983; Krupnick 1986; McGartland and Oates 1985; Spofford and Paulsen 1988; Portney 1990; O'Ryan 1996).[2] An important caveat, however, is that ambient concentrations in many receptor locations are higher under the ambient permit system than under the emission permit system or command and control instruments, while still meeting the pollution reduction target. As a result, the magnitude of the cost reductions from an ambient permit system stems both from the efficiency gains related to equalizing the pollutant reduction marginal costs or cost per unit of pollutant concentration at the receptor location—a true efficiency gain—and from the lower degree of overall required pollution control (Tietenberg 1985).[3]

1. This assumes that emissions from one zone do not affect the other zone, which is a simplification that allows implementing the system (www.aqmd.gov/reclaim).

2. This ranking of instruments based on cost-effectiveness assumes no uncertainty of benefits and costs, perfectly monitored emissions, complete enforcement, and no asymmetric information. The magnitude of the efficiency gains depends on numerous factors, including dispersion characteristics of the pollutant, relative size and abatement costs of sources, and number of emitting sources (see Tietenberg 1985).

3. This result is true for a unique or dominant receptor location under the ambient permit system. Otherwise, with many receptors the marginal cost of emission reduction for each source is equal to the sum across all receptors of the shadow price of the pollutant concentration at each receptor times the impact of the source's emissions on that receptor.

If no value is assigned to the higher overall level of pollution reduction achieved by the emission permit system and the command and control instruments, these instruments will be considered less desirable from a social perspective than the ambient permit system. The problem is that cost-effective approaches implicitly assign a shadow price of zero to improvements that exceed the target. If, "however, reduced concentrations below the level of the standards bring with them improvements in health or the environment, command and control instruments approaches will produce greater benefits than incentive based approaches" (Oates, Portney, and McGartland 1989, p. 1233). Consequently, comparision of instruments without correcting for these benefits is unfair and may be misleading. Two approaches can be used to overcome this problem. One is to eliminate the lower degree of required control component by requiring that all instruments comply with the same air quality standards in all receptor locations, as is done by O'Ryan (1996). The comparison in this case is still in a cost-effectiveness framework. A second approach is to determine the net benefits for each instrument, allowing for a more complete comparison using a cost–benefit analysis.

This article compares the net benefits of an ambient permit system, an emission permit system, and an emission standard, a command and control instrument, in Santiago, Chile, using cost–benefit analysis. Its contribution to the literature is to point out that regulatory schemes that are simpler to implement than the ambient permit system can also yield higher *net* benefits.[4] Which pollution control system yields the highest net benefits is an empirical question. The authors are not aware of any of the study that answers this question in a developing economy, and there are few studies that address the question in developed economies. In a comparison of a uniform standard and an ambient permit system in Baltimore, Md, Oates, Portney, and McGartland (1989) conclude that the resulting net benefits of the uniform standard are only slightly lower (US$6 million).

In developing economies, where few pollution control efforts have been undertaken, abatement costs are usually not very high and the health benefits of improving air quality can be significant. As a result, the net benefits of improving air quality may favor the use of the emission permit system and command and control instruments. The health benefits of improved air quality under these instruments will outweigh their relative cost disadvantage compared with an ambient permit system. To examine this hypothesis, Santiago's emission permit system, the Sistema de Compensaciones, is compared with an ambient permit system and an effluent concentration standard (a command and control instrument).

The next section presents an overview of the air pollution problem in Santiago. Section II addresses the compliance costs of reaching given air quality targets using market-based instruments and command and control instruments.

---

4. The additional benefits of reduced transaction costs from a simpler system are not evaluated in this analysis.

A linear programming model is used to establish the total costs of achieving a desired air quality standard for each instrument. The following sections present the population-based health benefits associated with each instrument, and then compare the net benefits of applying the ambient permit system and the two second-best policies. The last section presents the main policy conclusions and suggests future research lines.

I. SANTIAGO'S AIR POLLUTION PROBLEM

Santiago, Chile, like many large cities in developing economies, suffers from severe air pollution. During winter, concentrations of particulate matter of ten micrometers in diameter (PM10) constantly exceed the established ambient standards. An extensive international epidemiologic literature reports illness and premature deaths due to exposure to airborne particulate matters. Studies have found that 5.2 million inhabitants were affected in the city because of these high levels of PM10 pollution.[5] The city's policy-makers have been struggling since the early 1990s to improve air quality, implementing Decontamination Plans in 1990 and 1997 (for details, see ÓRyan and Larraguibel 2000).

For particulate matter emissions from large stationary sources—industrial boilers and processes, and large residential and commercial heaters—a relatively stringent effluent concentration standard was established in 1992. To introduce flexibility, an emission permit system for particulates was introduced in March 1992, under which existing pollution sources can sell or a buy permits, depending on whether their estimated emissions are below or above their grandfathered permits. The system does not consider emission banking. Permits are expressed in kilograms per day and are traded at a one-for-one ratio. All trades require approval by the regulatory agency. Annual compliance inspections reconcile emissions with the number of permits held by each source. A source that fails to cover its emissions with permits incurs heavy penalties, including the possibility of a temporary shutdown.[6] While an emission permit system was known to be suboptimal from a cost-effectiveness perspective, a more complicated ambient permit system was rejected because the required models for implementing it were not available and trades were believed to be unnecessarily complicated.[7] However, there was no explicit evaluation of this decision or of its impacts.

5. Ostro and others (1996) found a strong association between PM10 and daily mortality rates among Santiago residents after controlling for several potential other factors. Ostro and others (1999) found a statistically significant association between PM10 and medical visits for lower respiratory tract illness in children.

6. For an analysis of the emission trading Program see Montero, Sánchez, and Katz (2002) and O'Ryan (2002).

7. Ambient permit systems are difficult to implement because of information and model requirements. In particular, implementing such a system would require knowing the contribution to concentrations at different receptor locations of each of the sources included in the system. Additionally, the acceptability of the instrument by sources is negatively affected since two otherwise similar sources would face different trading rules simply because they are in different locations.

FɪɢURE 1. Baseline Emissions of Particulate Matter (PM10) in Santiago, Chile, in 1998
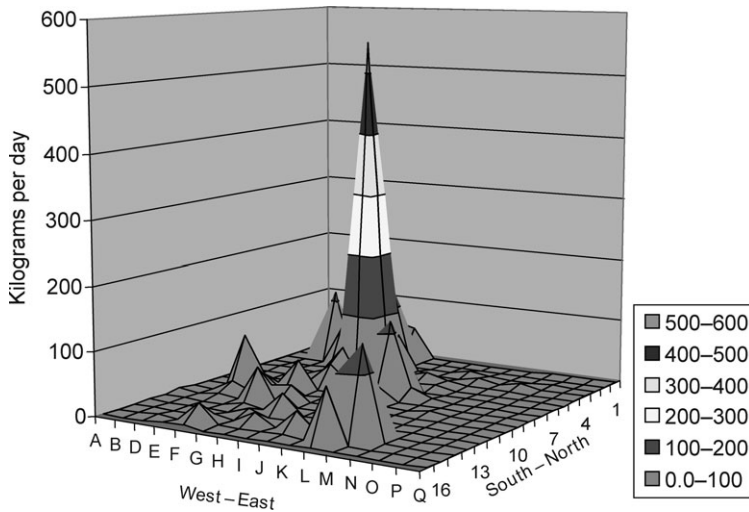
Source: Authors' analysis based on data from CONAMA (2000).

To examine the spatial configuration of emissions from fixed-point sources in Santiago, the city can be divided into a 34×34 kilometer grids of 289 (2×2 kilometers) cells that contain the relevant sources of air pollution in Santiago, as well as most of the exposed population. This area of the city contains 1,098 fixed-point sources. Total PM10 emissions in the city from these sources reached 2.55 tons a day in 1998 (CONAMA 2000).[8] Figure 1 presents average daily PM10 emissions from each cell in the grid, for that year. Point sources are clustered in a few zones. The cell with highest emissions is in the northwestern part of the city and emits 594 kilograms per day, 23 percent of the total PM10 emitted by point sources in the city.[9] Of the 289 cells of the grid, only 7 are highly polluting (emit more than three percent of total emissions) and the 14 most polluting cells account for 65 percent of total emissions. These emissions spread over the rest of the city, affecting air quality in each cell.

8. Even though this value seems low, together with emissions by mobile sources (roughly double those by fixed point sources) and the serious thermal inversion problem in Santiago, air quality concentrations exceed the standards discussed previously.

9. This cell includes a power plant with both natural gas- and diesel-powered generators, the largest single emitting source in the city and the only power plant in Santiago. Despite the magnitude of the source, it is included in this analysis since no strategic behavior should be observed. Additionally, there does not seem to be any important incentive for the power plant to hoard permits since it is the only power plant in the city and there is no possibility that another one will be authorized to operate in the city. As a result, the plant has been included in the current tradable permit program.

II. Costs of Improving Air Quality under Alternative Regulatory Instruments

The general setting is that there are n sources of pollution spatially distributed in the city. Air quality is measured at K receptor points, and a ton of pollution emitted by the firm i has a different impact on air quality at receptor k than a ton emitted by the firm j. Generally, the regulator wants to reach a vector $Q^* = (q_1^* \ldots q_k^* \ldots q_K^*)$ of maximum permitted ambient pollution concentrations. As is usual in policy formulation, the same standard is imposed on all locations—for all k, $q_k^* = q^*$.[10]

The Policy Instruments

Three policies are evaluated for Santiago: two market-based instruments (ambient permit system and emission permit system) and one command and control instrument (an effluent concentration or emission standard).

For the spatially differentiated ambient permit system, it is assumed that permits, defined in units of concentration at each receptor, are distributed to achieve the desired unique air quality goal at each receptor. Trades are not undertaken on a one-for-one emissions basis. This is the traditional cost-effective benchmark policy.

Under the marketable emission permits system, total allowable emissions are established for fixed sources in the airshed. Permits in an amount equal to these emissions are distributed to polluters, who can then buy and sell them on a one-for-one emissions basis. The number of permits each source buys or sells is the result of the cost minimization of compliance costs by each source.

Under the uniform effluent concentration standard, all point sources are required to emit at concentrations lower or equal to a unique stack concentration standard. Total compliance costs are then the sum of the compliance costs for each source needed to at least meet the stack concentration standard.

Conceptual Framework for Comparing the Compliance Costs of Each Instrument

To compare policy instruments, it is necessary to impose the condition that they reach the desired air quality goal at all receptor locations. However, different policy instruments typically result in different concentrations at each receptor location. To stay as close to reality as possible, it is usually accepted that the target has been reached when at least one receptor location has a concentration of q^*—the binding receptor—and the others are the same or lower. For this reason, to compare compliance costs, the command and control scheme and emission permit system will be defined so as to achieve the same air quality standard at their binding receptors as the ambient permit system.

---

10. Primary standards that are established to protect health are usually required by law to be the same everywhere in the country.

Formally, the cost-effective ambient permit system instrument is used to obtain the least cost solution to achieve a maximum permitted ambient concentration of q^* at K receptor points in the city. This can be expressed as the following problem (Montgomery 1972):

$$(1) \qquad \min_{\{e_i\}} \sum_{i=1}^{n} C_i(e_i) \qquad \text{s.t.} Q^* \geq ED; \quad E \geq 0$$

where D is an $n \times K$ dispersion matrix (d_{ik} is the impact of a ton of pollution emitted by source i on concentrations at receptor k), E is a $1 \times n$ vector of emissions by n firms in the city, $C_i(e_i)$ is the cost to firm i of emitting e_i, and Q^* the K-component vector of target concentrations.

Under the ambient permit system, there are K types of permits (one for each receptor) that give firms the right to increase ambient concentrations at each receptor. It is well known that as long as permits totaling q^* are given out for each receptor and the K sets of permits are traded in competitive markets, the ambient permit system minimizes the cost of achieving Q^* (Montgomery 1972).

Under an emission permit system, permits equal to E^* tons of emissions are distributed to the n firms, and firms trade emission permits one-for-one basis. If the permit market is competitive, the emission permit system is a solution to the following problem:

$$(2) \qquad \min_{\{e_i\}} \sum_{i=1}^{n} C_i(e_i) \quad \text{s.t.} E^* \geq \sum_i e_i.$$

The emission vector that solves problem (2), E', implies a vector $Q' = E'D$. Plotting total costs against the largest element of Q' (q_{max}) gives the cost of achieving $q^* = q_{max}$ under the emission permit system. If problem (1) has been solved for a given q^*, then E^* has to be varied until the largest elements of Q' coincide with q^*.

Under the emission standard, each source's emissions depend on the size of the source (gas flow) and hours of operation per day. The resulting emissions vector after the standard is applied, E^c, will imply a vector of ambient concentrations, $Q^c = E^c D$. The standard that would make the largest elements of Q^c coincide with q^* is the standard to be compared with the ambient permit system and emission permit system.

Empirical Estimation of Abatement Costs and Concentrations

To estimate the abatement costs under each instrument required to reach the concentration target, it is necessary to know both the abatement cost function, $C_i(e_i)$, and the matrix D relating the vector of emissions to concentrations. The cost of abatement for each source depends on the applicable control

alternatives. On the basis of the literature (Bretschneider and Kurfurst 1987; Vatavuk 1990; Aranda 1996; Bravo 2000) and expert opinion, two categories of abatement alternatives were identified for the main processes in Santiago: collection devices such as cyclones, multicyclones, bag filters, and wet scrubbers, and for some sources, a change of fuel. Each control option was also assigned an abatement efficiency value.[11]

The costs of collection devices were estimated based on estimates of the net discounted cash flow of total capital investments and net annual operating costs incurred each year over the useful life of the equipment. The present value of switching to cleaner fuels was stimated based on estimates of the cost of transformation and the cost differential associated with using a different fuel. Control devices of different sizes were costed. Analytical cost relations were established for each control alternative (see supplemental Appendix S.1). For each option, the minimum cost required to reach the required standard was used. However, the lack of flexibility may impose a high cost on some point sources, resulting in overall costs that are higher than that under an ambient permit system.

To relate concentrations to emissions, the natural systems model is represented by the environmental "transfer" coefficient, d_{ik}, of the dispersion matrix D. A tool that simulates the dispersion process for Santiago was used to obtain these coefficients, based on a multiple cell model that is solved using mass conservation equations.[12] The wind fields had to be averaged over the day, and meteorological conditions reflecting episode conditions (days in which the air quality standard is exceeded) had to be selected.[13] A total of 28 episode days were used, and the corresponding transfer coefficients were averaged. As a result, the transfer coefficients reflect the impact of a unit of emissions on concentration levels in each cell of the grid for adverse meteorological conditions.[14]

The Simulation Model

Each policy instrument is defined using different policy targets: air quality at each receptor location for the ambient permit system, total emissions for the emission permit system, and a uniform stack concentration target for the emission standard. To compare the compliance costs of these policy instruments,

11. These are presented in supplemental Appendix S.1. For the model each source was assigned only the options applicable to it. It is not assumed that existing abatement technologies are dismantled when there is a fuel switch, and the conservative assumption is made that no extra reductions are obtained when control equipment exists. Mixtures of more abatement and fuel switching were not considered, based on expert opinion that suggested that the technical options that were economically feasible are those considered in table S.1 of supplemental Appendix S.1.

12. The coefficients are derived in Muñoz (1993).

13. The results are presented in supplemental Appendix S.2, "Transfer Coefficients," and discussed in detail in Muñoz (1993).

14. These concentrations do not include secondary particulate matter generated by nitrogen dioxide and sulfur dioxide emissions, as there are no models available for this for Santiago. However, efforts are being initiated to estimate the impact of these emissions in the city.

TABLE 1. Level of Application of Emission Permit System and Emission Standard to Reach Target Concentration

Ambient permit system concentration target (micrograms per cubic meter)	Emission permit system (kilograms per day)	Required emission standard (micrograms per cubic meter)
29.1	1,832	90.0
27.8	1,556	37.0
26.5	1,324	13.0
25.2	1,063	8.0
22.9	955	2.3

Source: Authors' analysis based on data from Bravo (2000).

the ambient concentration reached in the binding receptor under each instrument is used as a common target.[15] Specifically, different reductions in pollution concentrations relative to the binding receptor are used as targets for each policy instrument. Table 1 presents the level of application of each instrument required to reach the same concentration target as defined by the ambient permit system.

For the simulation exercise, the problem for each instrument is specified as a linear programming model with binary variables. For the ambient permit system, the model considers the objective function and environmental constraint—a concentration target at each receptor location—presented above. The solution determines which control option was used by each of the sources considered to comply at minimum cost. Summing individual compliance costs over all sources results in total compliance costs.

To simulate the other two policy instruments, only the environmental constraint has to be modified. The emission permit system is similar to the ambient permit system, but must comply with an overall emission target—total emissions must be lower than a predetermined target. Under the emission standard, each emitting source must comply with a target effluent concentration.[16]

Once each source has made its cost-minimizing decision, the resulting emissions in each cell are added to obtain the aggregate emissions on an average episode day. These emissions are then transformed into concentrations at each point of the grid using transfer coefficients, making it possible to compare the average daily concentration reductions in episode days under each instrument and the costs of reaching these reductions.

Specifically, the compliance costs under each policy instrument are estimated using the following model that considers an objective function and two constraints: a technological constraint common to all instruments and an environmental constraint specific to each one. The model considers a total of 1,098

15. This guarantees that in all other receptor locations, air quality is the same or better.
16. The result of global minimization of costs is identical in the case of the emission standard to the individual cost minimization problem and for this reason the same model can be applied. In both cases, the source will choose the unique technology that enables complying at minimum cost.

emission sources and a maximum of 10 abatement options for each source. Each of the 289 cells of the Santiago grid is a receptor location. The simulation is carried out for a linear programming model with a binary variable. The model is formulated using the GAMS software and results are obtained with CPLEX solver.

The general model is as follows:

$$\text{Objective function}: \quad \text{Min} \sum_{i=1}^{1,089} \sum_{t=1}^{10} CT_{i,t} X_{i,t}$$

where $CT_{i,t}$ is the annual cost of applying technology t to source I, and $X_{i,\,t}$ is the binary variable that determines whether technology t is applied to source i.

$$\text{Technological constraint}: \quad \sum_{t=1}^{10} X_{i,t} = 1 \quad \forall i = 1, \ldots, 1,098$$

Environmental constraint: specific to each instrument.

For the *ambient permit system*, the specific environmental constraints are

$$\sum_{i=1}^{1,098} \sum_{t=1}^{10} \sum_{k=1}^{289} E_i \text{HO}_i \alpha_{k',k} \text{UB}_{i,k'} (1 - \text{EFF}_{i,t}) X_{i,t}$$

$$\leq Q_k \quad \forall k = 1, \ldots, 289$$

where $\alpha_{k',k}$ is the transfer coefficient representing the effect emissions in zone k have on concentrations at location k', HO_i is the hours of operation of source i per day, $\text{UB}_{i,k}$ is a dummy variable taking a value of one if source i is located in cell k and zero otherwise, $\text{EFF}_{i,t}$ is the efficiency in emission reductions of technology t being applied to source i, and Q_k is the air quality target for location k (and for all receptor locations).

For the *emission permit system*, the specific environmental constraints are

$$\sum_{i=1}^{1,098} \sum_{t=1}^{10} E_i \text{HO}_i (1 - \text{EFF}_{i,t}) X_{i,t} \leq E$$

where E_i is the total emission of source i (in kilograms per hour), and E is the aggregate emission target.

For the *emission standard*, the specific environmental constraints are

$$\sum_{t=1}^{10} \Omega_i(1 - \text{EFF}_{i,t})X_{i,t} \leq \Omega \quad \forall i = 1,\ldots,1,098$$

where Ω_i is the effluent concentration level (in milligrams per cubic meter) of source i obtained as emissions divided by flow, and Ω is the effluent concentration standard (milligrams per cubic meter).

For programming purposes, the targets defined by each instrument are set through Q_k, E, and Ω. Targets implying concentrations at each receptor location ranging from 29.1 to 22.9 micrograms per cubic meter were evaluated. Lower targets are not possible in the worst receptor location without reducing activity at some sources or closing them down, options not considered in this study.[17]

Compliance Costs under Alternative Policies

The model yields both costs and concentrations per cell of the grid. Before presenting the results, it is necessary to make a correction to current emissions. Natural gas had only been introduced in 1998 in Santiago, and many sources that could profitably switch to this fuel had not done so yet. To eliminate any distortionary effect on source decisions, it is assumed that all sources that can profit from switching to natural gas do so at the start of the program. Consequently, only the costs and benefits from additional reductions are considered.

As expected, the ambient permit system is clearly the most cost-effective instrument. The maximum reduction can be obtained with an annual cost for participating sources of almost US$20 million, less than half the cost for the other policy instruments.

The annualized compliance costs and resulting reductions in concentrations for each policy for fixed-point sources in Santiago are presented in table 2.[18] The reduction in compliance costs for the ambient permit system is considerable. The emission permit system is particularly expensive when small reductions are required, for example, for a 29.1 micrograms per cubic meter concentration, the target emission permit system costs are 45 times those of ambient permit system. However, over the range of reduction options for concentration targets lower than 28.7 micrograms per cubic meter, the costs for similar reductions are only 3–20 times higher with an emission permit system than with an ambient permit system. Compliance costs under the emission standard are even more expensive, between 3 and 35 times higher than the ambient permit system for most of the reduction range. The emission standard

17. Even with the best available control technology, concentrations cannot be reduced more for the thermoelectric megasource.

18. Concentrations consider only fixed-point sources. When mobile sources are included, concentrations increase about 50 percent.

TABLE 2. Annualized Compliance Costs for Different Concentration Reductions Relative to the Worst Cell (in million U.S. dollars)

Concentration target (micrograms per cubic meter)	Ambient permit system	Emission permit system	Emission standard
29.3[a]	—	—	—
29.1	0.01	0.45	0.07
28.7	0.06	0.96	0.56
28.2	0.1	2	2
27.8	0.2	2	3
27.4	0.2	2	7
26.9	0.3	3	9
26.5	0.5	4	9
26.1	0.7	6	10
25.6	1.1	7	12
25.2	2	12	17
24.8	2	14	21
24.3	3	14	30
23.9	6	24	31
23.4	10	27	34
23.0	13	45	38
22.9	19	51	48

[a]Current concentration level in the binding receptor.

Source: Authors' analysis based on data from Bravo (2000).

is also more expensive than the emission permit system for most of the reduction range except for extremely small or large reductions.

For very low and very high values of the target, the emission permit system is more costly than the command and control instrument. This is not an unexpected result, because the emission permit system is not cost-effective and so can be more costly than a command and control instrument for some specific reduction goals. This type of result, documented in other studies (Tietenberg 1985), depends on the target, the relative compliance cost functions, and the relative size and number of sources (O'Ryan 2006).

Air Quality at Each Receptor Location and Population–Weighted Concentrations

A key result is that concentration reductions are different in each receptor location—for the same target—under each policy instrument. This shows that part of the cost reductions from the ambient permit system is not related to efficiency gains, but is because of the lower degree of required control. Since health effects are related both to pollutant concentrations and to the size of the exposed population in each cell, estimation of pollution exposure under each instrument for each target requires estimation of population-weighted concentrations in each cell and summation of them across all cells. For the four receptors with the highest pollutant concentrations, population-weighted

FIGURE 2. Population-Weighted Pollution Concentrations as a Function of the Target Concentration for Selected Receptor Locations, by Pollution Control Instrument
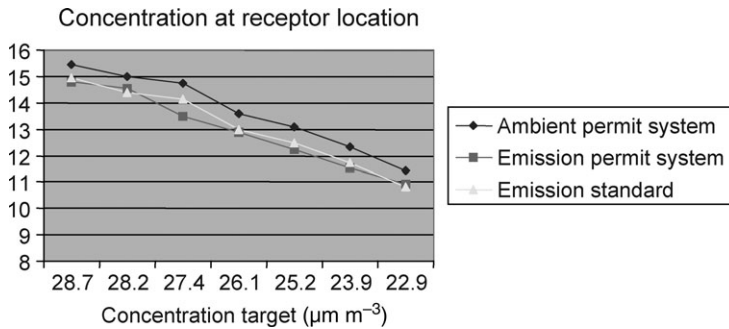
Source: Authors' analysis based on data from Chilean National Institute of Statistics.

concentrations are higher under the ambient permit system than under the other two instruments, reflecting the lower degree of abatement required under the ambient permit system (Figure 2).

As a consequence, the health benefits from applying each instrument will be different. In particular, the emission permit system, which imposes larger improvements in population–weighted air quality, would be expected to result in higher health benefits than the other instruments. The following section estimates these health benefits.

III. THE HEALTH BENEFITS OF IMPROVED AIR QUALITY

The damage function approach, frequently used in environmental cost–benefit analysis, is used to estimate the health-related benefits of improved air quality (see, for example, Ostro 1996; Environment Canada 1997; European Commission 1998; USEPA 2000). The methodology involves four steps. First, the change in emissions is determined for each policy instrument. Second, the resulting impact on concentrations is estimated. Third, the effects of the reductions in pollutant concentration on various health outcomes are estimated. The changes in health outcomes are quantified using dose–response functions for a set of health effects for which there are well-established statistical relations in the environmental epidemiologic literature. These dose–response functions are applied to the exposed population to determine the population-weighted health effects. Forth, these health effects are valued in monetary units and summed over the different effects, the individuals exposed, and time.

Dose–Response Functions

The dose–response functions used were obtained from the environmental epidemiologic literature. For mortality the dose–response function used was

TABLE 3. Dose–Response Coefficients for Human Health Effects from Particulate Matter (PM10) and Unit Costs of Health Effects

Source	Health effect category	Concentration response parameter	Unit cost in 1998 (in U.S. dollars)[a]
Ostro and others (1996)	Acute mortality (ICD 460) (percent increase per one microgram per cubic meter change in annual average PM10 concentration)	0.1%	700,000
Burnett and others (1995)	Hospital admissions for respiratory illness (ICD 480–86) (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	6.73×10^{-4}	1,600
Burnett and others (1995)	Hospital admissions for cardiac illness (ICD 410, 413, 427, and 428) (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	6.4×10^{-4}	3,500
	Emergency room visits for respiratory illness (a parameter that relates total emergency room visits to the total number of hospital admissions in 1995 is used instead of a dose–response function. Emergency room visits were six times the number of hospital admissions)		80
Ostro (1990)	Restricted activity days, adult population (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	0.0168	16
Dockery and others (1996)	Lower respiratory illness in children (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	0.0011	170
Abbey and others (1993)	Chronic bronchitis, population over age 25 (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	6.1×10^{-5}	140,000
Krupnick, Harrington, and Ostro (1990)	Acute respiratory symptoms (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	0.1679	9
Whittemore and Korn (1980)	Asthma attacks, among asthmatic population (individual risk factor per one microgram per cubic meter change in annual average PM10 concentration)	0.059	170

Note: ICD is international classification of diseases.

[a]Numbers have been rounded up to avoid giving a sense of false precision.

Source: Authors' analyses based on data sources shown in table 3 and Holz and Sánchez (2000) for unit costs.

estimated for Santiago (Ostro and others 1996). For the other health effects, the functions were obtained from epidemiologic studies estimated for other populations, although the selection criteria used followed Ostro and others (1996).

A large body of literature relates adverse health effects with ambient concentrations of PM10. The concentration response parameters reported in table 3 are typically obtained as the mean value reported by epidemiologic studies selected as providing the most reliable results. Most of the studies estimated linear and log-linear models, which imply a continuum of effects even at low concentration levels. This is justified by the fact that studies have failed to find thresholds for effects associated with particulate matter. In addition, many recent epidemiologic studies have found an association between particulate matter and health effects throughout the whole range of concentrations, even for levels under the primary air quality standards of the U.S. Environmental Protection Agency. There is also little evidence that the slopes of the dose–response functions diminish significantly at lower concentrations (Ostro 1996, p. 4). As a consequence, the functions used in this study assume that the slope of the dose–response function is the same regardless of the concentration level.[19]

Finally, since these dose–response functions consider average annual PM10 concentrations, the average daily episode concentrations estimated previously had to be converted to annual values. For this, the factors estimated by Jorquera (2002a, b) were used, which represent average dispersion conditions for each month in Santiago at four different receptor locations. Since his results do not vary much by location, the average results for the four locations were used. Average dispersion conditions in the worst winter month (June) are more than four times as bad as in the best month (January) (table 4). To estimate the average annual reduction in PM10 concentrations, these factors are assumed to represent the average dispersion conditions for each month relative to the episode conditions (which has a factor of 1). Consequently, the average is a weighted average, where the weights are the number of days in the month relative to the total annual number of days times the relative dispersion factor.

Monetary Valuation of Health Effects

For valuing a reduction in mortality from lowering pollution levels, the concept of the value of a statistical life is used, estimated from willingness to pay studies. The value of a statistical life is the average of 13 studies selected

---

19. See also European Commission (1998, vol. 7, pp. 133–134): "for many of these pollutants, there is clearly a threshold at the individual level, in the sense that most people are not realistically at risk of severe acute health effects at current background levels of air pollution. There is however no good evidence of a threshold at the population level; i.e., it appears that, for a large population even at low background concentrations, some vulnerable people are exposed some of the time to concentrations which do have an adverse effect. This understanding first grew in the context of ambient particles, where the no threshold concept is now well established as a basis for understanding and for policy."

TABLE 4. Relative Dispersion Factors for Each Month

Month	Relative dispersion factor	Number of days
January	0.239	31
February	0.279	28
March	0.366	31
April	0.579	30
May	0.805	31
June	1.000	30
July	0.859	31
August	0.646	31
September	0.431	30
October	0.279	31
November	0.251	30
December	0.251	31

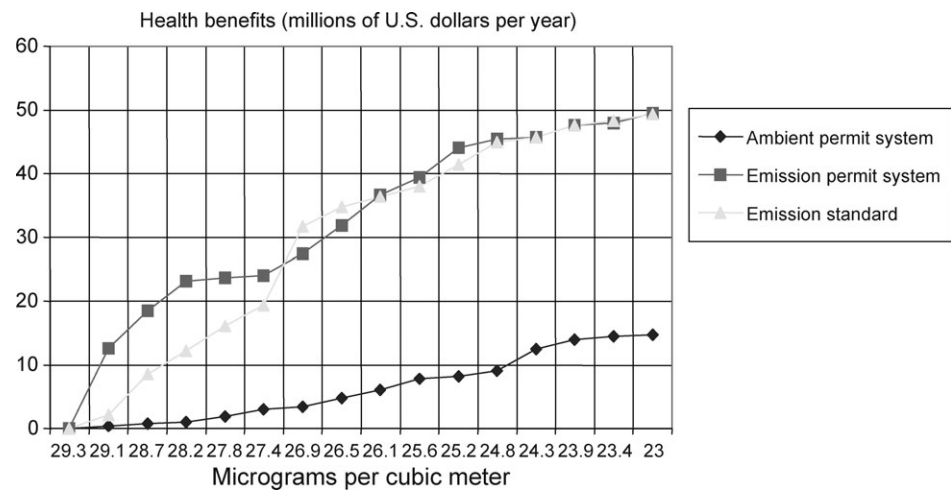Source: Authors' analysis based on data from Jorquera (2002a, b).

by the U.S. Environmental Protection Agency that report the lowest values. The values were deflated using the gross national product (GNP) per capita in purchasing power parity terms estimated for 1999 by the World Bank to account for differences in GNP per capita between the United States and Chile. For reductions in illnesses, no willingness to pay studies are available; therefore, the cost of illness estimates from Holz and Sánchez (2000) was used.[20] This approach considers direct treatment costs plus lost income as a measure of productivity loss during illness. This method is simple, but it has several limitations. It is a lower bound estimate of the true willingness to pay for reductions in illness because it does not consider other costs, such as pain and inconvenience. In addition, it does not consider the fact that people can take defensive actions. The third column of table 3 presents the unit values for each health effect used for the monetary valuation in this analysis.

Health Benefits

The ambient permit system results in substantially lower health benefits than the emission permit system and the emission standard (figure 3). The differences are largely because each policy imposes different reductions in each cell. The annual benefits obtained are on the order of tens of millions of dollars a year, similar to the annualized costs of reducing emissions.

20. The value of a statistical life estimate used in this study is lower than that estimated by Rizzi and Ortúzar (2003) for Chile using a stated choice approach in which individuals are asked to choose among alternatives. Their estimation adjusted to the 1998 U.S. dollars is approximately $800,000. However, in a recent paper by Rizzi (2005), also using stated-choice surveys, estimated a value of a statistical life for Chile of between US$200,000 and US$300,000.

FIGURE 3. Annual Population-Weighted Health Benefits Associated with Ambient Particulate Matter (PM10) Concentration Targets, by Pollution Control Instrument
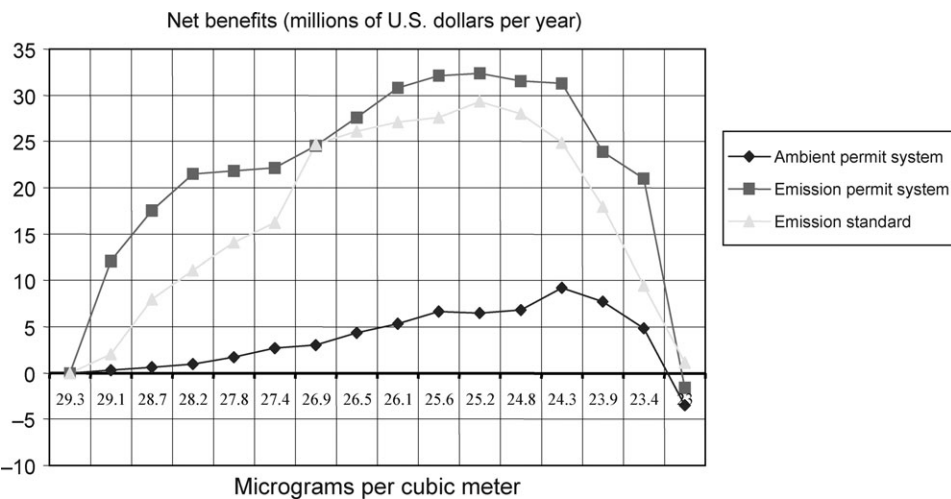

Source: Authors' analysis based on data from Chilean National Institute of Statistics.

IV. COMPARING COSTS AND BENEFITS

Subtracting the annual costs of each policy instrument from the annual benefits yields the net annual benefits to be expected from each policy instrument (figure 4). The net benefits are significantly higher for the emission permit

FIGURE 4. Annual Net Benefits Associated with Ambient Particulate Matter (PM10) Concentration Targets, by Pollution Control Instrument

Source: Authors' analysis based on data from Bravo (2000) and Chilean National Institute of Statistics.

system and the emission standard than that for the ambient permit system. The maximum net benefit is obtained at a PM10 concentration of 25.2 micrograms per cubic meter using the emission permit system. These net benefits are approximately US$32 million per year, almost four times higher than the maximum net benefits under the ambient permit system.

On average, net benefits are ten times higher under the emission permit system and almost six times higher under the emission standard than under the ambient permit system, with the difference even higher in many cases. For example, for a 28.2 micrograms per cubic meter concentration target, the net benefits of the emission permit system are 22 times higher than those from the ambient permit system, and 12 times higher than those of the emission standard. In other cases, the difference is small. For example, for a PM10 concentration level of 24.3 micrograms per cubic meter, net benefits from the emission permit system are only 3.4 higher and those from the emission standard are only 2.7 times higher than those from the ambient permit system.

Requirements to achieve concentration levels below 23 micrograms per cubic meter have negative net benefits because of the sharp increases in cost, even when using flexible instruments. The implication is that the regulatory authority must determine the reduction targets carefully to capture most of the net benefits. A difference as small as three micrograms per cubic meter in the required reduction target can result in significantly lower net benefits.

For mortality, different values of a statistical life do not change the ranking of instruments and the net benefit-maximizing concentration target. For example, with a lower value of a statistical life of US$300,000 the maximum net benefits are achieved with an emission permit system at a concentration level of 25.6 micrograms per cubic meter. The net benefits are, of course, lower, reaching only US$19 million a year.

As conjectured, in a developing country such as Chile, where little effort has previously been undertaken to reduce air pollution, the benefits of better air quality associated with an emission permit system or an emission standard outweigh the relatively small compliance cost reductions obtained with the more cost-effective ambient permit system. Clearly, the decision to apply an emission permit system for Santiago is correct when both costs and benefits are taken into account.

V. Conclusions

The choice of instrument to regulate PM10 pollution that yields the highest net benefit is an empirical matter. For Santiago, a simulation model was used to rank policy instruments using given transfer coefficients, emission coefficients, cost estimates, coefficients for health effects, and unit costs of health effects. The analysis assumed away some of the issues currently being discussed in the theoretical literature on instrument choice, such as imperfect emission

monitoring, information asymmetries, dynamic incentives for innovation, and incomplete enforcement of regulation.

Correcting for the difference in benefits associated with each instrument makes a significant difference in the choice of policy instrument to be used when the air quality goal is fixed and uniform across the airshed, as is usually the case. When only a cost-effectiveness criterion is used, the ambient permit system is clearly the preferred option for Santiago, reducing costs significantly compared with the emission permit system and the emission standard over a relevant range of pollution concentration levels. However, when the benefits associated with the overcontrol achieved using these two instruments are included, the emission permit system has the highest net benefits and the ambient permit system has the lowest net benefits over a wide range of plausible reduction possibilities.

In this latter case one of the main advantages of an ambient permit system plays against it. Since it is able to impose reductions that closely match the uniform standard in different parts of the city, it does not take advantage of the significant health benefits from reducing concentrations more than required by the standard. The efficiency gains of the ambient permit system are much smaller than the economic losses from the health impacts resulting from the higher pollutant concentrations allowed by this instrument. While in principle an ambient permit system could be designed to exactly emulate the concentrations reached by the other two instruments, and this would then clearly be the best option, in practice regulators set up a uniform air quality standard within an airshed rather than a system of differentiated standards.

The emission permit system and the emission standard have higher net benefits than the ambient permit system. An emission permit system is a particularly good policy choice for Santiago. Even though there are efficiency losses compared with an ambient permit system, these are more than compensated for by the health benefits obtained as a result of the reductions in pollutant concentrations in excess of the required standard. An emission permit system is also much simpler to implement than a trading system that involves spatial complexities in each trade.

These results may be applicable to other developing economies where control costs are not extremely high because emissions control is at an early stage. The health benefits from an emission permit system or an emission standard may outweigh the lower abatement costs from an ambient permit system. For developed economies, which do not face the same initial conditions (Oates, Portney, and McGartland 1989), the significant reductions in control costs associated with an ambient permit system might outweigh the losses in health benefits compared with other policies.

VI. Supplementary Material

Supplementary material is available online at http://wber.oxfordjournals.org/

REFERENCES

Abbey, D.E., F. Petersen, P.K. Mills, and W.L. Beeson. 1993. "Long Term Ambient Concentrations of Total Suspended Particulates, Pzone and Sulphur Dioxide and Respiratory Symptoms in a Non-Smoking Population." *Archives of Environmental Health* 48(1):33–46.

Aranda, P. 1996. "Costos y Efectividades de Opciones de Control para la Contaminación del Aire en Santiago." Thesis, Universidad de Chile, Santiago.

Atkinson, S., and D. Lewis. 1974. "A Cost-effectiveness Analysis of Alternative Air Quality Control Strategies." *Journal of Environmental Economics and Management* 1(3):237–50.

Bravo, R. 2000. "Proposición y Evaluación de Instrumentos de Incentivo Económico para mejorar la Calidad del Aire en Santiago: Aplicación al Caso de las Fuentes Fijas." Thesis, Universidad de Chile, Santiago.

Bretschneider, B., and J. Kurfurst. 1987. *Air Pollution Control Technology. Fundamental Aspects of Pollution Control and Environmental Science*. Vol. 8. New York: Elsevier.

Burnett, R.T., D. Krewski, R. Vincent, T. Dann, and J. Brook. 1995. "Associations Between Ambient Particulate Sulfate and Admissions to Ontario Hospitals for Cardiac and Respiratory Diseases." *American Journal of Epidemiology* 142(1):15–22.

CONAMA (Chilean National Environmental Commission). 2000. *Mejoramiento del Inventario de Emisiones de la Región Metropolitana*. Final Report. Santiago.

Dockery, D.W., J. Cunningham, A.I. Damokosh, L.M. Neas, J.D. Spengler, P. Koutrakis, J.H. Ware, M. Raizenne, and F.E. Speizer. 1996. "Health Effects of Acid Aerosols on North American Children: Respiratory Symptoms." *Environmental Health Perspectives* 104(5):500–5.

Environment Canada. 1997. "Human Health and Environmental Benefits of Reducing Sulphur Dioxide Emissions." In *Towards a National Acid Rain Strategy*. Submitted to the National Air Issues Coordinating Committee by The Acidifying Emissions Task Group October 1997. (www.ec.gc.ca/acidrain/acidr/acidr_toc_e.htm).

European Commission. 1998. "ExternE: Externalities of Energy—A Research Project of the European Commission. Vol. 7: Methodology 1998 Update." European Commission, Directorate-General XII, Science Research and Development, Office for Official Publications of the European Communities, L-2920 Luxembourg. (www.externe.info).

Hahn, R., and R. Noll. 1982. "Designing a Market for Tradable Emission Permits." In W. Magat ed., *Reform of Environmental Regulation*. Cambridge, Mass.: Ballinger.

Holz, J.C., and J.M. Sánchez. 2000. "Estimation of Unitary Costs of Mortality and Morbidity and its Application to Assess the Health Benefits of the Santiago Decontamination Plan." Instituto de Economía, Pontificia Universidad Católica de Chile, Santiago.

Jorquera, H. 2002a. "Air Quality in Santiago, Chile: A Box Modeling Approach: I. Carbon Monoxide, Nitrogen Oxides, and Sulfur Dioxide." *Atmospheric Environment* 36(2):315–30.

———. 2002b. "Air Quality in Santiago, Chile: A Box Modeling Approach: II. PM2.5, Coarse and PM10 Particulate Matter Fractions." *Atmospheric Environment* 36(2):331–44.

Krupnick, A. 1986. "Costs of Alternative Policies for the Control of Nitrogen Dioxide in Baltimore." *Journal of Environmental Economics and Management* 13(2):189–97.

Krupnick, A.J., W. Harrington, and B. Ostro. 1990. "Ambient Ozone and Acute Health Effects: Evidence from Daily Data." *Journal of Environmental Economics and Management* 18(1):1–18.

McGartland, A., and W. Oates. 1985. "Marketable Permits for the Prevention of Environmental Degradation." *Journal of Environmental Economics and Management* 12(3):207–28.

Montero, J.P., J.M. Sánchez, and R. Katz. 2002. "A Market-based Environmental Policy Experiment in Chile." *Journal of Law and Economics* 45(1):267–87.

Montgomery, David. 1972. "Markets in Licenses and Efficient Pollution Control Programs." *Journal of Economic Theory* 5(3):395–418.

Muñoz, Ricardo. 1993. "Desarrollo de un Modelo de Celdas Múltiples para la Simulación de la Dispersión de Contaminantes en la Atmósfera: Aplicación a la Ciudad de Santiago." Thesis, Universidad de Chile, Santiago.

Oates, W., W. Portney, and A. McGartland. 1989. "The Net Benefits of Incentive-Based Regulation: A Case Study of Environmental Standard Setting." *The American Economic Review* 79(5):1233–48.

O'Ryan, R. 1996. "Cost-effective Policies to Improve Urban Air Quality in Santiago, Chile." *Journal of Environmental Economics and Management* 31(3):302–13.

———. 2002. "Emissions Trading in Santiago: Why Has it Not Worked, but Been Successful?" Paper presented at Concerted Action for Tradable Emissions Permits Workshop, University College, March 24–25, London.

———. 2006. "Factors that Determine the Cost-effectiveness Ranking of Second-best Instruments for Environmental Regulation." *Journal of Regulatory Economics* 30(2):179–98.

O'Ryan, R., and L. Larraguibel. 2000. "Contaminación del Aire en Santiago: Estado Actual y Soluciones." Center for Applied Economics Working Paper 75. Universidad de Chile, Santiago.

Ostro, B. 1990. "Associations Between Morbidity and Alternative Measures of Particulate Matter." *Risk Analysis* 10(3):421–27.

———. 1996. "A Methodology for Estimating Air Pollution Health Effects." World Health Organization, Office of Global and Integrated Environmental Health, Geneva.

Ostro, B., J.M. Sánchez, C. Aranda, and G. Eskeland. 1996. "Air Pollution and Mortality: Results from a Study of Santiago, Chile." *Journal of Exposure Analysis and Environmental Epidemiology* 6(1):97–114.

Ostro, B., G. Eskeland, J.M. Sánchez, and T. Feyzioglu. 1999. "Air Pollution and Health Effects: A Study of Medical Visits among Children in Santiago, Chile." *Environmental Health Perspectives* 107(1):69–73.

Portney, P.R. 1990. *Public Policies for Environmental Protection*. Washington, D.C.: Resources for the Future.

Rizzi, L.I. 2005. "The Subjective Value of Road Safety in Chile." Proceedings of Road Safety Conference on Four Continents, Warsaw, October 5–7, Poland.

Rizzi, L.I., and J.D. Ortúzar. 2003. "Stated Preference in the Valuation of Interurban Road Safety." *Accident Analysis and Prevention* 35(1):9–22.

Seskin, E., R. Anderson, and R. Reid. 1983. "An Empirical Analysis of Economic Strategies for Controlling Air Pollution." *Journal of Environmental Economics and Management* 10(2):112–24.

Spofford, W., and C. Paulsen. 1988. "Efficiency Properties of Source Control Policies for Air Pollution Control: An Empirical Application to the Lower Delaware Valley." Resources for the Future Discussion Paper QE18-13. Resources for the Future, Washington, D.C.

Tietenberg, T.H. 1985. *Emissions Trading, an Exercise in Reforming Pollution Policy*. Washington, D.C.: Resources for the Future.

USEPA (United States Environmental Protection Agency). 2000. Regulatory Impact Analysis: "Heavy Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements." Office of Air and Radiation. Washington, D.C.

Vatavuk, W. 1990. *Estimating Costs of Air Pollution Control*. Chelsea, Mich.: Lewis Publishers.

Whittemore, A. S., and E. L. Korn. 1980. "Asthma and Air Pollution in the Los Angeles Area." *American Journal of Public Health* 70(7):687–96.

Women's Power, Conditional Cash Transfers, and Schooling in Nicaragua

Seth R. Gitter and Bradford L. Barham

The Social Safety Net (Red de Protección Social, RPS) program in Nicaragua is one of many conditional cash transfer programs that pay households cash stipends in exchange for school attendance and regular visits to health clinics by the children. A key feature is that payments go to the female head of household. Previous research suggests that exogenous transfers to women are more likely to be spent on their children's health, nutrition, and education and thus to reinforce the goals of these programs. Randomized experimental data from RPS are used to test for heterogeneous program impacts on school enrollment and spending based on a woman's power, as proxied by her years of schooling relative to her husband's years of schooling. The results confirm previous findings that more household resources are devoted to children when women are more powerful. However, when a woman's power greatly exceeds her husband's, additional female power reduces school enrollment. RPS impacts on schooling are much larger than the expected income effects estimated from the control group, although no evidence is found that female power alters the impact of RPS on school enrollment. The conditionality of RPS is probably decisive. While RPS significantly increases food and education expenditures, the impact is attributable primarily to income effects. JEL codes: D13, H31, I20

Many poverty alleviation programs in developing countries stipulate that payments or benefits be given to the female head of household (Rawlings and Rubio 2005). The justification for targeting women is based on theoretical models and empirical findings that show that payments received by women are more likely to be spent on improving the welfare of children (for theoretical work, see Kanbur and Haddad 1994; Haddad, Hoddinott, and Alderman 1997; Basu 2006; for empirical research, see Schultz 1990; Thomas 1990; Doss 1996). This article explores the impact of this requirement in Nicaragua's

Social Safety Net (Red de Protección Social, RPS), a conditional cash transfer program that pays women cash if their children attend school and they make regular visits to health care clinics.

Empirical evidence is limited on the effectiveness of targeting conditional cash transfers to women in order to raise school enrollment and affect other consumption outcomes. Three critical components of conditional cash transfers confound efforts to cleanly identify the impacts on school enrollment: income and two nonincome effects, conditionality and intrahousehold effects. The nonincome effects of targeted conditional cash transfers potentially include both the conditionality requirements of program participation (essentially a price effect) and the intrahousehold effects of providing women with the transfer. In other words, the education outcomes are also shaped by two distinct effects that are both part of the program's treatment.

Would a cash transfer without conditions achieve similar school enrollment outcomes (because education is a normal or even superior good for low-income families)? As for identifying the intrahousehold impacts of cash transfers targeted to women, an ideal experiment would randomly provide some transfers to men and some to women to determine how the impacts differ. Absent such a study design, one could examine household spending patterns of the treatment group (that are not conditional) to determine whether intrahousehold effects matter and in what ways. One could also look at the effects of intrahousehold differences in the control group (or the baseline data of the treatment group) to determine whether preexisting differences in education and spending patterns are consistent with power differences between men and women.

This article explores how RPS shapes education and spending patterns, with an eye on all three effects: income, conditionality, and intrahousehold impacts. On the intrahousehold side, the intention is to identify whether preexisting gender power structures are at work and to determine whether they are mitigated by the program, either through conditionality or by targeting transfers to women. By providing transfers directly to women, RPS also has the potential to empower women by increasing the resources they control. However, household resources are potentially fungible, raising a concern that other family resources may be reallocated away from children, offsetting the impact of the transfer. This phenomenon could be captured empirically by demonstrating smaller effects of conditional cash transfers on key outcomes in households in which men have more power. By targeting transfers to women, RPS has the implicit goal of helping ensure that money is spent on women and children, who might otherwise receive smaller shares of household resources in male-dominated households. Thus, it is also possible that the impacts of conditional cash transfer programs could be higher in male-dominated households if the transfers have the effect of changing behavior in the family that did not contribute to salutary outcomes for children.

The empirical analysis uses experimental methods that compare treatment and control groups. It adds to previous studies of the impact of conditional cash transfers by estimating heterogeneous program impacts based on

intrahousehold power differences. The power measure used is based on the ratio of years of school completed by the female and male heads of household. Women's intrahousehold power is assumed to increase as the female to male education ratio rises. This measure is arguably better in terms of exogeneity than male and female wage earnings used in some other studies, because earnings are endogenous to intrahousehold decision making and correlated with child wages, both of which could affect schooling decisions.

The article is organized as follows. Section I places this work within the context of the current literature and identifies its conceptual contributions. Section II presents the empirical approach to analyzing the impact of power and RPS on schooling and household spending. Section III provides background information on RPS along with descriptive statistics on variables of interest. Results of the estimations are reported in section IV, with conclusions and suggestions for further study provided in the last section.

I. Literature Review

This article links three related streams of literature. The first is the intrahousehold bargaining literature, which suggests that heterogeneous preferences between men and women can lead to different household decisions depending on power relations. The second attempts to measure the impacts of conditional cash transfer programs, with a focus on which aspects of the program (conditions or cash) are more effective in obtaining the desired results. The third seeks to determine whether there are demonstrable effects of targeting conditional cash transfers to women.

The theoretical and empirical literature on how households make decisions is well developed (Schultz 2002; Basu 2006). Two basic types of household models have been used to study decisions on child schooling and labor and the allocation of consumption expenditures between private and shared goods. Unitary models assume either that there is a benevolent dictator or that household members share the same preferences and pool their resources to maximize a single household utility function (Becker 1981). Households with heterogeneous preferences and a set balance of power are guided by a single utility function, even when one member is a nonbenevolent dictator. In these models, targeting transfers to women should have no impact on a household's allocation of spending (except through household income effects; Attanasio and Lechene 2002).

Nonunitary models generally examine decisions made by men and women who have distinct preferences and make decisions somewhere along a spectrum between full cooperation and conflict (McElroy and Horney 1981; Chiappori 1992; Basu 2006). Differences in bargaining power influence whose preferences gain greater expression in the household's choices. These models often assume that women have stronger preferences for child schooling and health outcomes; they therefore predict distinct effects of increases in nonwage income

depending on who receives the transfer. The motivation for giving conditional cash transfers to women is the assumption that women's higher propensity to spend on household shared goods will augment program effects.

Power relations between fathers and mothers have been shown empirically to affect child schooling outcomes (Binder 1999; Adato and others 2003; Iyigun and Walsh 2007), with relative income increases for women raising child school attendance. Thomas (1990) and Schultz (1990) show that nonwage income received by mothers is more likely than income received by fathers to be spent on children's health or schooling. The child's gender may also affect the resources received. Thomas (1994) shows that Brazilian mothers' nonwage income positively affected their daughter's health but not their sons'. Duflo (2003) shows that the impacts of exogenous income transfers through old-age pensions in South Africa were more likely to increase health outcomes of granddaughters of grandmothers than any other grandparent–grandchild relation. Emerson and Souza (2007) find that in Brazil fathers' education has a greater impact than mothers' education on sons' attainment, while mothers' education matters more to daughters' attainment.

Attanasio and Lechene (2002) and Adato and others (2003) examine the intrahousehold decision-making effects of conditional cash transfer programs. Both consider Progresa (now known as Oportunidades), a Mexican conditional cash transfer program. Attanasio and Lechene test the impact of Progresa and women's bargaining power as measured by the relative wages (potential and actual) of men and women on the share of household expenditures devoted to different goods (food, alcohol, transportation, services, and clothing).[1] The importance of women's power is supported by results that show that an increase in the relative income of women, including from Progresa's targeted cash transfer, has a positive relation to the share of expenditures allocated to children's clothing and food. Using a qualitative approach, Adato and others (2003) find that Progresa decreased the likelihood that husbands reported being the sole decision maker regarding spending on child health care, school attendance, and clothing, suggesting that the targeted cash transfer increased women's bargaining power.

One critical methodological and empirical issue in the intrahousehold literature is how to measure bargaining power. Adato and others (2003) suggest that each member's bargaining power is based on four factors: control over resources, influence over the bargaining process, interpersonal networks, and basic attitudinal attributes. Most research suggests that those with greater own assets or income (actual or potential) can exert more power, because they can withdraw from the household more easily (Doss 1996). In this sense, conditional cash transfers could increase the viability of women's exit options and

---

1. Because wage data are not available in the RPS sample, the results of Attanasio and Lechene cannot be compared with those obtained here.

strengthen their bargaining power, as long as women receive the transfer even if they leave the household.

This article uses the ratio of the number of years of school completed by the female to the number of years of school completed by the male head of household as a measure of power.[2] It assumes that as the female to male education ratio increases, women are likely to have more decision-making power. This measure is similar to but less crude than the literacy ratio Basu, Narayan, and Ravallion (2001) use (they use literacy because they assume that a literate member can withhold information from illiterate members to gain an advantage). One advantage of the education ranking approach over power measures that rely on relative wages or income is that education is exogenous to current income levels, which are themselves endogenous to fundamental household decisions regarding labor allocation. (Wages could not have been used in any case, because the RPS sample data did not collect wage information.)

Many studies suggest that women's power is both positively and monotonically related to spending on children and school enrollment. This assumption has been questioned by some recent work, however (Felkey 2005; Basu 2006; Lancaster, Maitra, and Ray 2006). Using an intrahousehold theoretical framework, Basu (2006) shows that if the woman has more power than the man, she will garner a greater share of the income produced by child labor. Based on this result, he posits that as her power continues to increase, she will receive more benefits from child labor, while the benefits of schooling may stay the same. He therefore concludes that if women become sufficiently more powerful than men, additional female power may actually result in a decline in school enrollment. Lancaster, Maitra, and Ray (2006) and Felkey (2005) provide empirical evidence in support of Basu's hypothesis, using samples from India and Bulgaria.

In Nicaragua even when women have as much education as their husbands, they still may not have equal power, because of cultural norms. However, at a certain point women with more education than their husbands could have sufficient power to sustain the nonmonotonic result suggested by Basu.

Basu's hypothesis is tested here by examining the nonlinear effects of the female to male education ratio on child school enrollment and household spending outcomes. The article also adds to these previous studies by testing whether cash payments made to mothers are likely to increase their power.

Previous work has shown that RPS and Progresa have been effective at increasing school enrollment rates and encouraging spending on food (for RPS, see Maluccio and Flores 2004; for Progresa, see Schultz 2004; Hoddinott and Skoufias 2004). The regression specification used by Hoddinott and Skoufias includes total consumption (including the transfer) as well as program

---

2. The female–male power ratio is usually given on a 0 to 1 scale. The ratio here ranges from 0 to 9. Though it could be normalized to a 0 to 1 scale, a non-normalized ratio is used for easier interpretation of the coefficients.

participation indicator variables. This combination helps provide estimates for the income and nonincome impacts of Progresa on food spending.

Hoddinott and Skoufias (2004) find that nonincome effects account for about half of the total impact of Progresa for total food expenditures and a higher percentage for expenditures on fruit, vegetables, and animal products. They place much of the credit for these impacts on lectures women received as part of Progresa that encourage proper nutrition through expenditures on fruit, vegetables, and milk. Attanasio and Lechene (2002) contend that the impacts may also be tied to targeting payments to women. Both could be correct: the health education lectures provided by Progresa could shape preferences, and targeted transfers could enhance women's bargaining power and thus their capacity to reveal those preferences. What is not clear is whether those expenditures may also have been viewed implicitly by the recipients as part of the conditionality of Progresa. In a simulation of the Bolsa Escola Program, a Brazilian conditional cash transfer, Bourguignon, Ferreira, and Leite (2003) find that both the conditionality of school attendance and income effects increase school enrollment.

Other research suggests that preexisting household conditions can shape the impact of a transfer. de Janvry and Sadoulet (2006) argue that conditional cash transfer programs can improve their results by moving from a uniform transfer size to one tied to easily observable household characteristics that alter program impacts. The relative education levels of parents are used here as an easily observable characteristic that may create heterogeneous impacts based on differences in preferences and power between men and women. Attanasio and Lechene (2002) find that payments made to women increase expenditures on food and schooling by increasing women's power (as measured by the ratio of female to male income), but they do not test for nonlinearities in this relation. It is possible that transfers to less powerful women may increase their power enough to participate in decision making and thus augment the targeting effect (as suggested by Adato and others 2003). Another possibility is that less powerful women may not be able to keep the whole transfer or that men may withdraw funds from the household to increase personal expenditures or leisure time. In the case of RPS, there is little evidence for this occurring, as increases in total household consumption were equivalent to the size of the transfer (Maluccio and Flores 2004).

de Janvry and Sadoulet (2006) include parental literacy in their estimation of the impact of Progresa on the child schooling decision. The impacts of the literacy of the mother and father are estimated as separate effects, not relative to one another as a measure of power. They find that both father's and mother's literacy increase schooling and decrease the size of the transfer required for the child to attend school. Their regression does not include controls for income, however, so parental schooling may well be capturing an income effect. Most important, they do not compare across households with

different female to male education ratios or other relative power measures to test for intrahousehold effects.

II. AN EMPIRICAL STRATEGY FOR ESTIMATING THE IMPACTS OF POWER AND RPS

Three components of household schooling and resource allocation decisions are examined here. The first is the effect of power structures ex ante of program effects on education outcomes and household spending patterns. The goal of this test is to see whether the power measure provides results that are consistent with the previously cited literature—that is, whether children of more powerful women are more likely to attend school and receive a larger share of resources.

The second component is an estimate that identifies income and nonincome effects. The control group is used to estimate income impacts on schooling and household spending. The income effects of a cash transfer in the control group which is the size of that of the RPS are then compared with the total effects of RPS, with the difference being an estimate of nonincome effects.The third component is the effect of women's power on program impacts on school attendance and household expenditure patterns. This component is measured by interacting variables that measure program impact and the power measure to test for heterogeneous program impacts by power.

The conventional approach to analyzing the treatment effects of conditional cash transfer programs is to use cross-sectional or panel data to compare outcomes in treatment and control groups. When the dependent variable of interest (school enrollment or consumption share) is not substantially different in the baseline year in control and treatment communities, program impacts can be measured using cross-sectional data in the treatment year. However, if initial conditions (in either the dependent or independent variables) are different in the treatment and control communities, then the full panel data should be used.

Difference-in-difference is the standard method used to measure impacts when initial conditions are not the same in control and treatment communities. This method measures the difference in the changes of the outcome of interest in treatment and control communities between the first year of treatment (year 1) and the baseline (year 0). If, for example, the outcome of interest in time period t is denoted as C_t for control communities and I_t for those in the treatment (intervention) group, the difference-in-difference program impact, denoted δ_1, is determined by $\delta_1 = (I_1 - I_0) - (C_1 - C_0)$. If through randomization in the baseline the outcome of interest is equally likely in both groups, the difference-in-difference impact is equivalent to $I_1 - C_1$.

Maluccio and Flores (2004) present a basic estimation equation for difference-in-difference (equation 1). Program impacts are measured using the difference-in-difference variables; δ_1 the coefficient on the term *Treat*, which is the interaction of two binary dummy variables for treatment year ($T = 1$); and the treatment status of the household ($RPS = 1$ for households in a treatment community).[3]

$$(1) \qquad E_{ict} = \alpha_0 + \alpha_1 A_1 + \alpha_2 A_2 + \delta_0 RPS + \delta_t Treat + u_{ic} + v_{ict}$$

where $E_{ict} =$ outcome variable of interest for household (or individual) i in community c at time t, $A_1 = 1$ if year is 2001, $A_2 = 1$ if year is 2002, *Treat* $= 1$ if treatment year is 2001 or 2002 and household is in RPS intervention in community c, $u_{ic} =$ all (observed and unobserved) household-level (or individual-level) time-invariant factors, $v_{ict} =$ unobserved idiosyncratic household (or individual) and time-varying errors, and α's and δ's $=$ unknown parameters.

The number of years of school completed by the female head of household divided by the number of years of school completed by the male head of household—relative female power by schooling years (*rFPSY*)—is used to measure power. As 49 percent of males have completed zero years of school, 1 is added to both numbers of school years to create a defined ratio for all households:

$$rFPSY = \frac{(Number\,of\,years\,of\,school\,completed\,by\,female\,head\,+\,1)}{(Number\,of\,years\,of\,school\,completed\,by\,male\,head\,+\,1)}$$

The variable *rFPSY* is used to measure the impact of female power on school enrollment and household expenditures. The average *rFPSY* of both control and treatment groups was 1.4; comparison between treatment and control groups does not show statistically significant differences between the two groups.[4] The square of *rFPSY* is also used in order to test for the possible nonlinearity of the relation between power and these outcomes. The power measure is interacted with the treatment impact measure *Treat* to estimate the interactive effects of the power measure and *RPS*. The square of the power measure and the treatment impact measure are interacted to test for a nonlinear relation between power and *RPS* impacts. Schooling of the male and female heads of household (male_schooling and female_schooling) is added directly into the equation to control for the impact of the individual education levels. Finally, total per capita consumption (PCC) is included to estimate and control for income effects, including those from RPS transfers. When PCC is included,

3. For ease of interpretation, the two years are combined into a single measure of the impact of the treatment in a treatment year; doing so does not substantially affect the results.

4. A simple *t*-test of the mean of *rFPSY* between the treatment and control group yields a *t*-statistic of 0.37.

the estimated impacts of nonincome effects *Treat*RPS* in equation (2) for all households is represented by δ_1. The estimated impacts of power on RPS effects are represented by δ_2 and δ_3, respectively.

$$
\begin{aligned}
E_{ict} = {} & \alpha_0 + \alpha_1 A_1 + \alpha_2 A_2 + +\alpha_3 Male\_Schooling \\
& + \alpha_4 Female\_Schooling + \alpha_5 rFPSY + \alpha_6 rFPSY\char`^2 \\
& + \delta_0 RPS + \delta_1 Treat + \delta_2 Treat * rFPSY + \delta_3 Treat * rFPSY\char`^2 \\
& + \beta_1 \ln Consumption_{ct} + \beta_2 \ln Size_{ct} + u_{ic} + v_{ict}
\end{aligned}
$$

(2)

where $E_{ict} = 1$ if child i in community c at time t is enrolled in school, and 0 otherwise;

or, for expenditure data, E is expenditures for household i in community c at time t; $\ln Consumption = \log (total\ consumption)$ for household c in year t (baseline); and $\ln Size_t = \log(household\ size)$ in year t.

For the first two components, this regression specification is similar to that of Hoddinott and Skoufias (2004), who estimate the impact of Progresa on food consumption—with some important distinctions. Their specification includes household characteristics, including the education of the head; the specification presented here includes the education of both the household head and his or her spouse separately and as a power measure. The same method used by Hoddinott and Skoufias (2004) is adopted here to separate income effects from nonincome effects by including total consumption in the regression as a control for income (including the transfer) as well as program effect measures. As Hoddinott and Skoufias note, if a conditional cash transfer alters consumption other than directly through transfers, total consumption becomes endogenous and may bias the results. This does not appear to be the case, as Maluccio and Flores (2004) find that the ex post increases in consumption for the treatment group are not statistically significantly different from the transfer.

The final component of the specification is the measurement of heterogeneous impacts of RPS based on household characteristics. The approach used is similar to that of two previous studies that measure the effect of economic shocks on RPS (Maluccio 2005) and Progresa (de Janvry and others 2006). In these studies, the heterogeneity across households is determined by exposure to these shocks. A measure of exposure to shocks is then interacted with the program eligibility variable. The approach here is similar, except that heterogeneity comes from the power measure rather than exposure to shock.

Both types of models (school enrollment and expenditure levels) are estimated using ordinary least squares (OLS). Estimating marginal effects is difficult using qualitative variable methods because of the interaction terms. Gitter and Barham (2006) find that OLS estimations of the enrollment impacts of

RPS are similar to probit predictions. In all of the estimations, errors are clustered at the community level to control for unobserved heterogeneity across communities. Because the household decision on school attendance may be different for boys and girls, separate estimates are performed for boys and girls.

III. SUMMARY OF THE RPS PROGRAM AND DESCRIPTIVE STATISTICS

The RPS was implemented in 21 randomly selected communities in northwestern Nicaragua (in Madríz and Matagalpa). Another 21 communities in the region served as the control group. Three survey rounds were conducted in all 42 communities, one in 2000, before program implementation, and two during the program, in 2001 and 2002. This analysis uses a subsample of the 1,300 total households in which there is a head of household who is married. This subsample includes 1,129 households.[5]

Participation in treatment communities was extremely high, with uptake rates of more than 95 percent of those eligible to participate.[6] Benefits include a C$2,880 ($224) annual food security transfer.[7] Households with children ages 7–13 who had not completed the fourth grade were eligible for a bimonthly transfer for school attendance of C$1,440 per year and an additional C$275 for school supplies. The average household received C$3,885, or about 18 percent of total annual household consumption expenditures.

Baseline comparisons between treatment and control groups on outcomes and explanatory variables support the use of experimental methods to test for impact results. The average school enrollment for children of eligible age in the baseline sample was 77 percent, with about a 0.1 percent difference between treatment and control groups. The difference in aggregate total consumption and other consumption measures in treatment and control groups was not significantly different from zero.

In more than 40 percent of the households, male and female heads had completed the same number of years of school. The other 60 percent of households were divided evenly between those in which the male head had more schooling and those in which the female head had more schooling (table 1). The control and treatment groups had a similar average *rFPSY*. However, the control group had slightly more (45 percent compared with 40 percent) households in which the male and female households had the same number of years of completed

5. See Maluccio and Flores (2004) for information on the program design. They show that sample attrition rates were similar in both control and treatment communities.

6. Ninety-five percent of households were eligible to participate (Maluccio and Flores 2004). Program participation does not appear to have been affected by adult literacy, household income, or marital status.

7. Exchange rate of C$12.85 = US$1 is as of September 2000.

TABLE 1. Descriptive Statistics of Total Household Consumption and School Enrollment

Item	rFPSY[a]	Baseline			Difference-in-difference	
		Control	Treatment	T-statistic[b]	2001[c]	2002[d]
Total household consumption (córdobas)	rFPSY < 1	25,160	24,427	0.51	6,694**	2,773**
	rFPSY = 1	22,206	21,634	0.64	6,492**	5,114**
	rFPSY > 1	24,291	24,051	0.81	4,164**	4,489**
	Total	23,624	23,147	0.70	5,851*	4,284**
School enrollment, ages 7–13 (percent)	rFPSY < 1	78	78	−0.04	17**	14**
	rFPSY = 1	72	74	−0.65	15**	10**
	rFPSY > 1	86	80	1.50	20**	12**
	Total	77	77	−0.21	17**	11**

*Significant at the 10 percent level.

**Significant at the 5 percent level.

[a](years of schooling completed by female head +1)/(years of schooling completed by male head +1).

[b]Comparison of baseline control and treatment.

[c]$(\text{Treatment}_{2001} - \text{Control}_{2001}) - (\text{Treatment}_{2000} - \text{Control}_{2000})$.

[d]$(\text{Treatment}_{2002} - \text{Control}_{2002}) - (\text{Treatment}_{2000} - \text{Control}_{2000})$.

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

schooling, while the treatment group had slightly more (31 percent compared with 26 percent) households in which women had more years of schooling.

Consumption in households with more powerful females ($rFPSY > 1$) is similar to that in households with more powerful males ($rFPSY < 1$)[8]; consumption is lower in households in which $rFPSY = 1$. This result likely reflects the fact that this group includes a significant number of households in which neither spouse completed a year of school.

Previously cited literature suggests that female power is linked to higher school attendance and spending on children. The predicted relation is found in table 1, which shows that households with more powerful women ($rFPSY > 1$) have average baseline school enrollment rates of 82 percent (86 percent and 80 percent for the control and treatment groups, respectively), while households in which $rFPSY < 1$ have school enrollment of 78 percent. A *t*-test on average enrollment between the two groups yields a *t*-statistic of 1.78.

The relation between power and spending can be seen in some of the other explanatory variables (table 2). The previously cited literature suggests that households in which women have more power spend more on food and education of their children. However, in the RPS sample, there is weak evidence in

8. A *t*-test comparing the total consumption of households with $rFPSY < 1$ and $rFPSY > 1$ yields a *t*-statistic of 0.6. Relative to households with unequal levels of schooling, households in which $rFPSY = 1$ have a *t*-statistic of 3.6.

TABLE 2. Descriptive Statistics of PCC of Food, Milk, and Education (córdobas)

| Item | rFPSY[a] | Baseline | | T-statistic[b] | Difference-in-difference | |
		Control	Treatment		2001[c]	2002[d]
Per capita food expenditures	rFPSY < 1	2,838	3,100	−0.97	673**	716
	rFPSY = 1	2,755	2,829	−0.31	745**	321
	rFPSY > 1	2,744	2,919	−0.77	611	603
	Total	2,801	2,969	1.20	618**	514
Per capita milk expenditures	rFPSY < 1	53	64	−0.65	57**	64**
	rFPSY = 1	29	30	−0.09	24	34**
	rFPSY > 1	75	66	0.52	67**	31**
	Total	49	50	0.20	48**	42**
Per capita education expenditures	rFPSY < 1	70	52	−0.93	23	75**
	rFPSY = 1	30	51	−1.90	7	19
	rFPSY > 1	78	94	−0.50	−22	47
	Total	56	65	−0.70	2	43

**Significant at the 5 percent level.

[a](years of schooling completed by female head +1)/(years of schooling completed by male head +1).

[b]Comparison of baseline control and treatment.

[c]$(Treatment_{2001} - Control_{2001}) - (Treatment_{2000} - Control_{2000})$.

[d]$(Treatment_{2002} - Control_{2002}) - (Treatment_{2000} - Control_{2000})$.

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

terms of total food spending, though food expenditures account for such a high proportion of total consumption (70 percent) that the deep poverty of these families may blunt differences in food expenditures evident elsewhere. Expenditure data come from self-reported household surveys on food consumption over a two-week period, which was scaled up for a year (Maluccio and Flores 2004).

Children in households with a powerful woman might receive a larger proportion of the household's food. Unfortunately, data on individual food consumption are not available, however. One way of determining whether this is the case is to look at milk consumption (including infant formula), which is more likely to benefit children. Milk consumption does appear to be related to women's power: in the baseline data, households with $rFPSY > 1$ consume more milk than those with $rFPSY < 1$ (the difference is significant at the 10 percent level using a simple t-test).

Maluccio and Flores (2004) use difference-in-difference estimates to measure program outcomes in their analysis of the total impact of RPS. Tables 1 and 2 provide basic difference-in-difference estimations for each of the *rFSPSY* measures for the outcomes of school enrollment, expenditures, and expenditure shares. In terms of school enrollment, the impacts are larger in households in

which the woman is powerful. All households saw at least a 15 percentage-point increase in enrollment the first year and a 10 percentage point increase the second year. Given that enrollment was at least 95 percent in all treatment communities, conditionality appears to be playing the dominant role. The effects were greater, however, in households with more powerful women.

One common concern is that men might withdraw money from the household for shared goods as women receive income from the transfer and use it for private consumption. If this concern were evident in the data, one would expect male-dominated households to have smaller expenditure impacts from RPS. In fact, in all cases except milk expenditures in 2002, impacts from RPS treatment as measured by difference-in-difference estimates show larger impacts for male-dominated households than for female-dominated households. This suggests that RPS transfers to women are having the intended impact of strengthening their potential to influence household consumption and investment choices rather than being captured by men who had pretransfer power advantages.

IV. Econometric Results

This section presents the results of econometric estimations of factors shaping school enrollment and household expenditures. There are three major components of these influences: the effect of female power ex ante of program impacts, income versus nonincome impacts, and variation in program impacts by female power. Two sets of regression results are reported: impacts on school enrollment and impacts on per capita expenditures for food, education, and milk. The regression specification is supported by the finding of an ex ante impact of female power on school enrollment and household expenditures on education. The results also show both income and nonincome effects from RPS, with nonincome effects being more important for schooling and income effects being more important for household spending patterns.

The econometric analysis of school enrollment outcomes for children ages 7–13 includes three sets of regressions: one for all children and one each for boys and girls (table 3). The impact ex ante of gender power differences can be seen through the two *rFPSY* measures. The coefficients on both *rFPSY* (positive) and *rFPSY*^{2} (negative) are statistically significant for the sample of all children and girls. Children's schooling is positively associated with maternal power, except when the *rFPSY* ratio is larger than 5 (the case for about 3 percent of the children in the sample), at which point further maternal schooling begins to reduce enrollment. These results are consistent with the nonmonotonic relation between power and schooling found by Basu (2006). However, as the negative effect is observed only at the far tail of the distribution, it could also indicate that there are monotonic but diminishing returns to power and schooling. For boys the quadratic term is not statistically significant, and the results suggest a positively monotonic relation between female power (*rFPSY*) and school enrollment.

TABLE 3. Regression on School Enrollment: Impacts of Power and RPS

Variable	All Children		Boys		Girls	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
RPS 1 if treatment group	-0.018	0.018	0.026	0.025	-0.068**	0.024**
1 if year 2001	0.048**	0.016**	0.055**	0.024**	0.036	0.022
1 if year 2002	0.062**	0.016**	0.074**	0.024**	0.044**	0.022**
Male household head school years	0.016**	0.004**	0.022**	0.006**	0.010	0.006
Female household head school years	0.008**	0.004**	0.008	0.006	0.009	0.006
rFPSY	0.053**ᵃ	0.018**	0.056**ᵃ	0.028	0.044**ᵃ	0.024**
rFPSY^2	-0.006**	0.002**	-0.004	0.004	-0.007**	0.003**
1 if RPS group and treatment year (Treat)	0.166**	0.030**	0.153**	0.045**	0.184**	0.041**
Treat*rFPSY	-0.030	0.026	-0.034	0.041	-0.025	0.033
Treat*rFPSY^2	0.004	0.005	0.003	0.008	0.005	0.006
Ln(household consumption)	0.054**	0.010**	0.047**	0.015**	0.063**	0.014**
Ln(household size)	0.000	0.016	-0.002	0.022	-0.003	0.023
Constant	0.153	0.098	0.172	0.146	0.137	0.132
R-squared	0.09		0.10		0.09	
Number of observations	4,593		2,337		2,256	

**Significant at the 5 percent level.

[a]Joint F-test of coefficients on $rFPSY = rFPSY^2 = 0$ significant at the 5 percent level.

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

Additional controls were added to the regressions for the number of school years completed by the male and female heads of household. The pooled regression shows that an additional year of school for the male head is twice as important as an additional year for the female head. These differences break mainly along child gender lines. In the sample with just boys, schooling of the male head has a larger impact (0.022) for each year of schooling compared with an extra year for the female head (0.008), although an *F*-test of *male_years* = *female_years* is not statistically significantly different from zero. This result suggests that additional school years of the male head and female head of household may have equal impacts. The coefficient estimates on both heads of household are equal for girls, although neither is statistically significant.

The second component of interest—the comparison of income and nonincome effects—is captured by the RPS impact measures (*Treat*), because income effects are controlled for by using total household consumption (including RPS transfers). The RPS nonincome impacts on school enrollment for both years are measured at 16.6 percent for the total sample, with the impact on girls slightly higher but not statistically significantly so. This estimate is lower than that of Maluccio and Flores (2004), who estimate total impacts (income and nonincome) of 22 percent for 2001 and 18 percent for 2002. This difference suggests that the income effects are on the order of 1.4–5.4 percentage points, or about 25–33 percent of the nonincome effects.

Another way to estimate income effects is to use the coefficient estimate on the variable of the natural log of total household consumption, ln*Consumption*. The difference in the average ln*Consumption* between treatment and control was 0.35 in 2001 and 0.24 in 2002. With a coefficient estimate of 0.054 on total household consumption, these differences would suggest that transferring the size of RPS would increase schooling by 1–2 percentage points.[9] This impact is slightly less than but consistent in magnitude with the difference between the estimated nonincome effects obtained here and the total effects obtained by Maluccio and Flores (2004).

The combined impacts of power and RPS on school enrollment are examined through the interaction of the nonincome treatment impact measure (*Treat*) and the power ratio (*rFPSY*). This interaction term and its square are not statistically significant, suggesting that the impacts of RPS treatment do not vary depending on the power of the female head of household. Furthermore, when the interaction of *Treat* and $FPSY^2$ is omitted, the relation between RPS impacts and power (*Treat*rFSPY*) is negative but not statistically significant.[10]

---

9. One concern is that with treatment the impact of total consumption on schooling may vary when compared with ex ante consumption patterns. Models that separately estimate the impact of consumption on schooling for only the control group yield coefficients that are not substantially different from the model presented above. These results are available from the authors on request.

10. These results are omitted because of space constraints; they are available from the authors on request.

TABLE 4. Regression on per Capita Expenditures by Category: Impacts of Power and RPS

Variable	Per Capita Spending on Food		Per Capita Spending on Education		Per Capita Spending on Milk	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
RPS 1 if treatment group	56.3	106.0**	1.4	11.9	1.8	8.7
1 if year 2001	−370.8**	99.3**	33.3**	11.2**	21.1**	8.0**
1 if year 2002	−423.0**	96.9**	56.4	10.9**	3.4	7.8
Male household head school years	−20.7	25.8**	4.6	2.9	1.1	3.1
Female household head school years	58.2**	25.0**	5.7**	2.8**	9.6**	4.4**
Relative female power by schooling years ($rFPSY$)	−171.3	113.1	18.2	12.7	−11.3	11.7
$rFPSY^2$	16.7	14.7	−2.7	1.7	1.6	1.4
1 if RPS group and treatment year (*Treat*)	445.2**	175.9	−0.7	19.8	72.0**	26.6**
Treat*rFPSY	−237.0[a]	144.7	1.1	16.3	−45.6[a]	32.3
Treat**rFPSY^2*	25.3	24.5	0.1	2.8	8.4	7.0
Ln(*household consumption*)	1,895.3**	52.3**	113.0**	5.9**	24.0**	7.2**
Constant	−1,5524.0	521.0	−1,078.7	58.6	−193.8	70.2
R-squared		0.37		0.16		0.03
Number of observations		2,550		2,550		2,550

**Significant at the 5 percent level.

[a]Joint *F*-test of coefficients on $rFPSY = rFPSY^2 = 0$ or Treat*rFPSY = Treat**rFPSY^2* = 0 significant at 5 percent level.

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

The impacts of power, RPS, and income on three types of expenditures (education, food, and milk) are estimated next (table 4). Consistent with the enrollment results, for most households female power as measured by *rFPSY* has a positive relation with spending on education and a negative quadratic effect. Similar to the enrollment results, the maximum value of power for enrollment occurs at an *rFPSY* of about 4 (which applies to about 4 percent of the sample). Unlike education, spending on food or milk in particular does not show a statistically significant relation with power. However, all three expenditure categories show statistically significant impacts of the number of years of schooling of the female household head. The interactive effects of female

power and RPS on the three types of expenditures do not yield statistically significant coefficients.

The impacts of RPS on expenditures can be seen through two variables: *Treat*, which represents nonincome effects, and ln*Consumption*, which captures income effects measured through household consumption. Consumption of food, education, and milk increased with an increase in total household consumption, including consumption increases from RPS. Examination of the nonincome impacts of RPS as measured by the variable *Treat* shows significant positive impacts on spending for milk and food but not for education. The nonincome impacts on milk and food are substantial. The estimated nonincome impact on milk expenditures per capita is $C72, more than twice the average baseline consumption. The estimated impact of RPS on food consumption per capita is $C445, nearly a 15 percent increase over baseline consumption.

The empirical analysis yields three main results. First, more female power generally leads to higher school enrollment and greater spending on education. However, consistent with the emerging literature, for households with extremely powerful women, more female power may begin to reduce schooling or at least have no additional marginal impact. Second, nonincome effects of RPS are extremely important for school enrollment, which may not be surprising given the conditionality of the program. Nonincome impacts are evident on both food and milk per capita expenditures. Although the RPS program encourages spending on these items, such spending is not required, suggesting that nonincome effects other than conditionality had an impact. Two likely possibilities are the targeting of transfers to women and the nutrition education programs. Third, there is no evidence of a decreased impact of RPS on spending or schooling when women are less powerful.[11] Overall, these results support the hypothesis that the goals of school enrollment and nutrition can be improved by directing funds to women and requiring school attendance.

V. Conclusion

A large body of literature on intrahousehold bargaining suggests a positive relation between women's power and the amount of resources devoted to children. This article uses a power measure based on the ratio of years of schooling of female to male household heads to study the impacts of a conditional cash transfer program in Nicaragua. This measure is generally consistent with the expected positive relation between women's power and child schooling, although, as suggested by Basu's (2006) model, past a certain

---

11. In a separate model that omits the quadratic interaction term, *Treat**FPSY^2, the linear term is negative and statistically significant.

point the marginal impact of additional female power on children's enrollment may be zero or negative.

In targeting transfers to women, RPS and other conditional cash transfer programs seek to increase women's potential to spend money on children's schooling and other goods, such as food, that can improve children's human capital. The analysis provides evidence of the effectiveness of RPS transfers in improving the allocation of household resources toward women and children. The nonincome effects of the program are responsible for most of the nearly 20 percent increase in school enrollment; the targeting of transfers to women plays a secondary role.

Running the enrollment regressions separately for girls and boys reveals that the mother's relative education level always has a positive impact on boys' education outcomes. The results for girls are consistent with the nonlinear relation suggested by Basu (2006): when women's power passes a certain threshold girls' enrollment falls. Basu hypothesizes that parental power may influence the percentage of benefits from child labor garnered by each adult. This percentage may also depend on the child's gender. The nonmonotonic relation for girls but not boys suggests that when girls leave school, the percentage of the benefits received by the female head of household is larger than it is for boys.

The expenditure analysis supports the effectiveness of targeting transfers to women: RPS nonincome effects accounted for a more than doubling of milk expenditures and 15 percent of the increase in food expenditures. This effect may be shaped as much by women's education as it is by their power. However, the expenditure analysis shows that the education level of the female head has a positive impact on expenditures, but that the impact of their relative power is weaker. Overall, the empirical results suggest that targeting transfers to women has been effective at increasing key welfare outcomes for all households, even those with greater male power. But these estimates are inferences from econometric analyses and not direct measures of treatment effects of targeting transfers to women from a randomized experiment. If one goal of conditional cash transfer programs is to strengthen and broaden the quality of information regarding the efficacy of targeting transfers to women, more detailed questions on how households allocate their resources or possibly experiments that provide targeted and nontargeted transfers should be used in future program designs.

References

Adato, M., B. de la Briere, D. Mindek, and A.R. Quisumbing. 2003. "The Impact of PROGRESA on Women's Status and Intrahousehold Relations." In A.R. Quisumbing, ed., *Household Decisions, Gender, and Development: A Synthesis of Recent Research*. Washington, D.C.: International Food Policy Research Institute.

Attanasio, O., and V. Lechene 2002. "Test of Income Pooling in Household Decisions." *Review of Economic Dynamics* 5(4):720–48.

Basu, K. 2006. "Gender and Say: A Model of Household Behavior with Endogenously Determined Balance of Power." *Economic Journal* 116(511):558–80.

Basu, K., A. Narayan, and M. Ravallion. 2001. "Is Literacy within Households Shared?" *Labor Economics* 8(6):649–65.

Becker, G. 1981. *A Treatise on the Family.* Cambridge, Mass.: Harvard University Press.

Binder, M. 1999. "Schooling Indicators during Mexico's 'Lost Decade.'" *Economics of Education Review* 18(2):183–99.

Bourguignon, F., F. Ferreira, and P. Leite. 2003. "Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Brazil's Bolsa Escola Program." *World Bank Economic Review* 17(2):229–54.

Chiappori, P. 1992. "Collective Labor Supply and Welfare." *Journal of Political Economy* 100(3): 437–67.

de Janvry, A., F. Finan, E. Sadoulet, and R. Vakis. 2006. "Can Conditional Cash Transfers Serve as Safety Nets to Keep Children at School and out of the Labor Market?" *Journal of Development Economics* 79(2):349–73.

de Janvry, A., and E. Sadoulet. 2006. "Making Conditional Cash Transfers More Efficient: Designing for Maximum Effect of the Conditionality." *World Bank Economic Review* 20(1):1–29.

Doss, C. 1996. "Testing among Models of Intrahousehold Resource Allocation." *World Development* 24(10):1597–609.

Duflo, E. 2003. "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa." *World Bank Economic Review* 17(1):1–25.

Emerson, P., and A.P. Souza. 2007. "Child Labor, School Attendance, and Intrahousehold Gender Bias in Brazil." *World Bank Economic Review* 21(2):301–16.

Felkey, A. 2005. *Husbands, Wives, and the Peculiar Economics of Household Public Goods.* Working paper. Ithaca, N.Y.: Cornell University, Department of Economics.

Gitter, S.R., and B.L. Barham. 2006. "Conditional Cash Transfers, Credit, Shocks, and Education: An Impact Evaluation of Nicaragua's RPS." University of Wisconsin, Madison, Wisc.

Haddad, L., J. Hoddinott, and H. Alderman. eds. 1997. *Intrahousehold Resource Allocation in Developing Countries: Methods Models and Policy.* Baltimore, Md.: John Hopkins University Press.

Hoddinott, J., and E. Skoufias. 2004. "The Impact of PROGRESA on Food Consumption." *Economic Development and Cultural Change* 53(1):37–61.

Iyigun, M., and Randall P. Walsh. 2007. "Endogenous Gender Power, Household Labor Supply and the Demographic Transition." *Journal of Development Economics* 82(1):138–55.

Kanbur, R., and L. Haddad. 1994. "Are Better Off Households More Equal or Less Equal?" *Oxford Economic Papers* 46(3):445–58.

Lancaster, G., P. Maitra, and R. Ray. 2006. "Endogenous Intra-household Balance of Power and Its Impact on Expenditure Patterns: Evidence from India." *Economica* 73(291):435–60.

Maluccio, J. 2005. "Coping with the 'Coffee Crisis' in Central America: The Role of the Nicaraguan Red de Protección Social." FCND Discussion Paper 188. International Food Policy Research Institute, Washington D.C.

Maluccio, J., and F. Flores. 2004. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." FCND Discussion Paper 184. International Food Policy Research Institute, Washington D.C.

McElroy, M.B., and M.J. Horney. 1981. "Nash-Bargained Household Decisions: Toward Generalization of the Theory of Demand." *International Economic Review* 22(2):33–49.

Rawlings, L., and G. Rubio. 2005. "Evaluating the Impact of Conditional Cash Transfer Programs." *World Bank Research Observer* 20(1):29–55.

Schultz, T. 1990. "Testing the Neoclassical of Family Labor and Fertility." *Journal of Human Resources* 25(4):599–634.

———. 2002. "Why Governments Should Investment More in Girls' Education." *World Development* 30(2):207–25.

———. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74(1):199–250.

Thomas, D. 1990. "Intra-household Allocation: An Inferential Approach." *Journal of Human Resources* 25(4):635–64.

———. 1994. "Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height." *Journal of Human Resources* 29(4):950–88.

Does Aid for Education Educate Children?
Evidence from Panel Data

Axel Dreher, Peter Nunnenkamp, and Rainer Thiele

Most of the aid effectiveness literature has focused on the potential growth effects of aggregate aid, with inconclusive results. Considering that donors have repeatedly stressed the multidimensionality of their objectives, a more disaggregated view on aid effectiveness is warranted. The impact of aid on education is analyzed empirically for almost 100 countries over 1970–2004. The effectiveness of sector-specific aid is assessed within the framework of social production functions. The Millennium Development Goals related to education, particularly the goal of achieving universal primary school enrollment, are considered as outcome variables. The analysis suggests that higher per capita aid for education significantly increases primary school enrollment, while increased domestic government spending on education does not. This result is robust to the method of estimation, the use of instruments to control for the endogeneity of aid, and the set of control variables included in the estimations. JEL codes: F35, O11, H52, I22

There is heated debate over whether foreign aid is effective in promoting economic development. According to Sen (2006) and Tarp (2006), Easterly's (2006) claim that aid has done "so much ill and so little good" obscures the fact that aid can work if done right. Dalgaard, Hansen, and Tarp (2004) find that overall aid has indeed been effective. Even recent surveys of the literature on aid and growth come to sharply opposing conclusions. While Doucouliagos and Paldam (2005) conclude that the aid effectiveness literature has failed to

Axel Dreher (corresponding author) is a senior researcher at the KOF Swiss Economic Institute, ETH Zurich, and a research affiliate at the Institute for Economic Research, Center for Economic Studies (CESifo), at the University of Munich; his email address is mail@axel-dreher.de. Peter Nunnenkamp is Head of the Research Area Global Division of Labour at the Kiel Institute for the World Economy; his email address is peter.nunnenkamp@ifw-kiel.de. Rainer Thiele is Head of the Research Area Poverty Reduction, Equity, and Development at the Kiel Institute for the World Economy; his email address is rainer.thiele@ifw-kiel.de. The authors thank Christian Conrad, Martin Gassebner, Katja Michaelowa, participants at a seminar at the Kiel Institute for the World Economy, and three anonymous referees for their useful comments and suggestions, and Michaela Rank for excellent research assistance. A supplemental appendix to this article is available at www://wber. oxfordjournals.org

establish that aid works, McGillivray and others (2005) stress that almost all research published since the late 1990s finds that it does.

What both camps tend to ignore is that different types of aid are unlikely to have the same economic effects on recipient countries. In much of the literature, it is still common to run panel regressions with aggregate aid flows as the explanatory variable. Work by Clemens, Radelet, and Bhavnani (2004) on short-impact aid has initiated a shift toward using disaggregated aid data.

It is open to debate whether a verdict on the effectiveness of aid can be reached at all as long as the analysis is restricted to the aid-growth nexus. Donors have repeatedly stressed that they pursue multiple objectives when granting aid (see, for example, Isenman and Ehrenpreis 2003). Specific-purpose aid intended to support donors' policy statements, including the empowerment of the poor through better education, tends to escape analyses narrowly focused on aid and growth.

Against this backdrop, it seems appropriate to pursue a different avenue for assessing the effectiveness of aid. This article focuses on more specific outcome variables than growth. It uses disaggregated aid data to investigate the link between aid granted to the education sector and education outcomes.

Education figures prominently among the Millennium Development Goals. Donors have committed themselves to helping countries achieve universal primary education by 2015 and eliminate gender disparities in education. To this end, donors have devoted an increasing share of aid resources to the education sector (Thiele, Nunnenkamp, and Dreher 2007).

Yet, it is open to debate whether more resources necessarily translate into better education outcomes or how aid can play a role in achieving universal primary education (Roberts 2003). The effectiveness of aid for education is assessed here within the framework of social production functions. The results show that higher per capita aid significantly increases primary school enrollment. This outcome is robust to the method of estimation and to the set of control variables included.

The article is organized as follows. Section I provides the analytical background and discusses the literature on aid and education. Section II addresses data issues. The method of estimation and main results are presented in Section III, and various tests for robustness are performed in Section IV. The article closes with some concluding remarks about the implications of the findings for policy and future research.

I. ANALYTICAL BACKGROUND AND RELEVANT LITERATURE

A social production function with education-related outcomes as the left-hand-side variable is estimated in which aid for education enters as an explanatory variable. Schultz (1988) proposed a production–demand framework to model the educational system of countries in the 1980s. The concept of a social production function has also been used in the literature on the link

between government expenditure and social outcomes. Hanushek (1995, p. 2) considers this concept to be "most appealing and useful" to assess the relation between school outcomes and measurable educational inputs. Recent examples following this approach include Bennell (2002), Roberts (2003), and Baldacci, Guin-Siu, and de Mello (2003).

While the exact specification of social production functions varies, the common feature is that it includes major demand and supply factors (Roberts 2003). For the production function for education, demand factors typically considered include per capita income (a proxy for household poverty), adult literacy (a proxy for the educational status of parents), the relative size of the school population, and the level of urbanization. The "price" of schooling also affects demand, although it is typically not included in empirical cross-country analyses because of lack of data. School fees are supposed to inhibit enrollment, which is why free universal primary education has been advocated (see, for example, Bruns, Mingat, and Rakotomalala 2003).[1]

Public spending on education figures prominently among the supply factors considered in social production functions. Other potentially relevant supply factors include the pupil–teacher ratio and the unit costs of education. The regression analysis conducted below extends the production function concept by adding aid for education as an additional supply factor.

Various studies find that demand factors explain most of the variance in school attendance (enrollment, completion rates) and educational attainment (youth literacy, test scores) across countries. Surveying the literature, Roberts (2003) concludes that per capita income tends to be the most powerful driving force of school performance; supply-side factors, in particular public expenditure on education, are statistically insignificant in most instances.[2] Roberts' own cross-country regression analysis corroborates the finding of ineffective public expenditure and finds that adult literacy is the main demand-side factor. Clemens (2004) shows that school enrollment rises only slowly over time and that the impact of education policy is relatively small compared with that of long-term economic changes.[3]

Very few studies consider foreign aid for education as a possibly important supply-side factor in the production function; studies that do so (Michaelowa and Weber 2006; Wolf 2007) are inconclusive. The aid literature focuses on whether aid helps achieve economic and social objectives by providing

---

1. The success of the Mexican anti-poverty program Progresa, in which educational subsidies in the form of conditional transfer payments to poor families increased enrollment, provides evidence for this notion (Behrman, Sengupta, and Todd 2005).

2. See Clemens (2004) and the references given there. Cross-country studies include Filmer and Pritchett (1999), Mingat and Tan (2003), and Baldacci, Guin-Siu, and de Mello (2003).

3. While this rather bleak picture concerning the ability of government spending on education to raise education outcomes appears to represent the majority view in the literature, there are some notable exceptions, including Gupta, Verhoeven, and Tiongson (1999) and Baldacci and others (2004).

additional resources for financing pro-poor expenditure and on the extent to which aid is fungible. Although Pettersson (2006) finds a high degree of fungibility, some other studies show the aid–expenditure link to be important. Gomanee and others (2003) and Mosley, Hudson, and Verschoor (2004) find that aid alleviates poverty through its effect on public expenditure. Gomanee and others (2005) reach the opposite conclusion in a more recent version of their article.

The effects of aid working through public expenditure could be captured by estimating a system of equations that includes a public expenditure equation with aid as one explanatory variable (Mosley, Hudson, and Verschoor 2004).[4] However, such an approach suffers from several problems. It is conceptually demanding, because the specification of the equations should ideally be based on a complete theoretical model, and the determinants of all dependent variables would have to be included in the estimations. The more conventional approach of instrumenting potentially endogenous variables in the production function for education has the advantage that instruments need explain only some fraction of the variation in the instrumented variables.

Furthermore, the interpretation of coefficients in the public expenditure equation is plagued with problems, particularly regarding the aid coefficient. A coefficient that is not significantly different from zero does not necessarily imply that aid is highly fungible and that aid does not add to overall (foreign and domestic) resources devoted to education. This implication would hold only if (most) aid were accounted for in the public budget of the recipient country.[5] However, project aid for education—the most important mode of aid delivery, at least until recently—often remains outside the budget (Roberts 2003). If all aid remain outside the budget, full fungibility implies an aid coefficient of −1. As it is impossible to determine the proportion of aid outside and inside the budget, estimation of the public expenditure equation offers no meaningful insights. For this reason, estimating a system of equations is not the preferred option.

The approach taken here follows the seminal contribution of Borenszstein, de Gregorio, and Lee (1998) on the economic growth effects of foreign direct investment (FDI). They consider both foreign and domestic investment in assessing whether foreign investment is more productive in raising growth. Analogously, an enrollment equation is estimated here that includes both aid for education and government expenditure on education as explanatory variables.

Aid for education may be more effective in raising enrollment rates than government expenditure on education, for several reasons (Roberts 2003).

---

4. This approach was adopted in an earlier version of this article.

5. The authors thank an anonymous referee for alerting them to the critical importance of this assumption.

First, at least 75 percent of government expenditure typically consists of teacher salaries, which were not covered by donors until recently (Michaelowa and Weber 2006). Donors provide aid for building schools, supplying teaching materials, improving school management, and reforming curricula, in the hope of improving the learning environment, the efficiency of schools, and the quality of education, which may provide stronger incentives to attend school.

Second, government expenditure on education is often biased against the poor, the population segment for which enrollment rates tend to be relatively low (Bennell 2002). This contrasts with donor strategies emphasizing poor and disadvantaged groups, in particular girls, for whom school attendance is often lower than for boys.

Third, leakage of local funds appears to be high (Reinikka and Svensson 2001) and capture by producers and privileged consumers to be common. External donors may succeed at least partly in mitigating leakage and capture of aid funds by not channeling aid through the public budget of the recipient country, by involving local authorities in aid allocation processes, and by increasingly applying performance-based allocation rules. Measures such as these are recent, however; before the 1990 Jomtien Conference, which emphasized the importance of universal primary education, donor support concentrated on higher levels of education (Thiele, Nunnenkamp, and Dreher 2007). Consequently, it remains an empirical question whether foreign aid has been more effective than domestic public spending on education in promoting primary school enrollment.

II. Data Issues

The data for assessing the impact of aid for education on education outcomes are far from perfect. The aid data—on commitments of sector-specific aid, including aid for education—come from the Creditor Reporting System (CRS) of the Organisation for Economic Co-operation and Development/ Development Assistance Committee (OECD/DAC). These data are imperfect because commitment data tend to overstate actual aid flows (commitments may not be fully disbursed) and because sector-specific commitments go partly unreported (Michaelowa and Weber 2006). These measurement problems, which work in opposite directions, cannot be resolved, because sector-specific disbursement data are not available before 1990. The correlation between commitments and disbursements of aid for education over the period for which both series are available is fairly high, with correlation coefficients of 0.70 for 1990–94, 0.71 for 1995–99, and 0.80 for 2000–04.

It can be argued that employing sector-specific aid data understates the contribution of aid to education objectives in recent years. Several donors now favor general budget support over project aid for specific targets. The extent to which general budget support is ultimately used for educational objectives is not known.

A similar argument can be made about multisector aid. However, the evaluation of the composition of aid in Thiele, Nunnenkamp and Dreher (2007) suggests that this is unlikely to pose serious problems. In contrast to donor announcements, the shares of general budget support and multisector aid in total aid were actually lower in 2002–04 than in the early 1990s. Nevertheless, the robustness of results is checked by replicating the estimates for a shorter period of observation (excluding recent years, in which donors may have increasingly supported educational objectives through aid that is not picked up in sector-specific aid data).

Data limitations with respect to education outcome variables are well known (Roberts 2003; Bennell 2002). Ideally, the outcome variable of the production function should go beyond enrollment rates to include educational attainment and the quality of education. Enrollment rates may provide a misleading picture. Clemens (2004) draws on detailed country studies to show that rising enrollment rates came at the cost of deteriorating quality of education in some countries, as reflected by much higher pupil–teacher ratios, higher failure and repetition rates, and lower test scores. Furthermore, some countries report unreasonably high net enrollment ratios (more than 100).

Qualitative dimensions of education, such as improved literacy and test scores, are not available for a sufficiently large number of countries over a sufficiently long period of time. However, distortions resulting from the shortcomings of enrollment rates as an education outcome variable were minimized in several ways. First, completion rates were considered as an alternative indicator. Second, near universal enrollment rates were checked against reported ratios of boys to girls (enrollment rates of almost 100 percent are inconsistent with gender imbalance): except in Tajikistan, no major discrepancy was detected. Third, additional estimates were run for a reduced sample, eliminating countries with exceptionally large increases in enrollment rates.

Another data problem concerns the time-series dimension. In 2003, the United Nations Educational, Scientific, and Cultural Organization (UNESCO), the original source of the World Bank data used here, revised its estimates of net primary enrollment for 1998–2001. For some countries, this revision is associated with a major break in the series on primary enrollment. These countries were identified by comparing the old and new data in years for which both series are available (normally 1998–2000) (see Clemens 2004 for a similar approach). Discrepancies were minor (less than 2 percentage points) in 69 of the 119 sample countries for which this comparison was possible. The revision resulted in major discrepancies (more than 10 percentage points) in 15 countries, in 8 of which the old series appears to have overstated enrollment rates. This problem was dealt with in the tests for robustness by replicating the analysis for a shorter period of observation, 1970–97. In this way, the risk of inconsistencies over time can be reduced, even though the old series may suffer from systematic over- or underreporting by some countries.

III. Method of Estimation and Base Results

Net primary school enrollment is the dependent variable throughout this section. The main explanatory variables of interest are aid to the education sector and government spending on education. (Using aid and spending on primary education is conceptually superior but would leave an insufficient number of observations for estimation.) Aid and government expenditure are measured on a per capita basis. In aid-growth regressions, aid is typically defined relative to the recipient country's gross domestic product (GDP). This provides a reasonable measure of the importance of foreign support relative to the recipient country's overall resources. Aid per capita is more appropriate than the aid to GDP ratio in assessing aid effectiveness with respect to specific Millennium Development Goals. In particular, achieving universal primary education requires accounting for the number of people among whom the resources devoted to education must be shared.

For other relevant covariates, the analysis closely follows the literature on education production functions in considering four demand-side variables: adult literacy, per capita GDP,[6] share of the population under 15, and share of the urban population in total population. Lagged education outcome is included as an additional explanatory variable in order to account for the possible persistence in outcomes. Additional supply-side variables suggested in the literature are added later to test for the sensitivity of results.

Pooled time-series cross-section (panel data) regressions are estimated for a maximum of 96 low- and middle-income countries between 1970 and 2004 (with the exception of data on aid, which are available only since 1973).[7] As some of the data are not available on an annual basis, all data are five-year averages. (The definitions of and sources for all variables are listed in table A-1; summary statistics are reported in table A-2.)

The basic equation takes the following form:

$$\text{school}_{i,t} = \alpha + \beta_1 \text{school}_{i,t-1} + \beta_2 \text{aid}_{i,t} + \beta_3 \text{spending}_{i,t} + \text{BX} + \eta_i + \varepsilon_{i,t} \quad (1)$$

where $\text{school}_{i,t}$ represents the logarithm of primary school enrollment in country i in year t; $\text{aid}_{i,t}$ is per capita foreign aid to the education sector; and $\text{spending}_{i,t}$ is per capita government expenditure on education. X is the vector of control variables, η_i represents country fixed effects, and $\varepsilon_{i,t}$ represents the disturbance term.

The dependent variable, $\text{school}_{i,t}$, is limited by a lower bound of zero and an upper bound of 100. The upper bound may lead to biased results in the sense

6. Following most earlier studies estimating social production functions (for example, Gupta, Verhoeven, and Tiongson 1999; Roberts 2003; and Michaelowa and Weber 2006), per capita GDP is measured in levels.

7. High-income countries were excluded because they receive no aid. The World Bank (2007) defines high-income countries as those with a 2005 GNI per capita of at least $10,726.

that aid can have little effect on enrollment in recipient countries with enrollment rates of close to 100 percent.[8] One way to deal with this problem is through logistic rather than log linear estimation (Fielding, McGillivray, and Torres 2005). A different route is taken here. While the analysis follows the standard approach of the literature, which includes enrollment either in levels or in logs, it tests robustness by excluding recipient countries where enrollment rates exceed a certain threshold and may thus bias results downward.

Aid cannot reasonably be expected to be exogenous to school enrollment: donors typically grant more aid to countries that are less developed. Nevertheless, as a first step fixed- and random-effects models that ignore the potential endogeneity are estimated before we present specifications that allow for the endogeneity of aid and government expenditure.[9]

The qualitative findings do not depend on the inclusion of random or fixed country effects. However, the random-effects specification is rejected in favor of the fixed-effects model (Hausman test, $P = 0.00$), so only fixed-effects estimates are reported (table 1, column 1). An F-test also shows that fixed country effects cannot be omitted ($P = 0.00$). By contrast, fixed-period effects did not turn out to be significant, so the estimates do not include them.

The estimates reveal a considerable degree of inertia in primary school enrollment. The lagged dependent variable is highly significant.[10] Clearly, it has some explanatory power, rendering insignificant most of the other covariates included in the regression. The fixed-effects model indicates that literacy, the share of the population under the age of 15, and the degree of urbanization do not significantly affect enrollment.[11]

Most surprisingly perhaps, per capita GDP does not have the expected positive impact on school enrollment, even when per capita GDP is alternatively specified in logs (not shown). As noted earlier, cross-country evidence suggests that per capita income is an important determinant of enrollment.[12] By contrast, the explanatory power of per capita GDP is rather weak in the few panel studies that have been conducted. The insignificance of per capita GDP in table 1 is similar to the results obtained by Michaelowa and Weber (2006) as

---

8. The authors thank an anonymous referee for pointing this out.

9. The estimator is biased and inconsistent in a short panel because of the inclusion of the lagged dependent variable and fixed- or random-country effects. The GMM estimator presented later takes this into account.

10. A second lag of the dependent variable turned out to be completely insignificant.

11. Estimates based on a panel-corrected standard error model (not shown) point to a very strong relation between adult literacy and enrollment across countries, corroborating the findings of Gupta, Verhoeven, and Tiongson (1999) and Roberts (2003).

12. Even among cross-country studies, there are notable exceptions. Gupta, Verhoeven, and Tiongson (1999) do not find a robust effect of per capita GDP on enrollment. Mingat and Tan (1998) show that the cross-country correlation between per capita GDP and education is relatively weak when higher-income countries are excluded from the calculation, as they are in this article.

TABLE 1. School Enrollment and per Capita Aid to Education, 1970–2004

Variable	(1)	(2)	(3)	(4)
Lagged dependent	0.38 (3.82)***	0.65 (6.99)***	0.44 (4.63)***	0.40 (3.52)***
Expenditure on education (per capita)	0.000034 (0.37)	−0.000198 (1.47)	0.000220 (0.75)	0.001523 (1.04)
Aid for education (per capita)	0.0026 (2.89)***	0.0029 (2.34)**	0.0268 (2.69)***	0.0243 (2.24)**
Literacy rate	0.0034 (1.35)	0.0019 (1.75)*	0.0046 (1.19)	0.0081 (1.43)
GDP per capita	−0.000003 (0.61)	0.000008 (1.42)	0.000028 (1.29)	0.000003 (0.08)
Population under 15	0.001 (0.34)	0.001 (0.88)	0.019 (2.16)**	0.020 (2.17)**
Urbanization	−0.00022 (0.09)	−0.00015 (0.30)	0.00420 (0.97)	0.00035 (0.05)
Number of countries	94	94	61	61
Number of observations	267	267	156	156
Maximum number of periods	6	4	4	4
Method	Fixed effects	System GMM	2SLS	2SLS
R-squared (overall)	0.93		0.41	0.40
Number of instruments		53	9	9
Hansen test (Prob > χ^2)		0.60	0.11	0.14
Difference-in-Sargan test (Prob > χ^2)		0.17		
Arellano-Bond test (Prob > z)		0.24		
First-stage F-test			6.32	5.55

Note: Numbers in parentheses are *t*-statistics, estimated robustly. The dependent variable is the logarithm of primary school enrollment in country *i* at year *t*.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Source: Authors' analysis based on data described in table A-1.

well as Baldacci and others (2004).[13] The weak explanatory power of per capita GDP in panel studies may at least partly reflect the fact that these studies capture short- to medium-run effects through the time dimension, whereas the effects in cross-section analyses are purely long term.

The results show a positive correlation between aid and school enrollment, with coefficients significant at the 1 percent level. As in much of the previous literature, government expenditure on education does not affect enrollment significantly. Because both government expenditure and aid enter the regression, the significant coefficient of aid suggests that it is more productive than domestic government spending in raising school enrollment.

In the next step, the potential endogeneity of aid and government expenditure is taken into account (columns 2–4), starting with the system generalized method of moments (GMM) estimator, as suggested by Arellano and Bover (1995) and Blundell and Bond (1998). The dynamic panel GMM estimator exploits an assumption about the initial conditions to obtain moment conditions that remain informative even for persistent data and is considered most appropriate in the presence of endogenous regressors. Results are based on the two-step estimator implemented by Roodman (2005) in Stata, including Windmeijer's (2005) finite sample correction. Aid and government expenditure are treated as endogenous and the additional covariates as strictly exogenous. The validity of these assumptions was tested by applying the Hansen test (amounting to a test for the exogeneity of the covariates) and the Arellano-Bond test of second-order autocorrelation, which must be absent from the data for the estimator to be consistent. Both tests turned out to be borderline when the first lags of the aid variable and the dependant variable are included as instruments. This may suggest that aid once lagged is endogenous and thus invalid as an instrument. Therefore, the first lags were excluded from the list of instruments. The test statistics do then clearly not reject the specification at conventional levels of significance.[14] A difference-in-Sargan test was does also performed on the additional instruments in the system GMM. This test also does not reject the specification.

The results remain qualitatively unchanged when the system GMM estimator is employed (column 2 in table 1).[15] Results obtained by rerunning the regression using the Arellano and Bond (1991) GMM difference estimator (not reported) corroborate the main results of the system GMM estimation, although the literacy rate no longer significantly affects school enrollment.

13. Baldacci and others (2004), who do not include aid as a regressor, find per capita GDP to be a significant determinant of enrollment only in some of their estimated equations.

14. Additional tests were performed for third- and fourth-order serial correlation, which does not seem to be present in the data.

15. Excluding the first lag of the endogenous variables reduces the number of instruments to 53. The regression was replicated excluding the second and third lags to ensure that the results do not depend on this still substantial number of instruments. This step reduced the number of instruments to 28. The relevant test statistics do clearly not reject the specification, and the results are not affected. In particular, the difference-in-Sargan test (Prob $> \chi^2$) is 0.42.

Two-stage least squares (2SLS) regressions are estimated instrumenting for aid as an alternative procedure for addressing the endogeneity issue. The International Country Risk Guide, the Fraser index of economic freedom, and the mortality rate of children under age five serve as instruments. These variables—the first two proxies for governance, the third a proxy for need—have been shown to be related to aid allocations (McKinlay and Little 1977; Hout 2004; Thiele, Nunnenkamp, and Dreher 2007). They are indeed highly correlated with aid for education in the sample and not significantly correlated with school enrollment once the other relevant regressors are controlled for.[16] Government expenditure on education is taken as an exogenous regressor in column 3; both government expenditure and aid are instrumented in column 4. The International Country Risk Guide index for ethnic tensions is used as an instrument for government expenditure on education. While the previous literature does not come up with reliable instruments for government spending (Feldmann forthcoming), ethnic tensions have been shown to affect expenditure.[17]

The overidentifying restrictions are not rejected at conventional levels of significance, suggesting that the model is well specified, and there is no sign that the instruments are endogenous. The instruments are jointly significant at the 1 percent level in the first-stage regressions, indicating some power. However, the *F*-test statistic falls short of the rule of thumb threshold of 10 proposed by Staiger and Stock (1997). Nelson and Startz (1990) show that the distribution of the 2SLS estimator and the *t*-statistics are only poorly approximated by the asymptotic representation with weak instruments. This is likely to imply test statistics with nonnormal distributions, making 2SLS misleading (Staiger and Stock 1997). According to Cruz and Moreira (2005), weak instruments exacerbate the finite sample bias inherent in the 2SLS estimator. Given that the specifications include the lagged dependent variable, 2SLS estimations may also suffer from dynamic panel bias. These estimations, therefore, have to be interpreted with caution. As all relevant test statistics do not reject the GMM specification, the preferred results are those based on the GMM regression (column 2 of table 1).

The effect of aid for education on primary school enrollment remains positive and significant at least at the 5 percent level when the two instrumental variable techniques are used. The lagged dependent variable is significant at the 1 percent level; per capita GDP and urbanization are insignificant. As before, so is government expenditure on education. The results on adult literacy are mixed. The share of the population under 15 has a completely

16. Initially, a measure of democracy was also included in the list of instruments. However, given that democracy might be important for the effect of aid on enrollment (as argued below), only the former three variables are used. The results are not affected by this choice.

17. Von Hagen (2005) argues that ideological and ethnic divisions result in higher government spending because each segment of society tends to neglect the tax burden falling on other segments, exacerbating the common pool problem.

insignificant impact when estimated with GMM and a positive and significant impact in the potentially misspecified 2SLS regressions. The 2SLS results may appear counterintuitive, because having a larger share of the population under 15 would put pressure on the education system and would thus be expected to reduce enrollment rates. However, the positive coefficient in the 2SLS estimation is in line with the empirical evidence reported by Schultz (1988), according to whom a larger share of school-age children affects various educational variables negatively but primary school enrollment positively.

The quantitative impact of aid on school enrollment is modest but not negligible. With aid measured in per capita terms and the dependent variable in logarithms, the GMM specification implies that a $1 increase in aid (roughly equivalent to the increase in average aid for education between 1975–79 and 1980–84 across the sample countries) increases school enrollment by about 0.29 percent. School enrollment would increase by 1.5 percent if donors kept their promise to double aid efforts (from an average of $5 per capita to the education sector in 2000–04). This increase is roughly equal to the increase in average enrollment between 1995–99 and 2000–04.

The fixed-effects estimate implies a slightly smaller increase in school enrollment, of about 0.26 percent for each additional dollar of aid. By contrast, the quantitative impact indicated by the 2SLS regressions appears to be unrealistically high, with an additional dollar of per capita aid raising school enrollment from 2.4 to 2.7 percent.[18]

For comparison, regressions were also run with aid and government expenditure measured as a percentage of GDP rather than in per capita terms. In the fixed-effects specification, the qualitative results are very similar: aid increases primary schooling at the 5 percent level of significance. In the preferred GMM specification, however, aid does not significantly affect school enrollment when measured as a percentage of GDP. This finding supports the reasoning that the aid to GDP ratio is not as good a measure for assessing aid effectiveness with respect to specific Millennium Development Goals as per capita aid. (The results are shown in table S-1 in the supplemental appendix to this article, available at http://wber.oxfordjournals.org/)

Given the inclusion of the lagged dependent variable, it is also possible to calculate the long-run effect of aid on school enrollment. The estimates reported in column 2 of table 1, the preferred GMM specification, and coefficients of 0.003 for aid and 0.65 for lagged aid suggest that the long-term effect of an additional dollar of per capita aid is to raise school enrollment about 0.8 percent. The analysis was replicated excluding the population under the age of 15 as well as urbanization and per capita GDP (both of which were insignificant in the preferred specifications). The results do not change qualitatively or quantitatively: aid still increases enrollment significantly under all

---

18. A second lag of the dependent variable was included to test whether this affected the results. The coefficient turned out to be unaffected both quantitatively and qualitatively.

specifications. Government expenditure on education remains completely insignificant (see table S-2 in the online supplemental appendix).

IV. SENSITIVITY ANALYSIS

This section presents extensions of and tests for robustness of the previous results, focusing on GMM specifications (table 2).[19] It begins by considering different aid measures. Ideally, one would employ aid disbursements for primary or basic education. However, even commitment data on aid for primary education are scarce, leaving an insufficient number of observations for meaningful estimation. Disbursements of total aid for education are available only since 1990. Moreover, inspection of the data suggests substantial underreporting of disbursed aid for education by donors until recently.[20] In light of considerable measurement error, it is not surprising that the coefficient of aid disbursements remains insignificant (column 1).

Total aid disbursements, for which data are available for the entire period of observation, can be used instead of aid for education (column 2). But this aid variable also turns out to be insignificant in both the fixed-effects and GMM specifications. The finding that total aid has no impact on enrollment is not surprising, as aid for education, albeit of rising relative importance, accounted for just 8.2 percent of total aid in 2002–04 (Thiele, Nunnenkamp, and Dreher 2007).

School enrollment still rises with aid (at the 5 percent level of significance) when aid commitments per child below the age of 15 are used rather than commitments per capita (column 3). Primary completion rates have been suggested as superior to enrollment rates in measuring progress toward education-related Millennium Development Goals. However, data on primary completion are available only since 1988, reducing the sample size considerably. The small sample size may explain why aid for education does not significantly affect completion rates at conventional levels of significance (column 4).

In various other respects, the results prove fairly robust. In additional tests for robustness government expenditure on education is dropped from the equation, levels of school enrollment are used instead of logs, and the sample period is restricted to years before 1998. The concern that results may be biased because the dependent variable is bounded is addressed above, and two sets of outlying countries are omitted. The robustness of the results to the inclusion of additional variables suggested in the literature and to the exclusion of the lagged dependent variable is also tested. In most of these tests, the key result remains unaffected.

19. Corresponding fixed-effects estimates are shown in supplemental appendix table S-3.
20. Only in 2001–04 did disbursed aid exceed 50 percent of committed aid for education. Reported disbursements represented less than 10 percent of commitments in the early 1990s.

TABLE 2. School Enrollment and Aid: Tests for Robustness, System General Method of Moments, 1970–2004

Variable	(1) School enrollment	(2) School enrollment	(3) School enrollment	(4) Completion rates	(5) School enrollment	(6) School enrollment levels	(7) School enrollment 1970–97	(8) School enrollment[a]	(9) School enrollment[b]	(10) School enrollment[c]	(11) School enrollment	(12) School enrollment	(13) School enrollment	(14) School enrollment	(15) School enrollment
Lagged dependent	0.845 (7.36)***	0.530 (5.40)***	0.65 (8.46)***	0.821 (8.30)***	0.603 (5.20)***	0.516 (4.46)***	0.617 (6.42)***	0.574 (7.48)***	0.588 (5.98)***	0.617 (6.42)***	0.631 (5.26)***	0.626 (6.33)***	0.602 (7.05)***	0.594 (5.47)***	0.597 (6.50)***
Expenditure on education (per capita)	−0.00006 (0.60)	0.00007 (0.69)	−0.00009 (1.21)	0.00023 (1.41)		−0.00102 (0.16)	−0.00008 (0.94)	−0.00534 (0.73)	−0.00004 (0.57)	−0.00008 (0.94)	−0.00001 (0.10)	−0.00002 (0.24)	−0.00005 (0.74)	−0.00010 (1.16)	−0.00009 (1.07)
Aid for education, disbursements (per capita)	0.0033 (0.89)														
Total aid disbursements (per capita)		0.00004 (0.32)													
Aid for education commitments (per child)			0.0008 (2.47)**												
Aid for education commitments (per capita)				−0.002 (0.46)	0.002 (2.10)**	0.182 (1.79)*	0.003 (2.11)**	0.004 (1.75)*	0.002 (1.79)*	0.003 (2.11)**	0.002 (3.37)***	0.003 (2.54)**	0.001 (0.29)	0.003 (2.51)**	0.009 (0.97)
Literacy rate	0.001 (0.52)	0.003 (2.55)**	0.002 (2.28)**	0.001 (0.78)	0.002 (1.72)*	0.224 (2.77)***	0.002 (2.23)**	0.002 (2.30)**	0.003 (2.27)**	0.002 (2.23)**	0.002 (2.65)***	0.003 (2.30)**	0.003 (2.19)**	0.003 (2.23)**	0.003 (2.43)**
Pupil-teacher ratio											4.13E − 04 (0.23)				
Unit costs of enrollment												−0.002 (1.63)			
Aid-squared													0.00003 (0.61)		
International crisis														0.16 (1.91)*	
Democracy															0.004 (0.60)

Democracy*aid															−0.001 (0.78)
Number of countries	94	96	96	92	99	96	96	90	87	96	89	95	96	96	96
Number of observations	199	269	269	194	306	269	269	221	238	269	176	268	269	269	269
Maximum number of periods	3	6	6	3	6	6	6	6	6	6	3	6	6	6	6
Number of instruments	21	42	42	24	36	42	42	42	42	42	34	43	62	62	44
Hansen test (Prob > chi^2)	0.14	0.58	0.35	0.43	0.25	0.32	0.42	0.31	0.45	0.42	0.34	0.43	0.39	0.43	0.30
Differences-in-Sargan test (Prob > chi^2)	0.06	0.92	0.13	0.94	0.28	0.19	0.41	0.21	0.40	0.21	0.24	0.18	0.73	0.16	0.25
Arellano-Bond test (Prob > z)	.	0.12	0.18	.	0.06	0.61	0.19	0.29	0.82	0.19	.	0.20	0.19	0.28	0.18

Note: Numbers in parentheses are *t*-statistics, estimated robustly. The dependent variable is in logarithms.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

[a]Excludes countries in the highest enrollment quartile.

[b]Excludes Botswana, Indonesia, Malawi, Rwanda, Togo, and Uganda (poor countries for which enrollment rates have risen particularly rapidly, according to Clemens 2004).

[c]Excludes Angola, Kuwait, Liberia, Lesotho, Malawi, and Tanzania (reported enrollment rates increased by more than 20 percent in a single year and by 10 percentage points at least once over the period under consideration).

Source: Authors' analysis based on data described in table A-1.

The reason to exclude government expenditure on education is that the aid coefficient may be biased downward when government expenditure is included and (part of) aid runs through the budget. The aid coefficient is supposed to capture the expenditure-augmenting effect of aid in addition to effects on enrollment that are attributable to a higher productivity of aid relative to government expenditure. The fact that the coefficient of aid is hardly affected when comparing column 5 in table 2 with column 2 in table 1 suggests that the expenditure-increasing effect of aid is not relevant, either because aid is not accounted for in the budget or because it is highly fungible.

Using enrollment levels instead of logs (column 6) leaves aid significant at the 10 percent level. One additional dollar of per capita aid increases school enrollment by about 0.2 percentage points. The years after 1997 are excluded because the results may be distorted by the revision of educational data since 1998 and by the recent shift from sector-specific aid, including aid for education, toward general budget support and multisector aid. Yet the key result is not affected: aid increases enrollment at the 5 percent level of significance (column 7).

The results reported in column 8 are based on a restricted sample that excludes all countries in the highest quartile in enrollment rates. In this way, a check can be run to determine whether the upper bound of the dependent variable implies a downward bias for aid effectiveness when including recipient countries with enrollment rates already close to 100 percent. The aid coefficient turns out to be somewhat larger for the restricted sample compared with column 2 in table 1. The same pattern applies to the fixed-effects estimates (column 8 in supplemental appendix table S-3 and column 1 in table 1). Standard errors are larger for the restricted sample, which is not surprising given the smaller variance of enrollment rates and the smaller number of observations.

Next some potentially influential outliers are excluded, though to little effect. Botswana, Indonesia, Malawi, Rwanda, Togo, and Uganda are treated as outliers (column 9). Clemens (2004) identifies these countries as examples of poor countries for which enrollment rates reported by UNESCO have risen particularly rapidly, at least in some cases at the cost of deteriorating education quality (as reflected in high failure and repetition rates in Rwanda and Togo, steeply rising pupil–teacher ratios in Malawi, and lower test scores in Uganda). Alternatively, Angola, Kuwait, Liberia, Lesotho, Malawi, and Tanzania are excluded (column 10), because reported enrollment rates increased by more than 20 percent in a single year and by 10 percentage points at least once over the period under consideration.[21]

---

21. Both relative and absolute changes are considered when defining the cut-off point, because the two deviate widely at the tails of the distribution. A moderate absolute rise in enrollment from 20 to 25, for example, implies a relative increase of 25 percent.

Alternative specifications include the pupil–teacher ratio, the unit cost of education (government expenditure on education divided by the population under 15, as a percent of per capita GDP), and a dummy variable for a crisis of international scale in the recipient country as additional control variables (columns 11, 12, and 14). The pupil–teacher ratio and the unit cost of production are applied as additional supply factors in the social production function framework.[22] The dummy variable for international crisis is included because enrollment might be expected to decline in such years and crises might be correlated with aid.

With the exception of the crisis dummy variable, none of the additional variables is significant at conventional levels. In all cases, the impact of aid remains significant at the 5 percent level. Aid-squared has often been used in the literature on aid and growth (see, for example, Dalgaard, Hansen, and Tarp 2004), because aid may suffer from decreasing returns once its optimal degree is exceeded. Thus aid-squared is included in column 15. Although aid and aid-squared are not individually significant, they are jointly significant at the 5 percent level.

Finally, the possibility that the impact of aid on school enrollment might depend on democratic governance in the recipient countries is taken into account. This issue relates to the ongoing discussion of whether donors should target aid to better-governed countries. According to Svensson (1999), democratic institutions provide an institutionalized check on governments, encouraging them to use aid more productively. The impact of aid on education outcomes is thus hypothesized to be greater the greater the degree of democracy. The test of whether aid is more effective under conditions of good governance treats governance as exogenous to aid, which is in line with much of the previous literature.[23] The interaction between aid and democracy is supposed to reveal a differential impact of aid. As a proxy for democratic governance, an index is constructed with data provided by Freedom House (2004). Data on this variable are available for a large number of countries and over most of the years under study.

While foreign aid does not increase enrollment individually in a significant way, aid and its interaction with democracy are jointly significant at the 10 percent level (column 15). The level of democracy and its interaction with aid are neither individually nor jointly significant at conventional levels.

---

22. Ideally, one would also want to control for the "price" of schooling. The lack of data does not permit this to be done.

23. The seminal contribution of Burnside and Dollar (2000) triggered the debate on whether aid is more effective under good policy conditions, treating the policy variables (openness, inflation, budget surplus) as exogenous. In a subsequent article, Burnside and Dollar (2004, p. 4) note that "researchers coming from the left, the right, and the center have all concluded that aid as traditionally practiced has not had systematic, beneficial effects on institutions and policies." Clemens, Radelet, and Bhavnani (2004) use instruments for institutional and policy variables; their use does not affect their results significantly.

It therefore appears that, in contrast to the hypothesis derived by Svensson (1999), the impact of aid does not depend on democracy. Aid for education may help achieve universal education even in countries characterized by less-advanced democratic institutions. Whether this result is robust to the specification of the model and the measurement of democracy is an interesting question for future research.[24]

V. CONCLUSIONS

The effectiveness of sector-specific aid on education is assessed empirically within the framework of social production functions for almost 100 developing economies over the period 1970–2004, with the education Millennium Development Goals, particularly primary school enrollment, considered as outcome variables. The results suggest that higher per capita aid significantly increases primary school enrollment and domestic government spending on education does not. This result is robust to the method of estimation, the inclusion of instruments to control for the endogeneity of aid, and the set of control variables included in the estimations.

These findings are in sharp contrast to Easterly's (2006) verdict that foreign aid has done "so little good." At the same time, the analysis underscores the need to disaggregate aid in order to assess its effectiveness. Aid specifically devoted to the education sector modestly but not negligibly contributes to achieving universal primary education in developing economies. The preferred (GMM) specification (column 2 in table 1) implies that an additional dollar of per capita aid to the education sector increases school enrollment by about 0.3 percent. Consequently, school enrollment could improve considerably if donors kept their promise to double current aid efforts.

Aid that is effective in improving education should also have favorable long-term effects on economic growth, which might not be measurable with conventional econometric methods.[25] But even if the link between education and growth turned out to be weak, the improved education outcome would be important in its own right, because "schooling has a large number of direct beneficial effects beyond raising economic output, such as lower child mortality" (Pritchett 2001, p. 388).

The positive effects of aid on education outcomes in recipient countries notwithstanding, the analysis points to some caveats that donors should keep in mind when giving aid. In contrast to some other studies (Gomanee and others 2003; Mosley, Hudson, and Verschoor 2004), this study finds no evidence that aid works by increasing government spending on education. Estimates are

---

24. In column 16 of supplemental appendix table S-3, results are also reported after omitting the lagged dependent variable from the fixed-effects specification. The main result is unchanged. Under this specification, the literacy rate positively affects enrollment, with a highly significant coefficient.

25. The longer-term growth effects of aid are difficult if not impossible to capture, as Clemens, Radelet, and Bhavnani (2004) argue. For a different view, see Rajan and Subramanian (2005).

hardly affected when accounting for both the productivity-enhancing and the expenditure-augmenting effects of aid (by excluding government spending from the estimations). It remains open to debate whether this is mainly because donors deliberately decided to grant aid outside the budget or because aid was highly fungible in the past. This question may be resolved in future research if donors increasingly shift from project-related aid to general budget support. For the time being, the complete insignificance of government spending on education in the estimations cautions against expecting too much from the "new form of conditionality" proposed by Mosley, Hudson, and Verschoor (2004). It remains to be seen whether the ability to use aid for education as a means of strengthening the poverty orientation of government spending on education has improved with the advent of Poverty Reduction Strategy Papers.

Lacking conclusive evidence that conditionality works donors should be selective when determining the allocation of aid for education. The most obvious criterion for selectivity is the need for aid in education, as reflected in particularly low enrollment and completion rates. The targeting of aid for education according to need leaves much to be desired, as Thiele, Nunnenkamp, and Dreher (2007) show.

Another selectivity criterion stressed by many donors—the quality of governance in recipient countries—might be less important than widely believed. Investigation of this issue is left to future research.

Finally, it would be desirable to assess whether the effectiveness of aid for education could be enhanced by shifting donor resources within the sector toward basic education. Basic education accounted for about one-third of total aid for education by donors in 2002–04 (Thiele, Nunnenkamp, and Dreher 2007). This low level of spending is not only in conflict with Millennium Development Goal2, which would require greater concentration on basic education, but also with findings that social returns to primary education tend to be particularly high in low-income countries (World Bank 1995). Future research may be able to address this issue by disaggregating aid and education data more finely once longer time-series are available for a larger number of countries.

Appendix

Table A-1. Definitions and Sources

Variable	Description	Source
School enrollment	Ratio of number of children of official school age enrolled in school to number of children of official school age (net enrollment ratio)	World Bank (2005)

(Continued)

TABLE A-1. Continued

Variable	Description	Source
Primary completion rate	Number of students successfully completing the last year of (or graduating from) primary school in a given year divided by the number of children of official graduation age in the population	World Bank (2005)
Expenditure on education	Public spending on public education plus subsidies to private education at the primary, secondary, and tertiary levels. Variable measured both per capita and as percent of GDP	World Bank (2003, 2005)
Aid for education (commitments)	Aid commitments by all donors according to CRS Purpose Code 110. Includes aid for basic education, secondary education, postsecondary education, and unspecified levels of education. CRS guidelines require sector-specific program assistance and budget support in the form of sector-wide approaches to be subsumed under education when meant to benefit this sector. Variable measured both per capita, per child below age 15, and as percent of recipient country's GDP	OECD (2006)
Aid for education (disbursements)	Aid disbursements by all donors according to CRS Purpose Code 110, Form 2. Coverage is same as for commitments. Variable measured per capita	OECD (2006)
Total disbursements	Total aid (in all sectors) disbursed by all donors. Variable measured per capita	OECD (2006)
Literacy rate	Percentage of people 15 and older who can, with understanding, read and write a short simple statement on their everyday life	World Bank (2005)
GDP per capita	Per capita GDP in purchasing power parity terms (2000 international dollars)	World Bank (2005)
Population under 15	Percentage of total population under 15	World Bank (2005)
Urbanization	Share of total population living in areas defined as urban in each country	World Bank (2005)
Democracy index	[8 − (political rights index + civil liberties index)] /2	Freedom House (2004)
Pupil−teacher ratio	Number of pupils enrolled in primary school divided by number of primary school teachers (regardless of their teaching assignment)	World Bank (2005)

(Continued)

TABLE A-1. Continued

Variable	Description	Source
Unit costs of enrollment	Government expenditure on education divided by population under 15, as percent of per capita GDP	World Bank (2005)
International crisis	Dummy variable that takes value of 1 when a country is involved in an international crisis	Wilkenfeld and Brecher (2006)
Government stability	Assesses government's ability to carry out its declared programs and its ability to remain in office. Risk rating assigned is sum of three subcomponents (government unity, legislative strength, popular support), each with a maximum score of four points and a minimum score of 0 points. A score of 4 points indicates very low risk and a score of 0 points very high risk	ICRG (2005)
Under-five mortality	Probability that newborn baby will die before reaching age five if subject to current age-specific mortality rates. Variable expressed as rate per 1,000	World Bank (2005)
Economic Freedom	Composite 0–10 index of economic freedom; higher values reflect greater freedom	Gwartney and Lawson (2004)

Source: Authors' description based on cited data sources.

TABLE A-2. Summary Statistics (Estimation Sample, Table 1, Column 2)

Variable	Mean	Minimum	Maximum	Standard deviation
School enrollment (logarithm)	4.36	3.10	4.68	0.31
Primary completion rate (logarithm)	4.26	2.77	4.72	0.46
Expenditure on education (per capita)	95.78	3.56	697.50	111.94
Expenditure on education (percent of GDP)	4.28	0.83	11.58	1.88
Aid for education, commitments (per capita)	3.67	0.00	64.33	6.51
Aid for education, commitments (percent of GDP)	0.43	0.00	3.23	0.57
Aid for education, commitments per child	10.56	0.00	292.84	22.38
Aid for education, disbursements per capita	0.86	0.00	21.90	1.90
Total aid disbursements per capita	33.50	−2.34	304.07	37.50
Literacy rate	73.79	10.30	99.80	22.75
GDP per capita	4783	494	16050	3497
Population under 15	36.89	15.02	51.09	8.62
Urbanization	48.16	5.14	92.63	21.55
Democracy, index	4.04	1.00	7.00	1.76
Pupil-teacher ratio	31.53	10.28	77.03	13.20
International crisis	0.07	0.00	0.80	0.16
Unit costs of enrollment	13	1.76	38	7
Government stability	7.84	2.28	11.08	1.78

(Continued)

TABLE A-2. Continued

Variable	Mean	Minimum	Maximum	Standard deviation
Under-five mortality	75.14	7.50	320.00	65.48
Economic freedom	5.76	2.90	7.50	0.98

Source: Authors' analysis based on data described in table A-1.

REFERENCES

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests for Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58(2): 277–97.

Arellano, Manuel, and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-components Models." *Journal of Econometrics* 68(1):29–51.

Baldacci, Emanuele, Maria Teresa Guin-Siu, and Luiz R. de Mello. 2003. "More on the Effectiveness of Public Spending on Health Care and Education: A Covariance Structure Model." *Journal of International Development* 15(6):709–25.

Baldacci, Emanuele, Benedict J. Clements, Sanjeev Gupta, and Qiang Cui. 2004. *Aid, Social Spending, Human Capital, and Growth in Developing Countries: Implications for Achieving the MDGs*. IMF Working Paper WP/04/217. Washington, D.C: International Monetary Fund.

Behrman, Jere, Piyali Sengupta, and Petra Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico." *Economic Development and Cultural Change* 54(1):237–75.

Bennell, Paul. 2002. "Hitting the Target: Doubling Primary School Enrollments in Sub-Saharan Africa by 2015." *World Development* 30(7):1179–94.

Blundell, Richard, and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87(1):115–43.

Borenszstein, Eduardo, R., José. de Gregorio, and Jong-Wha Lee. 1998. "How Does Foreign Direct Investment Affect Economic Growth?" *Journal of International Economics* 45(1):115–35.

Bruns, Barbara, Alain Mingat, and Ramahatra Rakotomalala. 2003. *Achieving Universal Primary Education by 2015: A Chance for Every Child*. Policy Research Working Paper. World Bank, Washington, D.C.

Burnside, Craig, and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):847–68.

———. 2004. *Aid, Policies, and Growth: Revisiting the Evidence*. Policy Research Working Paper 3251. World Bank, Washington, D.C.

Clemens, Michael A. 2004. *The Long Walk to School: International Education Goals in Historical Perspective*. Center for Global Development Working Paper 37. Washington, D.C.

Clemens, Michael A., Steven Radelet, and Rikhil Bhavnani. 2004. *Counting Chickens When They Hatch: The Short-Term Effects of Aid on Growth*. Center for Global Development Working Paper 44, Washington, D.C.

Cruz, Luiz M., and Marcelo J. Moreira. 2005. "On the Validity of Econometric Techniques with Weak Instruments: Inference on Returns to Education Using Compulsory School Attendance Laws." *Journal of Human Resources* 40(2):393–410.

Dalgaard, Carl-Johan, Henrik Hansen, and Finn Tarp. 2004. "On the Empirics of Foreign Aid and Growth." *Economic Journal* 114(June):F191–216.

Doucouliagos, Hristos, and Martin Paldam. 2005. *The Aid Effectiveness Literature: The Sad Results of 40 Years of Research*. Department of Economics Working Paper 2005.05. Denmark: University of Aarhus.

Easterly, William. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York: Penguin Press.

Feldmann, Horst, "Government Size and Unemployment in Developing Countries." Forthcoming Applied Economics Letters.

Fielding, David, Mark McGillivray, and Sebastian Torres. 2005. "Synergies between Health, Wealth, Education, Fertility and Aid: Implications for Achieving the Millennium Development Goals." World Institute for Development Economics Research, Helsinki. (www.wider.unu.edu/research/2004-2005/2004-2005-1/mdg-papers/fielding-mcgillivray-torres-160805.pdf).

Filmer, Deon, and Lant Pritchett. 1999. "The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries." *Population and Development Review* 25(1):85–120.

Freedom House. 2004. *Annual Report*. (www.freedomhouse.org).

Gomanee, Karuna, Oliver Morrissey, Paul Mosley, and Arjan Verschoor. 2003. "Aid, Pro-Poor Government Expenditure and Welfare." CREDIT Research Paper 03/03. University of Nottingham, United Kingdom: Centre for Research in Economic Development and International Trade.

———. 2005. "Aid, Government Expenditure, and Aggregate Welfare." *World Development* 33(3):355–70.

Gupta, Sanjeev, Marijn Verhoeven, and Erwin Tiongson. 1999. *Does Higher Government Spending Buy Better Results in Education and Health Care?* IMF Working Paper WP/99/21. Washington, D.C.: International Monetary Fund.

Gwartney, James D., and Robert A. Lawson. 2004. *Economic Freedom of the World: 2004 Annual Report*. Vancouver: Fraser Institute. (www.freetheworld.com).

Hanushek, Erik A. 1995. "Interpreting Recent Research on Schooling in Developing Countries." *World Bank Research Observer* 10(2):227–46.

Hout, Wil. 2004. "Political Regimes and Development Assistance: The Political Economy of Aid Selectivity." *Critical Asian Studies* 36(4):591–613.

ICRG (International Country Risk Guide). 2005. International Country Risk Guide Political Risk Indicator. (www.countryrisk.com/guide/archives/cat_country_risk_ratings.html).

Isenman, Paul, and Dag Ehrenpreis. 2003. "Results of the OECD DAC/Development Centre Experts' Seminar on 'Aid Effectiveness and Selectivity: Integrating Multiple Objectives into Aid Allocations.'" *DAC Journal* 4(3):7–25.

McGillivray, Mark, Simon Feeny, Niels Hermes, and Robert Lensink. 2005. "It Works; It Doesn't; It Can, But That Depends. 50 Years of Controversy over the Macroeconomic Impact of Development Aid." Research Paper 2005/54. Helsinki: United Nations University, World Institute for Development Economics Research.

McKinlay, Robert D., and Richard Little. 1977. "A Foreign Policy Model of U.S. Bilateral Aid Allocation." *World Politics* 30(1):58–86.

Michaelowa, Katharina, and Anke Weber. 2006. "Aid Effectiveness Reconsidered: Panel Data Evidence for the Education Sector." HWWA Discussion Paper 264. Germany: Hamburg Institute of International Economics.

Mingat, Alain, and Jee-peng Tan. 1998. *The Mechanics of Progress in Education: Evidence from Cross-Country Data*. Policy Research Working Paper 2015, World Bank, Washington, D.C.

———. 2003. "On the Mechanics of Progress in Primary Education." *Economics of Education Review* 22(5):455–67.

Mosley, Paul, John Hudson, and Arjan Verschoor. 2004. "Aid, Poverty Reduction, and the 'New Conditionality.'" *Economic Journal* 114(496):F217–43.

Nelson, Charles R., and Richard Startz. 1990. "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator." *Econometrica* 58(4):967–76.

OECD (Organisation for Economic Co-operation and Development). 2006. International Development Statistics (IDS) Online: Databases on Aid and Other Resource Flows. (www.oecd.org/dataoecd/50/17/5037721.htm).

Pettersson, Jan. 2006. "Foreign Sectoral Aid Fungibility, Growth and Poverty Reduction." *Paper presented at the World Institute for Development Economics Research Conference "Aid: Principles, Policies and Performance,"* June 16–17, Helsinki.

Pritchett, Lant. 2001. "Where Has All the Education Gone?" *World Bank Economic Review* 15(3):367–91.

Rajan, Raghuram G., and Arvind Subramanian. 2005. *Aid and Growth: What Does the Evidence Really Show?* IMF Working Paper WP/05/127. Washington, D.C.: International Monetary Fund.

Reinikka, Ritva, and Jacob Svensson. 2001. "Explaining Leakage of Public Funds." WIDER Discussion Paper 2001/147, Helsinki: United Nations University, World Institute for Development Economics Research.

Roberts, John. 2003. *Poverty Reduction Outcomes in Education and Health: Public Expenditure and Aid*. ODI Working Paper 210. London: Overseas Development Institute.

Roodman, David. 2005. "Xtabond2: Stata Module to Extend Xtabond Dynamic Panel Data Estimator." Washington, D.C.: Center for Global Development.(http://econpapers.repec.org/software/ bocbocode/s435901.htm).

Schultz, T. Paul. 1988. "Expansion of Public School Expenditures and Enrollments: Intercountry Evidence on the Effects of Income, Prices, and Population Growth." *Economics of Education Review* 7(2):167–83.

Sen, Amartya. 2006. "The Man without a Plan." *Foreign Affairs* (March/April). (www.nyu.edu/fas/institute/dri/Easterly/File/Foreign%20Affairs%20-%20The%20Man%20Without%20a%20Plan%20-%20Amartya%20Sen.pdf).

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3):557–86.

Svensson, Jacob. 1999. "Aid, Growth and Democracy." *Economics & Politics* 11(3):275–97.

Tarp, Finn. 2006. "Scaling Up Aid: Challenges and Opportunities: A Few Comments." Paper presented at the World Institute for Development Economics Research Conference "Aid: Principles, Policies and Performance," June 16–17, Helsinki.

Thiele, Rainer, Peter Nunnenkamp, and Axel Dreher. 2007. "Do Donors Target Aid in Line with the Millennium Development Goals? A Sector Perspective of Aid Allocation." *Review of World Economics* 143(4):596–630.

Von Hagen, Juergen. 2005. "Political Economy of Fiscal Institutions." GESY Discussion Paper 149. Germany: Mannheim University.

Wilkenfeld, Jonathan, and Michael Brecher. 2006. "International Crisis Behavior Project." Center for International Development and Crisis Management, College Park, MD, USA. (www.cidcm.umd.edu/icb).

Windmeijer, Frank. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators." *Journal of Econometrics* 126(1):25–51.

Wolf, Susanna. 2007. "Does Aid Improve Public Service Delivery?" *Review of World Economics* 143(4):650–672.

World Bank. 1995. *Development in Practice: Priorities and Strategies for Education*. Washington D.C.

———. 2003. *World Development Indicators* CD-ROM. 2005, and 2007 Washington, D.C.

World Bank Lending and Financial Sector Development

Robert Cull and Laurie Effron

A new database of World Bank loans to support financial sector development is used to investigate whether countries that received such loans experienced more rapid growth on standard indicators of financial development than countries that did not. Self-selection is accounted for with treatment-effects regressions. The results indicate that borrowing countries had significantly more rapid growth in M2/GDP than non-borrowers and swifter reductions in interest rate spreads and cash holdings (as a share of M2). Borrowers also had higher private credit growth rates than nonborrowers in some treatment-effects regressions but not in standard panel regressions with fixed country effects. On the whole, the results indicate some significant advantages in financial development for borrowers over nonborrowers. JEL codes: F33, G21, O16.

The World Bank has been making loans to governments of member countries since 1946. Over time, World Bank lending shifted from supporting post-World War II reconstruction to supporting economic growth and poverty alleviation. During its first four decades, it concentrated on financing investments in infrastructure and directly productive activities in agriculture and industry. This approach was driven by the assumption that the scarcity of foreign exchange for capital investments was the main constraint hindering economic growth in developing countries.

With the shift in analytic focus on constraints to growth in the 1980s (initially to the economic policies of developing countries and later to their institutional capacities), the World Bank introduced fast-disbursing loans for balance of payments support conditioned on changes in policies and institutions.

Consistent with its sharper focus on the importance of good macroeconomic policies and institutions for economic growth, the World Bank became an early proponent of supporting appropriate policies and capable institutions in the financial sector, which, it asserted, could also contribute to economic growth (World Bank 1989). These positions were later supported by extensive research establishing a causal link between financial development and economic growth (Levine and Zervos 1998; Rajan and Zingales 1998; Beck, Levine, and Loayza 2000; Levine, Loayza, and Beck 2000; Levine 2005) and by research showing that less government control over financial systems and institutions leads, over time, to deeper, stabler, and more efficient systems (Caprio, Honohan, and Stiglitz 2001; Barth, Caprio, and Levine 2001a,b; La Porta, Lopez-de-Silanes, and Shliefer 2002; Beck, Demirguc-Kunt, and Levine 2003).

In the early 1990s, most developing countries maintained policies and institutions that were considered detrimental to financial sector development— namely, government-administered financial systems, fixed interest rates, directed credit, and market dominance of publicly owned financial institutions. The World Bank began to target lending support to financial sector reforms addressing these constraints. Between 1992 and 2003, about one quarter of World Bank lending—some $56 billion—included support for financial sector reforms aimed at reducing direct government control over credit allocation, interest rates, and financial institutions and increasing government oversight of domestic financial markets and institutions by strengthening banking supervision and prudential regulations. The basic objective of such support was to establish a strong enabling environment in which well-governed financial institutions would mobilize resources, allocate credit, and manage risks efficiently.

Most reforms focused on the banking sector and within the banking sector on the restructuring or privatization of state-owned banks. Reforms also sought to strengthen banking legislation, regulation, and supervision. An independent evaluation found that government ownership of banks in countries with Bank loans that included conditionality on bank privatization decreased substantially (with the percentage of banking assets owned by governments dropping 60 percentage points within a decade) and by more than that in countries that did not borrow for bank privatization (where the decrease was 35 percentage points) (IEG 2006). The results were more ambiguous for differences in banking legislation and regulation; information on changes in banking supervision was not available for many countries.

Did these changes increase the mobilization of resources, allocate more credit, and make the financial sector more efficient? This article addresses these questions. Using quantitative indicators to measure changes in depth, efficiency, and credit to the private sector, it examines whether Bank assistance between 1992 and 2003 helped develop financial sectors in client countries and tests whether progress was greater in countries receiving loans for this purpose than in client countries that did not receive such loans.

The article contributes to a broader literature on the effects of economic reform programs in developing countries. Much of that literature focuses on the impact

of adjustment lending by the World Bank and the International Monetary Fund (IMF) on broad macroeconomic aggregates, notably real per capita growth.[1] This article examines the effects of World Bank lending on measures of financial sector development, which has been linked to economic growth.

Improvements in indicators of financial development were generally significant for borrowers—and more pronounced for borrowers than nonborrowers. Treatment-effects regressions are used that explicitly account for nonrandom selection (the possibility that borrowers tended to be countries that were likely to have improved their financial sectors without the loans). Additional robustness checks test whether the findings are specific to particular regions and whether improvements in financial indicators preceded or followed World Bank loans. (If improvement preceded the loans, it would seem unlikely that the loans had a large causal impact in borrowing countries.) A final set of checks incorporates additional controls for countries' readiness for and experience with financial reform. The results of these tests reinforce the main findings: borrowing countries tended to experience substantial improvement in their financial indicators, significantly more than the typical improvement in nonborrowing countries, even after accounting for selection.

The remainder of the article is organized as follows. Section I describes the data, including the indicators used to assess outcomes and the variables that summarize World Bank lending in support of financial sector reform from 1992 to 2003. Sections II and III describe the basic regression models and summarize the base results, and section IV presents results using estimation techniques that address selection problems. Section V runs additional robustness checks, including regional regressions and models that attempt to control for countries' history of and readiness for reform. The last section briefly summarizes the results of the various methods.

I. Data

The analysis relies on standard indicators, such as M2/GDP and private credit/GDP, that have been shown to be robustly associated with long-run economic growth (Beck, Levine, and Loayza 2000; Levine, Loayza, and Beck 2000; Levine 2005).[2] This analysis is restricted to banking indicators, because banks hold the vast majority of financial sector assets in developing countries.[3]

---

1. Easterly (2005) describes this literature as including Barro and Lee (2002), Conway (1994), Corbo and Goldstein (1987), Corbo and Fischer (1995), Devarajan, Dollar, and Holmgren (2001), Dicks-Mireaux, Mecagni, and Schadler (2000), Goldstein and Montiel (1986), Haque and Khan (1998), Hutchinson (2001), Kapur, Lewis, and Webb (1997), Khan (1990), Killick, Gunatilaka, and Marr (1998), Knight and Santaella (1997), Summers and Pritchett (1993), Przeworski and Vreeland (2000), Svensson (2003), and Van de Walle (2001).

2. For descriptions of standard indicators of financial development and their use, see Beck, Demirgüç-Kunt, and Levine (2000).

3. The ratio of private credit to GDP can, however, include lending by nonbank financial institutions.

An advantage of these indicators is that they are available for many countries throughout the decade of this analysis. They do create some problems, however. For example, as part of the restructuring or privatization of problem banks, the value of nonperforming assets may be reduced or the loans eliminated from bank balance sheets, thereby reducing the private credit ratio. Successful restructuring efforts contribute to a healthier banking sector. Because such efforts reduce private credit, however, they would be viewed as detrimental to financial development in the models presented here.

Moreover, the private credit ratio does not provide information about which segments of society receive credit or about the quality of the loans made, because data on nonperforming loans are not available on a standardized basis across countries. The M2/GDP ratio provides information on deposit levels, but that information is not broken down by the income level of the depositors. Increases in the ratio may not mean that all segments of society are availing themselves of formal banking services.

Additional indicators—namely, the spread between the lending and deposit interest rates and the ratio of cash held outside of banks to M2 (a measure of the lack of confidence in the formal banking sector)—are incorporated to round out the assessment of financial development. These indicators were chosen largely because of data availability. They, too, have limitations. For example, interest rates were controlled in a number of developing countries at some point during the sample period; spreads are unlikely to be an accurate measure of efficiency in these instances. Measures of capital adequacy, portfolio quality, and profitability are not available in a standard format across countries.[4]

These caveats notwithstanding, it seems likely that taken together, M2/GDP, private credit/GDP, cash/M2, and interest rate spreads provide a reasonably complete picture of both short- and long-term banking development between 1992 and 2003.[5] In the short term, movements in the ratio of cash/M2 can show depositors' reactions to policy changes. The private credit ratio, while subject to short-term perturbations, tends to capture long-term financial development. M2/GDP and interest rate spreads are arguably somewhere between the two extremes.

4. In the robustness checks presented in the supplemental appendix, an index of financial sector efficiency and freedom developed by the Heritage Foundation is used as the dependent variable in the regressions. The results are qualitatively similar to those for the four quantitative indicators used here.

5. These measures could render an imprecise picture, because the opaqueness of the financial sector makes the effects of reform difficult to observe. As robustness checks, the base models were also rerun using credit ratings—which presumably offer a more nuanced assessment of the success of reforms— from three agencies (Moody's, Standard and Poor's, Fitch IBCA) as dependent variables. The unit of observation is the average rating of all sovereign debt issues for each country in each year. Though the coefficient on that variable was positive for borrowers and negative for nonborrowers, neither was significant. This is likely because there are so few observations to work with since only a subset of developing countries issued enough sovereign debt on a regular basis. This illustrates the key advantage of the quantitative indicators used in this paper, which is their yearly availability for a wide set of developing countries.

Standard financial outcome variables are combined with data on World Bank loans from a review by the World Bank's Independent Evaluation Group (IEG 2006) (formerly the Operations Evaluation Department). IEG examined 556 lending operations between 1992 and 2003 that involved support for reform of the financial sector,[6] including adjustment loans, technical assistance, and lines of credit.[7] The analysis focuses on adjustment loans because they were larger (in dollar terms) and included conditions aimed at spurring broad financial sector development. In contrast, lines of credit often focused on channeling funds through specific financial institutions; technical assistance loans tended to be much smaller than adjustment loans and focused on narrower issues.[8]

Sixty-eight countries received at least one adjustment loan with conditions tied to financial reform (table A-1). A control group of 38 countries received no such loans during the period under review. All developing countries for which reasonably complete financial indicator data were available were included in the analysis. Comparisons between the two groups form the basis of most of the statistical analysis that follows.[9]

II. Methodology

The base results rely on simple fixed-effects panel regressions of the form

$$(1) \quad Y_{it} = \alpha_i + \beta_{wb}t + \beta_{no-wb}t + \beta_1 \text{ first loan}_{it} + \beta_2 adj_{it} + \beta_3 \mathbf{ref}_{it} + \beta_4 \mathbf{X}_{it-1} + \varepsilon_{it}$$

where i represents the country and t represents time, measured in years since 1991. The time variable takes on values of 1 to 12. The estimated coefficient β_{wb} thus represents the average growth rate of the indicator of financial development (Y) for countries that received adjustment loans to support financial reform (table 1 provides descriptions and summary statistics for all the

6. The database of World Bank loans starts in 1992; country and financial sector data start in 1991. The data set was gathered as part of an IEG evaluation of Bank lending to the financial sector. IEG began its analysis in 1992 in an effort to avoid duplication of a previous study (which covered 1984–98) while still covering a sufficiently long period. This type of censoring could conceivably affect the results.

7. Lines of credit are funds passed through an intermediary for demand-driven purposes. The end-user has to repay the loan, usually with interest.

8. Regressions were run with various samples of loans. For models based solely on investment loans, technical assistance loans, or lines of credit, there were no robust significant differences between borrowers and nonborrowers. Including smaller subcategories, such as loans for technical assistance, along with the adjustment loans, did not change the qualitative differences between borrowers and nonborrowers for the adjustment loan-only sample. Variables based on small subcategories of financial sector loans (such as those devoted to pensions) did not produce stable significant differences between borrowers and nonborrowers.

9. The World Bank maintained a policy dialogue throughout the period with some countries in the control group. Nine of the 38 countries in the control group borrowed for financial sector reform before the period of study, although in almost all cases the borrowing consisted of a single loan, often granted as part of a multisector operation in which financial reforms were not central. These factors could make it harder to find statistically significant differences between the two groups in the regressions that follow.

TABLE 1. Variable Descriptions and Summary Statistics

Variable	Description	Mean	Median	High	Low
Growth rate	Equal to the year minus 1991. Estimated coefficient measures annual growth rate for dependent variable in question	6.68	7	12	1
Policy variables					
Adjustment loans	Cumulative number of adjustment loans by country in question at time *t*. In some models, total includes loans for technical assistance	0.70	0	6	0
Bank privatization	Number of adjustment loans with emphasis on bank privatization	0.24	0	4	0
Regulation and supervision	Number of adjustment loans with emphasis on bank regulation and supervision	0.39	0	4	0
Other banking	Number of adjustment loans with emphasis on banking reform other than privatization, regulation, or supervision	0.48	0	4	0
Auditing and accounting reform	Number of adjustment loans with emphasis on accounting and auditing reform	0.06	0	2	0
Capital market development	Number of adjustment loans with emphasis on capital and securities market development	0.19	0	3	0
General financial sector reform	Number of adjustment loans with emphasis on general financial sector development not covered under other variables	0.19	0	3	0
Rural finance	Number of adjustment loans with emphasis on rural financial sector development	0.03	0	2	0
Microfinance	Number of adjustment loans with emphasis on development of microfinance	0.01	0	1	0
Nonbank financial sector institutions	Number of adjustment loans with emphasis on development of nonbank financial institutions	0.04	0	2	0
Dependent variables					
Private credit/GDP	Claims on private sector (International Financial Statistics [IFS] line 22d) divided by GDP (IFS line 99b) multiplied by 100	25.1	17.7	158.5	0
M2/GDP	Money (IFS line 34) plus quasi-money (IFS line 35) divided by GDP (IFS line 99b) multiplied by 100	33.2	26.9	148.2	0.002
Cash/M2	Currency outside deposit money banks (IFS line 14a) divided by M2 (IFS line 34 + line 35) multiplied by 100	23.1	19.4	82.5	0
Interest rate spread	Lending rate (IFS line 60l) minus deposit rate (IFS line 60p), multiplied by 100	11.4	8.6	163.5	−6.9

(Continued)

TABLE 1. Continued

Variable	Description	Mean	Median	High	Low
Concentration	Percentage share of total banking sector assets held by three largest banks (based on asset information in Bankscope)	62.0	59.5	100.0	14.9
Macro/institutional controls					
CPIA score	Proxy for institutional development	3.23	3.28	5.35	1.0
Surplus (deficit)/ GDP	Overall budget balance, including grants, multiplied by 100 (World Bank)	−2.72	−2.11	10.26	−31.63
Annual GDP growth (percent)	Annual GDP growth (World Bank, various years)	3.00	3.94	106.3	−50.2
Inflation (percent)	GDP deflator (World Bank, various years)	78.1	9.41	6041.6	−25.2
Selection equation variables					
Government checks	Variable equals one if there is no chief executive. It rises by one under each of the following circumstances: there is a chief executive, the chief executive is competitively elected, and the opposition controls the legislature. In presidential systems, it rises by one for each chamber of the legislature, unless the president's party has a majority in the lower house and a closed-list system is in effect (implying stronger presidential control of the party and therefore of the legislature). It also rises by one for each party coded as allied with the president's party that has an ideological (left-right-center) orientation closer to that of the main opposition party than to that of the president's party. In parliamentary systems this variable rises by one for every party in the government coalition as long as the parties are needed to maintain a majority and for every party in the government coalition that has a position on economic issues (right-left-center) that is closer to that of the largest opposition party than to that of the party of the executive. In parliamentary systems, the prime minister's party is not counted if there is a closed rule in place (in this case the prime minister is presumed to fully control the party). The highest possible score is 18. Average checks 1991–2000 are calculated for each country. (World Bank Database on Political Institutions; see Beck and others 2001.)	2.65	2.50	10.1	1.0
Debt (as percent of GNI)	Average external debt 1970–89 (World Bank, various years)	56.2	48.1	222.2	4.0

(Continued)

TABLE 1. Continued

Variable	Description	Mean	Median	High	Low
IMF credit (millions of constant dollars)	Average IMF credit 1970–89 (World Bank, various years)	462.0	93.5	9,370.0	0
Total debt service (percent of GNI)	Average debt service 1990–99 (World Bank, various years)	5.3	4.2	0.3	20.3
Capital formation (percent of GDP)	Gross fixed capital formation 1990–99 (World Bank, various years)	22.3	21.0	6.9	64.8

Note: Figures are calculated over all observations for which at least one dependent variable was available.

variables). A test of whether $\beta_{wb} = \beta_{nonwb}$ indicates whether adjustment loans had a beneficial impact on financial sector development. To the extent that the growth rates for the control countries were the same as (or greater than) those of countries that received World Bank assistance, the value of that assistance could be questioned. All regressions also include α_i, a country-specific fixed effect. Results should be interpreted as changes relative to the country-specific mean for the indicator in question. (More direct methods for addressing potential selection problems are presented later.)

First loan measures the number of years since a country received its first loan with financial sector conditions. It is included because improvements in financial indicators were more likely to have materialized in countries that received loans early in the period.[10] Including the *first financial sector loan* variable offers a more precise test of whether improvements in financial indicators occurred after the receipt of loans. For example, if β_{wb} is positive and significant but the coefficient for the time since first loan variable is insignificant, it would suggest that as a group borrowing countries were more likely to improve their financial indicators regardless of when they received loans from the World Bank. By contrast, if the *first loan* variable is significant while the simple borrower dummy variable is not, it would suggest that improvement in indicators occurred after the receipt of World Bank loans.

---

10. For the base regressions, the fairest tests of whether World Bank lending contributed to financial development should include country-specific fixed effects; specification tests confirm that they should be included. Therefore, the dummy variable *received Bank adjustment loans* is set equal to one throughout the period, regardless of whether the country received its first loan in the first year or the twelfth year. The dummy for *no Bank adjustment loans* is set equal to one throughout the period for nonborrowers. Had the variables not been coded in this way, all countries that received no loans would have been lost from the observation set because of the country fixed effects; only countries whose borrowing status changed during the period would be used to examine the effects of Bank lending. Such models would have offered comparisons for borrowing countries before and after receiving a loan, but they would not have facilitated comparisons between borrowers and nonborrowers, the focus of this article.

Some specifications include *adj*, the cumulative number of adjustment loans. Some countries received as many as six adjustment loans with financial sector reform components between 1992 and 2003. Repeated structural adjustment lending from the World Bank or the IMF failed to produce improvement on multiple macroeconomic outcomes, including growth (Easterly 2005).[11] Models with *adj* therefore test whether similar results hold for the financial sector. As with the macroeconomic and institutional variables, all policy reform variables are lagged one year in the regressions.

Ref is a vector of variables summarizing reform areas covered under adjustment loans (bank privatization; bank regulation and supervision; banking reform not focused on privatization, regulation, or supervision; auditing and accounting reform; capital market development; reform of nonbank financial institutions; general financial sector reform; rural finance; and microfinance).[12] Because the data set is a country-level panel of financial sector outcomes, the project-level data must be aggregated into country-year reform packages. The cumulative number of loans that had conditions in the policy areas in question are explanatory variables in the regressions that follow.[13]

X is a vector of macroeconomic and institutional controls, including inflation, real growth, and M2/GDP.[14] All of the macroeconomic and institutional controls are lagged one year in the panel regressions that follow to mitigate problems arising from the simultaneous determination of the controls and the dependent variables. Inflation should slow financial development if it makes loan contracting over extended periods more difficult. Real growth will accelerate financial development, because it is likely to stimulate demand for financial services. Because World Bank lending to all sectors could spur growth and growth could spur financial development, this is an important control for isolating the effect of financial sector loans on indicators of financial development.

11. Easterly notes that the repeated extension of loans to a country is itself a sign that lending was not effective, "One might expect that it would take more than one loan to accomplish 'adjustment,' but it is hard to see why it would take such a large number" (2005, p. 6).

12. The intention was to specify the policy areas that had the greatest chance of improving financial indicators, which are largely bank based. Adjustment loans devoted solely to small- and medium-size enterprise finance or pensions were therefore excluded from the database (very few loans focused only on these areas).

13. Similar qualitative results hold when the number of loans in a given year covering that policy area, simple dummies indicating that a policy area was covered, or dummies indicating that the policy area was covered at some point during the sample period are used.

14. Government budget deficits were included in initial specifications, but they tended to be insignificant. Since inclusion of that variable reduced the sample size by almost half, it was eliminated from the final specifications. Its inclusion does not greatly alter the comparison between borrowers and nonborrowers.

The base model includes M2/GDP as a general measure of the level of financial development. It is not clear a priori whether the level of financial development should have a positive or negative effect on subsequent financial development. On the one hand, a high level of M2/GDP could signal a high level of future development. In that case, lagged M2/GDP can be viewed as a proxy for a country's willingness and ability to pursue financial sector reform. On the other hand, a low level of M2/GDP could signal greater potential for improvement.

The Country Policy and Institutional Assessment (CPIA) index is included as a broad measure of institutional development. The World Bank conducts this assessment annually to assess the quality of a country's policy and institutional framework. The index is based on 20 criteria, grouped into four clusters: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. The CPIA is available for a large sample of countries for the whole period. It incorporates information about the conduciveness of a country's policy framework for reform. Including it reduces concerns that the borrower variable might be picking up a country's general ability to achieve reform. Until 2005, this variable was not available outside the World Bank, and details of its construction were not well known. In the robustness checks, therefore, CPIA is replaced with proxies for institutional development that are more readily available and (arguably) less endogenous.

One could view the basic model as a one-lag vector autoregression in four variables (CPIA score, inflation, real growth, and M2/GDP). There is no guarantee, however, that this is the correct reduced-form model. A series of models tested for the appropriate included variables and lag lengths by adding lags for each of the explanatory variables until the last lag added was not significant. This measure was taken in order to ensure that the underlying model of the indicators of financial development is as complete as possible before adding the treatment variables. Including the additional lags reinforces the conclusions about the relative performance of borrowers and nonborrowers. Indeed, differences between borrowers and nonborrowers are larger in the specifications with multiple lags. To reduce clutter in the specifications and for ease of exposition, the analysis uses the one-lag models as the base specifications (results from the multiple-lag specifications are also discussed below).

III. RESULTS

In the base specifications, percentage changes in the indicators of financial development are measured by taking their logs (table 2).[15] Two of the simplest

---

15. Taking logs also helps reduce the influence of outliers in the estimated coefficients.

TABLE 2. Base Results: Fixed-Effects Panel Regression

	Log(M2/GDP)					Log(private credit/GDP)				
	2.1 (2.1)	2.2 (2.2)	2.3	2.4	2.5 (2.3)	2.6 (9.4)	2.7 (9.5)	2.8	2.9	2.10 (9.6)
Received World Bank adjustment loan	0.036***	0.017***	−0.009	−0.007	−0.008	0.029***	0.011*	0.006	0.009	0.022
	(0.004)	(0.003)	(0.007)	(0.007)	(0.007)	(0.005)	(0.006)	(0.015)	(0.015)	(0.015)
Did not receive World Bank adjustment loan	0.014***	0.002	0.003	0.002	0.003	0.037***	0.022***	0.022***	0.022***	0.022***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
H_0: $\beta_{wb} = \beta_{nonwb}$ rejected? (P =0.05)	Yes	Yes	No	No	No	No	No	No	No	No
Years since first adjustment loan			0.034***	0.037***	0.047***			0.006	0.011	0.018
			(0.009)	(0.009)	(0.009)			(0.018)	(0.018)	(0.018)
H_0: $\beta_{wb} = \beta_{yrs\ since\ 1st}$ rejected?			Yes	Yes	Yes			No	No	No
H_0: $\beta_{yrs\ since\ 1st} = \beta_{nonwb}$ rejected?			Yes	Yes	Yes			No	No	No
Macroeconomic and institutional controls										
CPIA score$_{t−1}$		0.004	−0.007	−0.004	0.001		0.016	0.014	0.019	0.033
		(0.017)	(0.017)	(0.017)	(0.016)		(0.033)	(0.034)	(0.034)	(0.034)
GDP growth$_{t−1}$		0.003**	0.003**	0.003**	0.003***		0.005**	0.005**	0.005**	0.004*
		(0.001)	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.002)	(0.002)
Inflation$_{t−1}$		−0.000001	−0.00003	−0.0001	0.0002		−0.00016	−0.00017	−0.0002	0.0002
		(0.0001)	(0.0001)	(0.0001)	(0.0001)		(0.0003)	(0.0003)	(0.0003)	(0.0003)
M2/GDP$_{t−1}$		0.015***	0.015***	0.015***	0.014***		0.009***	0.009***	0.009***	0.008***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.002)	(0.002)
Policy variables										
Cumulative adjustment loans				−0.019	−0.047***				−0.031	−0.042
				(0.013)	(0.016)				(0.028)	(0.032)
Number of loans for bank privatization					0.041*					0.018
					(0.022)					(0.046)
Number of loans for regulation/ supervision					0.042					−0.014
					(0.030)					(0.061)
Number of loans for other reforms					−0.047*					−0.156
					(0.025)					(0.053)
Number of loans for auditing/ accounting					0.040					0.319***
					(0.041)					(0.086)
Number of loans for capital market					0.083***					0.215***
					(0.029)					(0.061)
Number of loans for general financial					−0.015**					−0.019
					(0.025)					(0.052)
Rural finance					0.199**					0.233**
					(0.053)					(0.110)
Microfinance					−0.059					−0.247*
					(0.069)					(0.142)
Nonbank financial sector institutions					−0.199***					−0.143
					(0.044)					(0.092)
Constant	3.13***	2.71***	2.81***	2.80***	2.80***	2.63***	2.37***	2.39***	2.37***	2.30***
	(0.021)	(0.062)	(0.066)	(0.067)	(0.066)	(0.032)	(0.125)	(0.136)	(0.137)	(0.137)
Number of observations	866	611	611	611	611	899	610	610	610	610
Number of countries	90	89	89	89	89	94	89	89	89	89
R-squared (within)	0.13	0.42	0.43	0.44	0.49	0.06	0.09	0.09	0.10	0.17

	Log(cash/M2)					Log(interest spread)				
	2.11 (9.7)	2.12 (9.8)	2.13	2.14	2.15 (9.9)	2.16 (9.10)	2.17 (9.11)	2.18	2.19	2.20 (9.12)
Received World Bank adjustment loan	−0.028***	−0.034***	−0.007	−0.017	−0.024*	−0.005	−0.011	0.073***	0.076***	0.072***
	(0.003)	(0.005)	(0.013)	(0.013)	(0.014)	(0.007)	(0.009)	(0.021)	(0.022)	(0.022)

(Continued)

TABLE 2. Continued

	Log(cash/M2)					Log(interest spread)				
	2.11 (9.7)	2.12 (9.8)	2.13	2.14	2.15 (9.9)	2.16 (9.10)	2.17 (9.11)	2.18	2.19	2.20 (9.12)
Did not receive World Bank adjustment loan	−0.017***	0.001	0.0001	0.001	0.0005	0.015	0.027**	0.024**	0.024**	0.024**
	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)	(0.012)	(0.011)	(0.011)	(0.011)
$H_0: \beta_{wb} = \beta_{nonwb}$ rejected? (P =0.05)	Yes	Yes	No	No	No	No	Yes	Yes	Yes	No
Years since first adjustment loan			−0.034**	−0.042***	−0.045***			−0.116***	−0.113***	−0.111***
			(0.016)	(0.016)	(0.016)			(0.026)	(0.027)	(0.027)
$H_0: \beta_{wb} = \beta_{yrs\ since1st}$ rejected?			No	No	No			Yes	Yes	Yes
$H_0: \beta_{yrs\ since1st} = \beta_{nonwb}$ rejected?			Yes	Yes	Yes			Yes	Yes	Yes
Macroeconomic and institutional controls										
CPIA score$_{t-1}$		−0.025	−0.014	−0.030	−0.030		0.014	0.048	0.054	0.046
		(0.030)	(0.030)	(0.030)	(0.030)		(0.050)	(0.049)	(0.050)	(0.050)
GDP growth$_{t-1}$		0.001	0.0003	0.0001	0.001		−0.005	−0.006	−0.006	−0.007*
		(0.002)	(0.002)	(0.002)	(0.002)		(0.004)	(0.004)	(0.004)	(0.004)
Inflation $_{t-1}$		−0.0004*	−0.0004	−0.0003	−0.0006		0.001*	0.001**	0.001**	0.0005
		(0.0002)	(0.00024)	(0.0002)	(0.0003)		(0.0003)	(0.0003)	(0.0003)	(0.0004)
M2/GDP$_{t-1}$		−0.008***	−0.008***	−0.008***	−0.007***		0.008***	0.010***	0.010**	0.009***
		(0.002)	(0.002)	(0.002)	(0.002)		(0.003)	(0.003)	(0.003)	(0.003)
Policy variables										
Cumulative adjustment loans				0.086***	0.105***				−0.025	0.022
				(0.023)	(0.027)				(0.038)	(0.050)
Number of loans for bank privatization					0.037					0.069
					(0.038)					(0.073)
Number of loans for regulation/ supervision					0.041					0.099
					(0.052)					(0.111)
Number of loans for other reforms					0.030					−0.130
					(0.047)					(0.087)
Number of loans for auditing/ accounting					−0.206***					0.170
					(0.074)					(0.133)
Number of loans for capital market					−0.163***					−0.349***
					(0.052)					(0.114)
Number of loans for general financial					0.010					0.053
					(0.045)					(0.092)
Rural finance					−0.235**					−0.012
					(0.090)					(0.134)
Microfinance					0.121					−0.107
					(0.101)					(0.228)
Nonbank financial sector institutions					0.025					0.112
					(0.085)					(0.161)
Constant	3.03***	3.39***	3.29***	3.35***	3.38***	2.10***	1.77***	1.42***	1.40***	1.50***
	(0.017)	(0.109)	(0.118)	(0.117)	(0.118)	(0.044)	(0.185)	(0.196)	(0.199)	(0.205)
Number of observations	1119	671	671	671	671	532	366	366	366	366
Number of countries	98	87	87	87	87	60	57	57	57	57
R-squared (within)	0.09	0.15	0.16	0.18	0.22	0.01	0.07	0.12	0.13	0.16

*Significant at the 10 percent level; **significant at the 5 percent level; ***significant at the 1 percent level.

Note: All models include country fixed effects. All macroeconomic and institutional control variables are lagged one year. Standard errors appear in parentheses.

Source: Authors' analysis based on data sources described in the text and in table 1.

regressions (2.1 and 2.11) indicate that borrowers had significantly more rapid growth in M2/GDP and more rapid declines in cash/M2, both signs of better financial development. This pattern holds when macro/institutional controls

are included in the regressions (models 2.2 and 2.12).[16] Interest spreads tended to widen for nonborrowers and decline for borrowers, though neither coefficient is significant in model 2.16. When the macroeconomic and institutional controls are included in model 2.17, the nonborrower coefficient is positive and significant, and the borrower coefficient remains negative and insignificant. Borrowers thus tended to outperform nonborrowers in terms of spreads.

The pattern of results is different for private credit, which grew faster in nonborrowing than in borrowing countries, though the difference is not significant in either the simplest regression (2.6) or the regression that includes controls (2.7). In addition, private credit growth was positive and significant for borrowers in both specifications, which might come as a surprise given the number of borrowers that suffered financial crises during the period (see table A-1 for a list of borrowers).

If World Bank loans are designed to spur financial development, borrowers should enjoy significantly faster credit growth than nonborrowers. One possible explanation for the fact that they did not is that the nonborrower growth rates are "too high." Private credit growth rates are more than twice as large as M2/GDP growth rates for nonborrowers in the simplest specification, and the nonborrower coefficient does not achieve significance in the M2/GDP specification when the controls are introduced. Such a pattern may be possible over a short period; over longer periods, it is likely to be destabilizing and unsustainable. Indeed, several articles indicate that rapid growth in indicators of financial depth, particularly those related to credit, can be so destabilizing that they lead to crisis (Demirgüç-Kunt and Detragiache 1998; Kaminsky and Reinhart 1999; Honohan 2004; Loayza and Ranciere 2004).[17] By contrast, the growth rates for borrowers seem more reasonable (1–3 percent for private credit/GDP and 2–4 percent for M2/GDP).

Improvements were largest for indicators that are (arguably) better suited to capturing short- and medium-term financial sector development (cash/M2, spreads, and M2/GDP). Therefore, it could be argued that sufficient time had not elapsed to see the full effects of reform on private credit for borrowers. This does not explain why borrowers would perform worse than nonborrowers over this period, however. Tests described below indicate that selection could be driving the results in table 2: borrowers came to the World Bank partly because they were less likely than nonborrowers to generate private credit

16. These results and those that follow also hold when additional lags for the macroeconomic and institutional control variables are included. In the base specification for M2/GDP, for example, the borrower coefficient is 0.017, which is significant at the 1 percent level, while the nonborrower coefficient is 0.002, which is insignificant. For the multilag specification, both the borrower coefficient (0.032) and the nonborrower coefficient (0.017) are significant at the 1 percent level. In both the base and the full-lag specifications, the hypothesis that the borrower and nonborrower coefficients are equal is rejected at the 1 percent level. The full-lag specifications are available from the authors upon request.

17. Loayza and Ranciere (2004) show that a positive long-run relation between financial depth and growth coexists with a largely negative short-run relation.

growth on their own. In the third specification for each indicator in table 2 (models 2.3, 2.8, 2.13, and 2.18), the number of years since first adjustment loan variable is introduced. With the exception of the private credit specifications the *years since first* variable is significantly associated with improved financial development, as expected. This is consistent with the notion that financial reform being a gradual process, countries that received loans earlier in the period experienced greater improvements.

Perhaps more important, when the *years since first* variable enters the regression, the borrower variable is no longer significant. This provides a strong indication that improvements in financial indicators occurred after the receipt of World Bank loans. It, therefore, seems unlikely that borrowing countries were more likely than nonborrowers to improve their financial development indicators regardless of whether they received loans. The coefficient for the *years since first loan* variable (for borrowers) is also statistically distinguishable from that of the nonborrower variable for all indicators except private credit. These patterns also hold when the cumulative number of adjustment loans is included (in models 2.4, 2.9, 2.14, and 2.19). The cumulative loan variable is either insignificant or associated with less financial development across specifications. This result is consistent with findings on repeated structural adjustment lending (Easterly 2005).

In the private credit specifications, *years since first loan* is always positive, though never significant. The borrower variable is also positive and insignificant in all specifications. When the two coefficients are jointly evaluated, the null hypothesis that their sum is equal to zero is rejected at the 5 percent level or better in specifications 2.9 and 2.10. Thus there is some statistical support for the idea that private credit grew in borrowing countries and that the improvements occurred after the receipt of Bank loans. However, the null hypothesis that the difference between the coefficients for borrowers and nonborrowers is zero cannot be rejected. Borrowers did not outperform nonborrowers in private credit growth in any of the specifications in table 2.

The final specification for each indicator (models 2.5, 2.10, 2.15, and 2.20) includes variables that summarize the policy reform areas covered under World Bank loans. Their inclusion does not alter the comparisons between borrowers and nonborrowers, but the interpretation of the results changes slightly: the coefficient on the borrower variable now indicates the impact of participation if there were no conditions attached to loans in any of the policy areas that are controlled for. With the exception of the capital markets development and rural finance variables, the policy variables tend not to be significant across indicators, and the borrower coefficient is similar to that when policy variables are not included in the specification.

While one could come up with explanations for the patterns of the policy coefficients in table 2, it is best not to invest too much effort in this direction. The policy variables are the best that have been put together to study the

effects of reform on financial development, but they have some limitations. First, not all loans that covered a policy area did so in the same way. Some loans may have devoted substantial resources to the policy area, while others may not have done so. Second, because the classifications are based in part on the objectives stated in the documents describing the loans, these measures summarize ex ante indications of planned reform rather than actual ex post reforms. Therefore, the policy variables are a set of coarse controls, included to examine whether the primary results on borrowers versus nonborrowers hold up. The focus is on the simplest decisions—that is, whether or not to borrow and how many loans to take out—rather than on a painstaking qualification of the nature of the reforms to produce variables that are unlikely to explain variation in country-level aggregate financial indicators.

Finally, the control variables that are significant tend to be associated with the financial indicators in the ways one would expect (higher inflation and slower growth retard financial development, for example).[18] CPIA scores are not significant, perhaps because of collinearity with M2/GDP, which is associated with improved financial development for all indicators except interest spreads.[19] The M2/GDP coefficients suggest that the variable could be viewed as a proxy for a country's willingness and ability to undertake financial sector reform.

IV. Selection Effects

The sample of borrowers is unlikely to be random. Selection bias could work in either direction. Countries with the greatest potential for financial development might prefer to pursue reform on their own rather than incur World Bank debt and have to negotiate and adhere to conditions. Alternatively, countries that are ill prepared to achieve financial reform may find themselves ineligible for Bank adjustment loans on mutually acceptable terms.

Nonrandom selection of borrowers can be dealt with in at least two ways. One possibility is to use treatment-effects regressions, which consider the effect of an endogenously chosen binary treatment (in this case, the choice to borrow) on another endogenous continuous variable (in this case, indicators of financial development), conditional on two sets of independent variables. The first set of independent variables is used to estimate a selection equation that describes the participation choice. Information from the selection equation is then used in the financial development regression. The key

18. For robustness year dummy variables were also included in the base regressions to control for global factors that might have affected financial development in all countries. These dummy variables were significant only in the interest rate spread regressions; the qualitative comparisons between borrowers and nonborrowers were similar to those for the base regressions.

19. The CPIA variable becomes positive and significant when M2/GDP is dropped from the private credit specifications.

difficulty is finding an appropriate set of exogenous variables for use in the selection equation.

A second option for facilitating fairer comparisons between borrowers and nonborrowers is propensity-score matching. The intuition underlying this method is that certain country types (for example, the most institutionally sound) are more apt to respond favorably to the treatment than others. To the extent that the control group is more (or less) heavily weighted toward types that are less likely to respond favorably, comparisons with the treatment group can be misleading. The propensity-matching technique therefore matches treatment and control observations based on relevant observable characteristics: apples are compared with apples and oranges with oranges. However, it can be difficult to judge a good match when treatment and control group observations can be compared on multiple observable dimensions. Propensity-score matching can reduce that dimensionality by summarizing the impact of observables in a single equation. A standard probability model (logit or probit) is used to estimate the conditional probability of receiving the treatment (in this case adjustment loans) given a set of covariates. Because the equation is used only to reduce the dimensionality of the conditioning, no behavioral assumptions are attached to it. Thus, unlike in the treatment-effects regressions, the exogeneity of the covariates is not a concern. Contemporaneous variables can be used, and higher-order transformations of those variables are typical.

Applications of these techniques usually involve matching a relatively small set of treatment observations to a subset of a relatively large pool of nontreatment observations. In this case, the set of nontreatment observations is limited, because there are only 38 nonborrowers in the sample (see table A-1).[20] Treatment-effects regressions are, therefore, relied on. Propensity-matching techniques were also applied to these data (the results are presented and discussed in supplemental appendix S-1, available at http://wber.oxfordjournals.org/). In general, propensity matching yields results that favor borrowers a bit more than the base results do.

In many Heckman-type selection models, the dependent variable is observable only for individuals (or households or countries) that received the treatment. In this analysis, indicators of financial development are observable for borrowers and nonborrowers alike. Treatment-effects models are, therefore, estimated in which

$$(2) \qquad Y_i = \alpha + \beta \mathbf{X}_i + \delta Z_i + \varepsilon_i$$

20. In principle, it would be possible to increase the number of observations by going back to the panel data set. However, the nearest matching control group observations would almost certainly be from borrowing countries in years when no adjustment loan was in place. As in a fixed-effects regression with a dummy variable for current borrowing status, this would provide information only about those countries that changed their borrowing status during the period. Because the goal here is to compare countries that borrowed with those that did not, applying propensity matching to the panel data set was not appropriate.

where Y is an indicator of financial development; X is the vector of macroeconomic, institutional, and policy control variables; and Z is the endogenous treatment variable indicating whether or not country i borrowed.[21] As is typical in this literature, the decision to borrow is modeled as the outcome of an unobserved latent variable Z^*, which is a function of exogenous covariates W and a random component u:

$$(3) \qquad Z_i^* = \gamma W_i + u_i$$

The researcher observes that

$$(4) \qquad Z_i = 1, \text{if } Z_i^* > 0 \quad Z_i = 0 \text{ otherwise.}$$

Because there is an element of self-selection in borrowing from the World Bank and the error term of the model that summarizes this choice (equation 3) could be correlated with the error term in the regression of interest (equation 2), a valid set of instruments is needed. These instruments should be highly correlated with the endogenous regressor (the borrowing dummy variable) but contemporaneously uncorrelated with the error term in equation 2 (that is, truly exogenous).

It is difficult to find exogenous variables for use in the selection equation. It is very likely, for example, that proxies for borrowing needs, as reflected in measures of countries' fiscal health and indebtedness, are themselves endogenous.

Appropriate instruments are found by turning to the literature on the political economy of international financial institutions' lending to test whether strong or weak potential reformers are more likely to receive Bank adjustment loans to promote financial sector development.

Vreeland (2004) offers the following propositions regarding IMF lending. The head of the executive branch in a developing country is more likely to enter into a lending arrangement with the IMF when the governmental structure dictates that the executive face a large number of veto players. And the IMF is more likely to lend to countries that have fewer veto players.

The intuition underlying the first proposition is that reform-minded executives in developing countries use IMF support to help overcome opposition to potentially unpopular policies. The idea is that after the executive reaches an agreement with the IMF, failure to achieve reform is more costly, because rejection of those policies is also seen as a rejection of the IMF, which all domestic

---

21. As a robustness check, the total number of loans was also treated as endogenous in specifications that are not presented. The total number of adjustment loans was used to create a dummy variable for "high participation," defined as more than five loans. The high-participation dummy variable was then corrected for selection bias using a treatment-effects regression. These results, which are similar to those for the simple borrower/nonborrower dummy variable, are available upon request from the authors.

politicians and interest groups may recognize as costly.[22] The likelihood that a head of government uses IMF agreements in this way depends on the checks and balances the executive faces. Leaders facing no veto players (that is, dictators) would have no need for IMF support. Leaders facing too many veto players are unlikely to be able to overcome opposition despite IMF support.

Because the IMF prefers to finance successful reform projects, it is likely to be unwilling to enter into agreements with executives who face a large number of veto players. The combination of these two effects should result in a nonlinear relation between the number of veto players and the probability of a loan. In some intermediate range, IMF agreements should be most prevalent, because they are more likely to achieve the desired objective of overcoming the opposition of veto players. World Bank adjustment loans could serve a similar purpose.

These concepts are operationalized using data on the number of checks and balances stipulated in country constitutions (Beck and others 2001). The number of checks and the squared number of checks are included in the selection equations that follow to test whether Vreeland's hypotheses are valid for this data set. If they are, the coefficient for the checks variable should be positive and the checks-squared variable negative in the selection equation. Thus, the likelihood of receiving a World Bank loan for financial sector development would first increase as countries move away from dictatorship (as a result of self-selection by the country) and then decrease when the number of veto players passed some threshold value (as a result of the Bank's selection criteria).

A country's borrowing needs may also affect the likelihood of receiving World Bank adjustment loans. World Bank lending commitments are positively related to an increase in debt service payments and negatively related to the level of international reserves of the borrower (Ratha 2005). As noted, however, contemporaneous measures of countries' fiscal health and indebtedness are likely to be endogenous. Information on fiscal health and indebtedness from 1970 to 1989 is therefore included in the selection equation, which is by definition not contemporaneously correlated with the error term in the financial development regressions (which use data from 1992 to 2003).

It is also conceivable that developing countries—particularly countries with a relatively large stock of World Bank debt—use the proceeds of new Bank loans to repay old loans (evergreening). Beyond some point, however, debt accumulation becomes problematic, making future agreements less attractive, especially from the Bank's point of view. For countries with little past borrowing, predictions about future borrowing are difficult to make. If the lack of borrowing reflects a preference for self-reliance, one would expect little future borrowing. If demand for loans is cyclical, lending would decline during up cycles and increase during down cycles.

---

22. According to Vreeland (2004, p. 2), "The IMF may restrict access to loans, it may preclude debt rescheduling with creditors who require an IMF arrangement to be in good standing, and decreased investment may result if investors take cues from the IMF."

A variety of variables was used to measure countries' past and current indebtedness and overall fiscal health to test these hypotheses. Squared terms enter the selection equation to capture any nonlinearities between past indebtedness and the likelihood of receiving a World Bank loan.

Using historical data to predict whether countries borrowed for financial sector development makes it impossible to estimate a selection effect that varies by year for each country. The likelihood of receiving at least one adjustment loan since 1992 is estimated based on data from 1970 to 89. For this reason, the subscript t does not appear in equation (2). The (largely time-invariant) governmental checks variable is better suited to the cross-sectional regressions than to the panel regressions.

The coefficients from a simple probit regression that uses the borrower dummy variable as the dependent variable are as follows:

$$Borrower_i = \underset{(0.69)}{0.47} + \underset{(0.29)}{0.29 Checks_i} - \underset{(0.04)^*}{0.07 Checks_i^2} - \underset{(0.02)^*}{0.03 Debt_i}$$

$$+ \underset{(0.0002)^*}{0.0003 Debt_i^2} + \underset{(0.0008)^{**}}{0.0016 IMF\ Credit_i}$$

Number of observations: 79
Pseudo R-squared: 0.15
Standard errors in parentheses
*significant at the 10 percent level
**significant at the 5 percent level.

The coefficients from the probit regression and those from the selection equations in the treatment-effects models that follow provide support for the hypotheses in this section. The checks and checks-squared coefficients imply that Bank loans are most likely for an intermediate level of checks. Various measures of past fiscal health were tried, including the current account balance, tax revenues as a percentage of GDP, and the overall government budget balance. Because there are relatively few observations, only two such variables—total IMF borrowing (in millions of constant dollars) and total external debt (as a share of GDP) from 1970 to 1989—are included in the selection equations. The debt variable is negative and its square positive, implying a U-shaped relation with the probability of receiving Bank loans. Thus countries with little past borrowing were more likely to receive loans than those with intermediate levels, possibly indicating that borrowing needs are cyclical. However, heavily indebted countries from 1970 to 89 were the most likely to borrow for financial sector reform from 1992 to 2003, providing additional support for the evergreening hypothesis.[23] The positive coefficient on the IMF borrowing coefficient is also consistent with evergreening.

23. Only a linear and a quadratic term for debt are included in the selection equation, making it impossible to test whether the probability of borrowing eventually declines for extreme levels of indebtedness. The qualitative results for the financial development regressions are similar when the quadratic debt term is excluded from the selection equation.

Although the cross-sectional approach is more promising than the panel approach for handling the selection problems faced here, the approach could make it more difficult to find significant results, because standard errors are likely to be larger in regressions with few observations. Skeptics of the panel results above could argue that because the error terms from multiple observations from the same country are likely to be correlated, the number of independent observations is the same as the number of countries in the data set. Restricting the observation set to the cross-section of countries can, therefore, be viewed as an additional test of whether borrowing countries outperformed nonborrowers in terms of financial development.

For the treatment-effects regressions, growth in Y in year t is calculated as Y_t/Y_{t-1}. The average of annual growth rates over the whole period for each country is used to derive one observation per indicator per country. These country averages are used as dependent variables in the ordinary least squares (OLS) and treatment-effects regressions in table 3. The OLS results in table 3 are similar to those from the panel regressions in table 2, indicating that those results were not solely the product of multiple observations for each country. In particular, M2/GDP grew and cash/M2 declined significantly more rapidly among borrowers. Borrowers' interest spreads declined more rapidly than those of nonborrowers, but the result is not significant in the cross-sectional OLS regression, possibly because there are only 47 observations for that variable. As in the panel regression in table 2, borrowers had slower rates of private credit growth than nonborrowers, although the difference is not statistically significant.

After correction for self-selection using the treatment-effects model, the results show that borrowers outperformed nonborrowers by a wider margin.[24] The change is most pronounced for M2 growth (models 3.2 and 3.3) and private credit growth (models 3.5 and 3.6). At the risk of reading too much into these models, this suggests that the typical World Bank borrower had relatively poor prospects for financial development. Once this is accounted for econometrically, the positive effects of Bank involvement are easier to detect.

Treatment-effects regressions for cash/M2 are more volatile than those for private credit and M2/GDP. In model 3.8, which does not include control variables, the borrower dummy variable is insignificant. Multiple variables are significant in the selection equation, and the likelihood ratio test at the bottom of table 3 indicates that errors from the selection and cash/M2 equations are independent. Thus, the OLS results are valid, and there is no need to perform treatment-effects regression. In model 3.9, which includes institutional and macroeconomic controls, the borrower dummy variable is positive and

TABLE 3. Cross-sectional, OLS, and Treatment-Effect Regressions

Item	Average change in M2/GDP			Average change in private credit/GDP			Average change in cash/M2			Average change in interest rate spread		
	3.1 OLS	3.2 Treat1	3.3 Treat2	3.4 OLS	3.5 Treat1	3.6 Treat2	3.7 OLS	3.8 Treat1	3.9 Treat2	3.10 OLS	3.11 Treat1	3.12 Treat2
Received Bank adjustment loan	0.019* (0.009)	0.075*** (0.014)	0.074*** (0.014)	−0.0005 (0.019)	0.144** (0.020)	0.138*** (0.026)	−0.031*** (0.011)	−0.005 (0.036)	0.059*** (0.017)	−0.128 (0.090)	−0.159 (0.244)	−0.243 (0.183)
CPIA score	0.010 (0.010)		0.012 (0.010)	−0.0001 (0.022)		−0.008 (0.018)	−0.014 (0.012)		−0.010 (0.011)	−0.168* (0.094)		−0.165* (0.091)
Inflation	0.0001* (0.00005)		0.00005 (0.0001)	−0.0001 (0.0001)		−0.00016 (0.00012)	0.00002 (0.0001)		0.00001 (0.00004)	0.0001 (0.001)		0.0002 (0.0005)
Real growth	0.005* (0.003)		0.004 (0.003)	0.015*** (0.006)		0.011*** (0.004)	−0.005 (0.003)		−0.006** (0.003)	0.006 (0.025)		0.008 (0.025)
M2/GDP	−0.0001 (0.0002)		−0.0002 (0.0002)	−0.00005 (0.0005)			0.0001 (0.0003)			0.001 (0.002)		0.002 (0.002)
Constant	0.971*** (0.032)	0.987*** (0.010)	0.939*** (0.031)	0.988*** (0.066)	0.953*** (0.018)	0.950*** (0.055)	1.06*** (0.036)	0.984*** (0.021)	1.00*** (0.035)	1.54*** (0.287)	1.08*** (0.141)	1.59*** (0.283)
Selection equation												
External debt 1970–89 (percent of GDP)		−0.034* (0.018)	−0.029* (0.017)		−0.018 (0.011)	−0.005 (0.005)		−0.035 (0.022)	−0.018 (0.017)		−0.019 (0.035)	−0.011 (0.034)
External debt squared		0.0003** (0.0001)	0.0003* (0.0001)		0.0001 (0.00008)	0.00004** (0.00002)		0.0003* (0.0002)	0.0001 (0.0002)		0.0002 (0.0003)	0.0002 (0.0003)
Number of governmental checks		0.210 (0.228)	0.164 (0.227)		0.234 (0.145)	0.258 (0.170)		0.478 (0.315)	0.616* (0.357)		0.260 (0.538)	0.239 (0.483)
Number of governmental checks squared		−0.050 (0.034)	−0.044 (0.032)		−0.035* (0.020)	−0.034 (0.024)		−0.089** (0.045)	−0.097 (0.061)		−0.056 (0.081)	−0.047 (0.069)
IMF borrowing 1970–89 (constant $ millions)		0.0015** (0.0006)	0.0014** (0.0006)		0.00084* (0.00045)	0.0006 (0.0005)		0.0017** (0.0008)	0.0015** (0007)		0.0013 (0.0010)	
Constant		0.608 (0.541)	0.532 (0.534)		0.137 (0.294)	−0.233 (0.320)		0.140 (0.711)	−0.412 (0.546)	−0.061 (1.05)	−0.061 (1.05)	−0.089 (0.960)

(Continued)

TABLE 3. Continued

Item	Average change in M2/GDP			Average change in private credit/GDP			Average change in cash/M2			Average change in interest rate spread		
	3.1 OLS	3.2 Treat1	3.3 Treat2	3.4 OLS	3.5 Treat1	3.6 Treat2	3.7 OLS	3.8 Treat1	3.9 Treat2	3.10 OLS	3.11 Treat1	3.12 Treat2
Number of countries	73	72	72	73	76	76	70	74	74	47	47	46
R-squared	0.16			0.14			0.18			0.13		
Log likelihood		94.99	98.04		47.90	53.28		82.61	89.96		−36.76	−35.54
Wald Chi squared		29.24	37.08		50.67	45.38		0.02	32.13		0.43	5.71
Prob > Chi squared		0.000	0.000		0.000	0.000		0.8995	0.000		0.5138	0.35
Likelihood ratio test of independent eqns (rho = 0)		10.85	8.92		17.79	16.70		0.39	6.36		0.01	0.34
P > Chi squared		0.001	0.003		0.000	0.000		0.5308	0.0117		0.9191	0.5591

*Significant at the 10 percent level; **significant at the 5 percent level; ***significant at the 1 percent level.

Note: All policy and control variables (CPIA, inflation, growth, deficits/GDP, M2/GDP) are averaged over 1991–2000. External debt and IMF borrowing are annual averages for 1970–89. Number of governmental checks is from Beck and others (2001). Standard errors appear in parentheses.

Source: Authors' analysis based on data sources described in the text and in table 1.

significant, indicating less confidence in the financial system. Although fewer variables are significant in the selection equation than in model 3.8, the likelihood ratio test indicates that errors from the selection and cash/M2 equation are not independent; the treatment-effects results are thus preferred over the OLS results. Because the cash/M2 results are highly sensitive to slight perturbations in either the selection equation or the equation of primary interest, it is difficult to draw a strong conclusion for that variable based on table 3.

In contrast, the selection equations for M2/GDP produce many significant coefficients, and the likelihood ratio test indicates that the treatment-effects model is preferred to the OLS model. In the treatment models, the borrower coefficient is nearly identical whether or not controls are included. This relatively stable pattern of results lends credibility to the conclusion that borrowing countries performed better than nonborrowers on that dimension. Significance levels in the selection equation for private credit are somewhat lower than for M2/GDP, but the coefficients are similar.

As in the simple probit discussed above, the governmental checks and checks-squared coefficients from the selection equations indicate that the probability of receiving an adjustment loan increases from one to three checks but declines thereafter (figure 1). Loans from international financial institutions are therefore most likely for intermediate levels of checks. The debt and debt-squared coefficients indicate that countries with low levels of debt in the 1970s and 1980s were more likely to be borrowers in this data set than those with moderate levels of debt (figure 2), a finding that is consistent with the hypothesis that debt levels may be cyclical. The selection equations therefore provide plausible results in many of the treatment-effects regressions.

FIGURE 1. Probability of Receiving World Bank Adjustment Loan for Financial Sector Development as a Function of Number of Political Checks
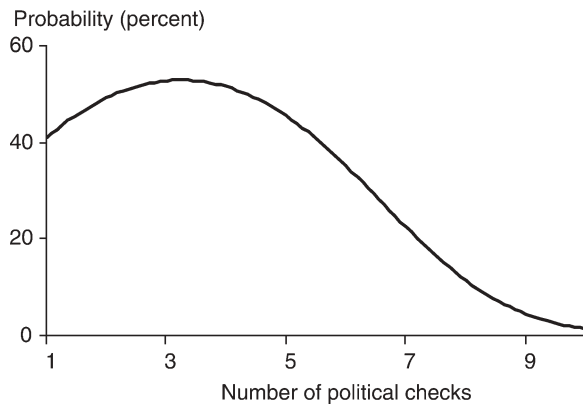
Source: Authors' calculations based on table 3, model 3.2. Data sources are as described in the text and in table 1.

FIGURE 2. Probability of Receiving World Bank Adjustment Loan for
Financial Sector Development as a Function of Level of External Debt
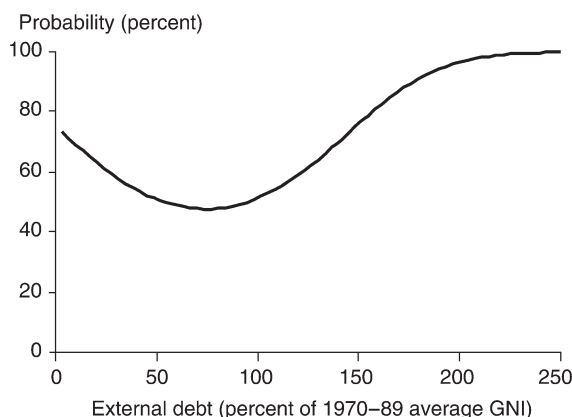
Source: Authors' calculations based on table 3, model 3.2. Data sources are as described in
the text and in table 1.

These regressions reinforce conclusions about the positive association
between borrowing and financial development. For M2/GDP and private
credit/GDP borrowers outperform nonborrowers in the treatment-effects
models. Unlike some of the base models for private credit that do not control
for selection, none of the treatment-effects specifications indicates that bor-
rowers underperform nonborrowers. For interest spreads the hypothesis that
the errors from the first- and second-stage regressions are independent cannot
be rejected, in which case no correction for selection is required. The base
results in table 2 and the OLS results in table 3 are thus valid. The coefficient
for borrowers is negative in both sets of regressions, highly significant in the
base results, and nearly significant in the OLS results. For cash/M2 the
treatment-effects results are unstable; it is thus not possible to draw strong con-
clusions from them.

V. ADDITIONAL ROBUSTNESS CHECKS

A series of tests indicates that the main findings are not driven by the regional
composition of borrowers and nonborrowers and are robust to the inclusion of
variables that measure a country's readiness for and experience with reform
and to the substitution of ratings of financial sector development for quantitat-
ive measures of financial sector development. In the readiness for reform
regressions, CPIA scores are replaced with a measure of the degree to which
countries adhere to the rule of law developed by the International Country
Risk Guide (ICRG).

Several findings emerge from this analysis. First, in Latin America and the
Caribbean financial sector development was stronger in countries that

borrowed from the World Bank than in countries that did not (supplemental appendix S2). Second, the basic pattern of results holds when countries from Europe and Central Asia and Sub-Saharan Africa are dropped from the analysis (supplemental appendix S3). Countries in Europe and Central Asia might have been driving the base results, because many of them began the period of study with low indicators of financial development that improved largely as a result of macroeconomic stabilization. Countries in Sub-Saharan Africa tended to be in the nonborrowing group; the base results might have been picking up differences in financial development between them and countries from other regions. Neither of those concerns is supported by the data. Third, the main findings are robust to the inclusion of controls for whether a country was ready for reform, what other reforms it had already taken when it received financial sector adjustment loans, and what other agencies were involved in its reforms (supplemental appendix S4). Fourth, results are similar when an index of banking and financial sector freedom replaces the quantitative indicators as the dependent variable (supplemental appendix S5): countries that borrowed from the World Bank experienced greater improvement on the index than those that did not.

VI. Conclusions

Evidence based on analysis of a new data set on Bank adjustment loans that supported financial sector reform from 1992 to 2003 indicates that borrowing countries performed better than nonborrowers on multiple measures of banking sector development, including M2/GDP, interest spreads, and cash/M2. They performed worse than nonborrowers on private credit/GDP in OLS regressions. Improvements in financial indicators occurred after the inception of adjustment lending, even after controlling for the adverse selection effects associated with repeated lending to the same country. The main findings hold both in panel regressions that incorporate fixed-country effects and in cross-sectional regressions that use average growth in financial indicators over the full period for each country as dependent variables. The cross-sectional regressions indicate that the panel results are not driven by multiple observations from the same country, which can artificially reduce standard errors.

A series of models accounts for potential selection effects. Nonlinear selection equations capture concepts from the political economy literature on the relations between international financial institutions and developing countries. This approach, therefore, distinguishes countries that prefer not to borrow from these institutions, because they are relatively self-sufficient from those that international financial institutions prefer not to deal with because reform is unlikely to succeed. Addressing nonrandom selection using treatment-effects regressions reveals that the rate of growth of private credit and M2 was significantly larger for borrowers than for nonborrowers. For interest rate spreads test statistics indicate that the errors from the selection and financial

development regressions are independent, obviating the need to correct for nonrandom selection. For cash/M2 the treatment-effects results are highly sensitive to small perturbations in the specification, but some models indicate that no correction for selection effects is necessary.

Robustness checks indicate that the results are not driven by the regional composition of borrowers and nonborrowers and are robust to the inclusion of proxies for countries' readiness and ability to reform. Taken in their entirety, these results suggest that the World Bank adjustment loans studied here had some positive effects on financial sector outcomes.

APPENDIX A-1. Countries That Did and Did Not Receive World Bank Adjustment Loans for Financial Sector Reform between 1992 and 2003

Countries that received World Bank adjustment loans	Countries that did not receive World Bank adjustment loans
Albania	Angola
Algeria	Benin
Argentina	Banngladesh
Armenia	Belarus
Azerbaijan	Botswana
Bolivia	Cambodia
Bosnia and Herzogovina	Chile
Brazil	China
Bulgaria	Congo, Dem. Rep.
Burkina Faso	Costa Rica
Cameroon	Côte d'Ivoire
Cape Verde	Czech Republic
Central African Rep.	Dominican Republic
Chad	Egypt, Arab Rep. of
Colombia	Estonia
Croatia	Ethiopia
Ecuador	Gabon
El Salvador	Gambia, The
Georgia	India
Ghana	Iran
Guatemala	Kenya
Guinea	Lebanon
Guyana	Lesotho
Honduras	Mali
Hungary	Mauritius
Indonesia	Nepal
Jamaica	Nigeria
Jordan	Panama
Kazakhstan	Papua New Guinea
Korea, Rep. of	Paraguay
Kyrgyz Rep.	Senegal
Lao, PDR	South Africa

(Continued)

APPENDIX A-1. Continued

Countries that received World Bank adjustment loans	Countries that did not receive World Bank adjustment loans
Latvia	Sri Lanka
Lithuania	Swaziland
Macedonia	Togo
Madagascar	Trinidad and Tobago
Malawi	Venezuela, R. B. de
Malaysia	Zimbabwe
Mauritania	
Mexico	
Moldova	
Mongolia	
Morocco	
Mozambique	
Nicaragua	
Niger	
Pakistan	
Peru	
Philippines	
Poland	
Romania	
Russia	
Rwanda	
Sierra Leone	
Slovak Republic	
Slovenia	
Tajikistan	
Tanzania	
Thailand	
Tunisia	
Turkey	
Uganda	
Ukraine	
Uruguay	
Uzbekistan	
Vietnam	
Yemen	
Zambia	

Note: The 106 countries in this table are those that appear in at least one regression. The maximum number of countries in any regression is 98.

Source: Independent Evaluation Group database of World Bank loans for financial sector reform.

REFERENCES

Barro, Robert, and Jong-Wha Lee. 2002. *IMF Lending: Who Is Chosen and What Are the Effects*? NBER Working Paper 8951. Cambridge, Mass.: National Bureau of Economic Research.

Barth, James, Gerard Caprio, Jr., and Ross Levine. 2001a. "Banking Systems around the Globe: Does Regulation and Ownership Affect Performance and Stability?" InFrederic Mishkined., *Prudential*

Supervision, What Works and What Doesn't. National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

———. 2001b. "Bank Regulation and Supervision: What Works Best?" Policy Research Working Paper 2725. World Bank, Washington, D.C.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2000. "A New Database on the Structure and Development of the Financial Sector." World Bank Economic Review 14(3):597–605.

———. 2003. "Bank Concentration and Crises." Policy Research Working Paper 3041. World Bank, Washington, D.C.

Beck, Thorsten, Ross Levine, and Norman Loayza. 2000. "Finance and the Sources of Growth." *Journal of Financial Economics* 58(1–2):261–300.

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

Caprio, Jr., Gerard, Patrick Honohan, and Joseph Stiglitzeds. 2001. Financial Liberalization: How Far, How Fast? New York: Cambridge University Press.

Conway, Patrick. 1994. "IMF Lending Programs: Participation and Impact." *Journal of Development Economics* 45(2):365–91.

Corbo, Vittorio, and Morris Goldstein, eds. 1987. Growth-Oriented Adjustment Programs. Washington, D.C.: International Monetary Fund.

Corbo, Vittorio, and Stanley Fischer. 1995. "Structural Adjustment, Stabilization and Policy Reform: Domestic and International Finance." InJere Behrman, and T.N. Srinivasan, eds., Handbook of Development Economics, Vol. III. Amsterdam: Elsevier Science.

Demirgüç-Kunt, Asli, and Enrica Detragiache. 1998. "The Determinants of Banking Crisis in Developing and Developed Countries." *IMF Staff Papers* 45(1):81–109.

Devarajan, Shantayanan, David Dollar, and Torgny Holmgren, eds. 2001. Aid and Reform in Africa. Washington D.C.: World Bank.

Dicks-Mireaux, Louis, Mauro Mecagni, and Susan Schadler. 2000. "Evaluating the Effect of IMF Lending to Low-Income Countries." *Journal of Development Economics* 61(2):495–526.

Easterly, William. 2005. "What Did Structural Adjustment Adjust? The Association of Policies and Growth with Repeated IMF and World Bank Adjustment Loans." *Journal of Development Economics* 76(1):1–22.

Goldstein, Morris, and Peter Montiel. 1986. "Evaluating Fund Stabilization Programs with Multi-Country Data: Some Methodological Pitfalls." *IMF Staff Papers* 33(2):304–44.

Haque, Nadeem Ul, and Moshin S. Khan, 1998. "Do IMF–Supported Programs Work? A Survey of the Cross-Country Empirical Evidence." IMF Working Paper WP/98/169. International Monetary Fund, Washington, D.C.

Honohan, Patrick. 2004. "Financial Development, Growth, and Poverty: How Close Are the Links?" Policy Research Working Paper 3203. World Bank, Washington, D.C.

Hutchinson, Michael M. 2001. A Cure Worse than the Disease? Currency Crises and the Output Costs of IMF–Supported Stabilization Programs. NBER Working Paper 8305. Cambridge, Mass.: National Bureau Economic Research.

IEG (Independent Evaluation Group). 2006. *IEG Review of World Bank Assistance for Financial Sector Reform.* Washington, D.C.: World Bank.

IMF (International Monetary Fund). Various years. International Financial Statistics. Washington, D.C.: International Monetary Fund.

Kaminsky, Graciela L., and Carmen M. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review* 89(3):473–500.

Kapur, Devesh, John P. Lewis, and Richard Webb. 1997. The World Bank: Its First Half Century. Vol. 1: The History. Washington, D.C.: Brookings Institution.

Khan, Moshin S. 1990. "The Macroeconomic Effects of Fund-Supported Adjustment Programs." *IMF Staff Papers* 37(2):195–231.

Killick, Tony, Rumani Gunatilaka, and Ana Marr. 1998. Aid and the Political Economy of Policy Change. London: Routledge.

Knight, Malcolm, and Julio A. Santaella. 1997. "Economic Determinants of IMF Financial Arrangements." *Journal of Development Economics* 54(2):405–36.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2002. "Government Ownership of Banks." *Journal of Finance* 57(1):265–301.

Levine, Ross. 2005. "Finance and Growth: Theory and Evidence." InPhilippe Aghion, and Steven Durlauf, eds., Handbook of Economic Growth. Amsterdam: Elsevier Science.

Levine, Ross, Norman Loayza, and Thorsten Beck. 2000. "Financial Intermediation and Growth: Causality and Causes." *Journal of Monetary Economics* 46(1):31–77.

Levine, Ross, and Sara Zervos. 1998. "Stock Markets, Banks, and Economic Growth." *American Economic Review* 88(3):537–58.

Loayza, Norman, and Romain Ranciere. 2004. "Financial Development, Financial Fragility, and Growth." Policy Research Working Paper 3431. World Bank, Washington, D.C.

Przeworski, Adam, and James Raymond Vreeland, 2000. "The Effect of IMF Programs on Economic Growth." *Journal of Development Economics* 62(2):385–421.

Rajan, Raghuram G., and Luigi Zingales. 1998 "Financial Dependence and Growth." *American Economic Review* 88(3):559–86.

Ratha, Dilip. 2005. "Demand for World Bank Lending." *Economic Systems* 29(4):408–21.

Summers, Lawrence H., and Lant H. Pritchett, 1993. "The Structural-Adjustment Debate." *American Economic Review* 83(2):383–89.

Svensson, Jakob. 2003. "Why Conditional Aid Doesn't Work and What Can Be Done about It?" *Journal of Development Economics* 70(2):381–402.

Van de Walle, Nicolas. 2001. African Economies and the Politics of Permanent Crisis, 1979–1999. Cambridge: Cambridge University Press.

Vreeland, James. 2004. "Institutional Determinants of IMF Agreements." Department of Political Science, Yale University, New Haven, Conn.

World Bank. 1989. World Development Report 1989: Financial Systems and Development. New York: Oxford University Press for the World Bank.

———. Various years. World Development Indicators. Washington, D.C.: World Bank.

HIV Pandemic, Medical Brain Drain, and Economic Development in Sub-Saharan Africa

Alok Bhargava and Frédéric Docquier

Country-level longitudinal data at three-year intervals over 1990–2004 are used to analyze the factors affecting emigration of physicians from Sub-Saharan countries and the effects of this medical brain drain on life expectancy and number of deaths due to AIDS. Data are compiled on emigrating African physicians from 16 receiving Organisation for Economic Co-operation and Development (OECD) countries. A comprehensive longitudinal database is developed by merging the medical brain drain variables with recent data on HIV prevalence rates, public health expenditures, physicians' wages, and economic and demographic variables. A triangular system of equations is estimated in a random effects framework using five time observations for medical brain drain rates, life expectancy, and number of deaths due to AIDS, taking into account the interdependence of these variables. Lower wages and higher HIV prevalence rates are strongly associated with the brain drain of physicians from Sub-Saharan African to OECD countries. In countries in which the HIV prevalence rate exceeds 3 percent, a doubling of the medical brain drain rate is associated with a 20 percent increase in adult deaths from AIDS; medical brain drain does not appear to affect life expectancy. These findings underscore the need to improve economic conditions for physicians in order to retain physicians in Sub-Saharan Africa, especially as antiretroviral treatment becomes more widely available. JEL codes: C33, C5, F22, I12, O11, O55.

The AIDS pandemic in Sub-Saharan Africa is affecting all dimensions of social and economic life. In many countries, gains in life expectancy achieved over the past several decades have been wiped out. Reductions in life expectancy are

detrimental at the macroeconomic level, because they reduce economic growth (Bhargava and others 2001). At the micro level, early parental deaths have created the enormous problem of AIDS orphans (USAID, UNICEF, and UNAIDS 2003; Subbarao and Coury 2004). Orphaned children have lower levels of psychological well-being and school attendance, which is critical for learning and for increasing awareness of HIV transmission routes (Bhargava 2005b). Informed policy formulation to deal with these problems requires analyses of data at the micro and macro levels.

The formulation of policies in the wake of the HIV/AIDS pandemic is complex and would benefit from research on broader biomedical and social science issues. For example, the World Health Organization (WHO) advocates antiretroviral treatment provided by public clinics for patients in HIV/AIDS Stage 4 or with a CD4 cell (type of white blood cell) count below 200 (Gutierrez and others 2004). Studies on the effects of antiretroviral treatment on productivity of undernourished populations can provide insights leading to enhanced treatment strategies. In a similar vein, a recent report by Physicians for Human Rights (2004) emphasizes the need for Sub-Saharan African countries to invest more in training physicians and nurses. The training of additional healthcare staff is hampered by low tertiary enrollment rates in Sub-Saharan African countries; the region could benefit from strategies such as training its physicians in Asia (Bhargava 2005a).

Individual-level surveys in six African countries indicate that more than half of all physicians would like to emigrate to developed countries, in search of better working conditions and more comfortable lifestyles (Awases and others 2003). Very large proportions of healthcare staff—38 percent in Ghana, 45 percent in Cameroon, 49 in Senegal, 58 percent in South Africa and Zimbabwe, 62 percent in Uganda—report being "stressed" by caring for HIV/AIDS patients. The risks associated with caring for HIV/AIDS patients and the possibility of children of healthcare staff contracting HIV as they enter adolescence may exacerbate medical brain drain (Awases and others 2003; Bhargava 2005a). Higher HIV prevalence rates can create a vicious circle, by increasing emigration of physicians and nurses, which can in turn increase deaths from AIDS and the numbers of orphaned children.

The trends underlying the international migration of skilled and unskilled labor are complex and reflect several aspects of labor supply and demand conditions in developing and developed countries (Ozden and Schiff 2006). Emigration of skilled workers in general increases remittances and creates business and information networks that can enhance economic performance in the countries of origin. The net effects of emigration of physicians and nurses from developing countries are likely to depend on the domestic demand for healthcare services over a long period. The AIDS pandemic makes it difficult for Sub-Saharan African countries to withstand attrition of already scarce healthcare workers.

The results of numerous household-level studies of HIV/AIDS in Sub-Saharan Africa indicate that factors such as multiple sex partners and

sexually transmitted infections can exacerbate HIV transmission (Caral and Holmes 2001). HIV prevalence rates constructed from individual-level data that take account of survival time after contracting HIV are useful for estimating life expectancy.

The effects of medical brain drain on indicators of well-being—such as life expectancy and the number of adult deaths due to AIDS—cannot be investigated using data from household surveys; country-level data are needed. Longitudinal data are useful for modeling the relation between HIV prevalence rates and brain drain. Analyses at the country level can provide insights for designing surveys investigating economic and social factors underlying medical brain drain.

Variables reflecting medical brain drain are not available for Sub-Saharan African countries. Fortunately, most statistical and medical agencies in receiving countries of the Organisation for Economic Co-operation and Development (OECD) keep longitudinal information on immigration of physicians (information on nurses is not compiled in the same detailed fashion). Databases such as the *World Development Indicators* (World Bank 2005) contain limited information on HIV prevalence rates. UNAIDS (2006) recently expanded AIDS-related variables on a longitudinal basis for Sub-Saharan African countries. Data on wages from the International Labour Organization (ILO) (2005) and on public health expenditures from the WHO (2006) can be merged with economic and demographic variables from the *World Development Indicators* to create a comprehensive longitudinal database.

This article estimates medical brain drain rates, the number of adult deaths due to AIDS, and life expectancy for 1990–2004 from longitudinal data. Alternative specifications are tested using econometric techniques. The article is structured as follows. Section I briefly describes the data on medical brain drain and other variables. Section II develops the analytical framework for specification of the relations, outlining the likely forms of interdependence among medical brain drain rates, the number of deaths due to AIDS, and life expectancy. Section III describes the empirical models, and section IV the econometric methods used to estimate and test the models. Section V presents the results from estimating random effects models for medical brain drain rates, the number of adult deaths due to AIDS, and life expectancy for Sub-Saharan African countries at five points between 1990 and 2004. Certain exogeneity hypotheses for the variables are tested to assess the validity of the model assumptions. The last section summarizes the article's main conclusions and identifies some areas for further research.

I. The Data

This article examines the 16 most important OECD countries (Australia, Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, New Zealand, Norway, Portugal, Sweden, Switzerland, the United Kingdom, and

the United States) for which longitudinal data on foreign-born physicians are available. These countries account for 93 percent of skilled immigrants in the OECD (Docquier and Marfouk 2006).

The medical brain drain can be evaluated in terms of stocks and rates, following Carrington and Detragiache (1999) and Docquier and Marfouk (2006). The rate of medical brain drain m for country i in time period t can be written as:

$$(1) \qquad m_{it} = \frac{M_{it}}{P_{it} + M_{it}}$$

where M_{it} denotes the stock of physicians from country i working abroad and P denotes the number of physicians working in the home country.

Docquier and Bhargava (2007) developed an annual database covering 1991–2004 from data provided by national agencies. For the data extracted from national censuses, two or three data points are usually available; data for the remaining years were interpolated using a log-linear adjustment. Data on the country of qualification of immigrants are available from medical associations in Canada, France, New Zealand, Norway, United Kingdom, and the United States; these data cover 73 percent of the sample. When the country of qualification could not be determined, data on country of birth were obtained from national censuses and registers in Australia, Austria, Belgium, Denmark, Ireland, and Sweden; these data cover 18 percent of the sample. For countries for which these data were not available (Italy, Germany, Portugal, and Switzerland), emigrants were defined according to their citizenship; these data cover 9 percent of the sample.[1]

The data reveal that medical brain drain rates from 44 Sub-Saharan African countries to 16 OECD countries rose in most countries between 1991 and 2000 (figure 1). Only Angola, Benin, Burkina Faso, Chad, The Gambia, Ghana, Kenya, Malawi, Mauritania, Mozambique, Niger, Senegal, South Africa, and Uganda experienced declines in medical brain drain over this period.

Comprehensive longitudinal data on HIV prevalence rates and the number of adult deaths due to AIDS were recently released by UNAIDS for 1990–2004 (UNAIDS 2006). They reveal skyrocketing levels of HIV prevalence rates in many countries between 1991 and 2000 (figure 2).

Longitudinal information on government expenditures on health by Sub-Saharan African countries are available for 1996–2004 (WHO 2006). Because of missing observations for 1990–95, average government health expenditures during 1996–2004 are treated as time-invariant variables in the

1. Because the medical brain drain rate is treated as an endogenous variable in the estimation, alternative definitions of emigrants are not critical. Moreover, highlighting differences in the definitions across OECD countries should promote a more unified approach in the future.

FɪɢᴜʀE 1. Medical Brain Drain Rates for Sub-Saharan Africa, 1991 and 2000
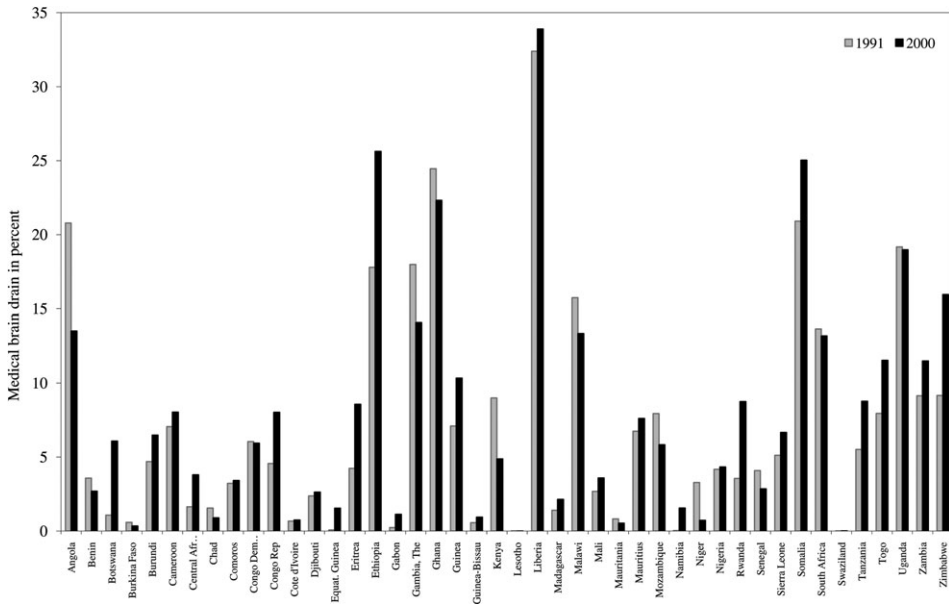
Source: Docquier and Bhargava 2007.

FɪɢᴜʀE 2. HIV Prevalence Rates in Sub-Saharan Africa, 1991 and 2000
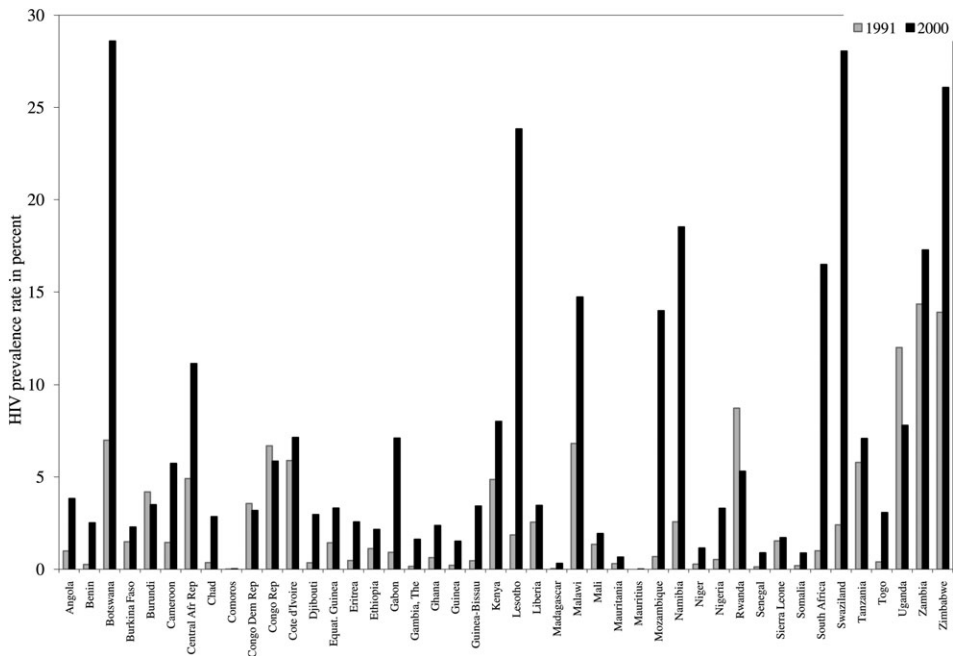
Source: UNAIDS 2006.

econometric modeling. The ILO (2005) provides data on physicians' wages expressed in terms of average physicians' wages in the United States (see also Vujicic and others 2004).

Data on gross domestic product (GDP) (in 2000 dollars) are from the *World Development Indicators* (World Bank 2005). Additional information from UNESCO (2004) and the World Bank (2006) were used to reduce the numbers of missing observations.

Estimates of the proportion of the labor force with secondary or tertiary education are from Barro and Lee (2001), Docquier and Marfouk (2006), and Cohen and Soto (2007). The variables in the database were averaged to create five three-year interval time points over 1990–2004. Alternative data sets were created using two- and four-year averages; the three-year average figures are used here because of the structure of the data and the stochastic properties of the variables (Bhargava 2001); interpolations of variables such as life expectancy can create difficulties for econometric modeling.

A steady increase in HIV prevalence rates and the number of adult deaths due to AIDS is evident from the sample means (table 1). The number of physicians emigrating from Sub-Saharan African countries rose between 1990 and 2004, as did medical brain drain rates. Average life expectancy fell about two years between 1991 and 2003. There was an increase over time in net school enrollment rates in primary and secondary education. Physicians' wages in Sub-Saharan Africa declined slightly relative to physicians' wages in the United States.

II. ANALYTICAL FRAMEWORK

The relations between HIV prevalence rates, medical brain drain, number of adult deaths due to AIDS, and life expectancy estimated using country-level data are of interest to policymakers. The nature of HIV transmission through sexual intercourse and the lags between contracting HIV and the onset of AIDS have important implications for the specification of macroeconometric models that go beyond the usual difficulties of deducing the effects of disease prevalence rates on aggregate economic indicators. It is even more complex to explain HIV prevalence rates at the country level, where information on the average number of sex partners, the prevalence of sexually transmitted infections, and patterns of migrant labor are unavailable. The effects of HIV prevalence rates on medical brain drain can nevertheless be analyzed, while allowing for the possibility that the HIV prevalence rate is potentially an endogenous variable in the system and may be influenced by medical brain drain.

Most rural residents in Sub-Saharan Africa have limited access to basic healthcare, and only a small proportion of people with HIV receive antiretroviral treatment. The lags between HIV infection and the onset of AIDS are thus likely to depend mainly on the natural rate of disease progression. That

TABLE 1. Sample Means and Standard Deviations of Selected Variables in Sub-Saharan Africa, 1990–2004 (3-year averages)

Variable	1991[a]		1994		1997		2000		2003	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
HIV prevalence (percent)	2.98	3.72	5.037	5.592	6.594	7.047	7.073	7.907	7.085	8.096
Number of AIDS deaths	3,960.82	7,777.64	10,253.64	16,672.51	19,713.29	28,002.43	30,245.11	40,974.4	36,963.61	54,538.9
Number of physicians per 1,000 population	0.15	0.24	0.15	0.22	0.16	0.24	0.16	0.24	0.16	0.24
Number of physicians emigrating	151.35	528.16	169.10	586.36	188.77	650.39	211.26	687.94	269.75	925.89
Medical brain drain[b]	0.09	0.11	0.10	0.12	0.10	0.11	0.10	0.11	0.11	0.12
Life expectancy (years)	50.79	8.21	50.13	8.28	49.36	7.75	48.59	7.88	48.21	8.56
GDP per capita (2000$)	752.84	1,134.56	749.14	1,179.60	818.13	1,292.17	876.92	1,400.12	906.38	1,415.34
GDP per capita growth rate (percent)	1.67	1.55	1.95	2.13	4.18	8.23	2.81	2.86	3.63	3.74
Population (thousands)	11,536	16,747	12,511	18,310	13,515	19,799	14,539	21,223	15,581	22,692
Literacy rate (percent of adult population)	50.76	19.73	53.73	19.72	56.76	19.62	59.70	19.46	62.57	19.21

(Continued)

TABLE 1. Continued

Variable	1991[a]		1994		1997		2000		2003	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Primary-school enrollment (percent)	58.11	25.01	60.20	25.41	61.00	23.31	62.57	20.98	67.00	19.71
Secondary-school enrollment (percent)	19.11	15.95	20.64	17.34	22.06	18.22	23.91	18.72	25.58	18.62
Percentage of population with secondary or tertiary education	13.2	10.2	14.8	11.7	16.0	12.6	17.0	12.9	18.0	13.7
Government health expenditure (percent of GDP)	4.32	1.31	4.68	1.42	4.84	1.75	4.87	1.79	4.98	1.75
Physicians' wages in home country/physicians' wages in United States	0.124	0.133	0.104	0.110	0.098	0.106	0.082	0.072	0.083	0.074
Investment (percent of GDP)	18.98	9.67	20.93	11.19	21.77	13.64	21.36	13.57	20.04	13.45

Note: There are at most 47 countries in the sample at each time point.

[a]Based on medical brain drain rates in 1991 and 1992.

[b]Ratio of number of physicians from Sub-Saharan Africa working in 16 OECD countries to the number of physicians working in Sub-Saharan and OECD countries.

Source: Docquier and Bhargava 2007.

progression is accelerated by poor nutritional status, including anemia (Belperio and Rhew 2004). Survival time for individuals contracting HIV in Sub-Saharan Africa is about 10 years (Jaffar and others 2004). These biological aspects are useful for interpreting empirical evidence from analyses of country-level data.

Medical brain drain should increase the number of adult deaths due to AIDS, especially among urban populations that receive some form of medical care. However, a 15-year period may not be long enough to observe the adverse effects of medical brain drain on HIV prevalence rates or on life expectancy, which is already low in Sub-Saharan Africa.

It is useful to outline the pathways through which HIV prevalence can depress economic activity. In developed countries, where antiretroviral treatment is widely available, high HIV prevalence rates may not significantly reduce productivity. By contrast, in developing countries AIDS-related work absenteeism is likely to be common. The effects of HIV prevalence rates on aggregate indicators of economic activity may be dampened somewhat by the fact that many people in Sub-Saharan Africa are engaged in household tasks and work on their own farms; when they fall ill, other household members can step in to perform their tasks (Bhargava 1997). Thus while the effects of HIV prevalence rates on GDP growth rates may not be evident, reductions in life expectancy due to HIV/AIDS can hamper economic growth (Bhargava and others 2001).

The relations among HIV prevalence, medical brain drain, the number of adult deaths caused by AIDS, and life expectancy can be embodied in a tri-angular system of equations. First, one can specify a model for medical brain drain that is likely to be affected by physicians' wages, GDP levels, and HIV prevalence rates, which increase the risks in the work environment. Second, HIV prevalence rates and medical brain drain are likely to increase the number of adult deaths due to AIDS and reduce life expectancy. Longitudinal data are available at five time points for estimating econometric models for the outcomes medical brain drain rates, number of adult deaths due to AIDS, and life expectancy; the simultaneous equation system is a "block triangular" system (the triangular system of equations contains longitudinal observations that are themselves triangular in form because of time orderings; see equation (8) below).

Appropriate techniques for efficiently estimating all the equations simultaneously are not available in the econometrics literature. Instead, each model can be estimated separately using longitudinal estimation methods that take into account the potential endogeneity of variables such as HIV prevalence rates. The models for medical brain drain rates, number of adult deaths due to AIDS, and life expectancy should be dynamic in nature, because past realization of these variables is critical for explaining their current levels.

III. THE EMPIRICAL MODELS

The first equation in the simultaneous equations system is the model for the rate of medical brain drain. It is given by

$$(2) \quad \text{Logistic } (\textit{medical brain drain rate})_{it} = a_0$$
$$+ a_1 \ln(\textit{physicians wage}/\textit{USA ratio})_{it}$$
$$+ a_2 \ln(\textit{school enrollment secondary})_{it} + a_3 \ln(\textit{GDP per capita})_{it}$$
$$+ a_4 \ln(\textit{HIV prevalence rate})_{it} + a_5 \text{ Logistic } (\textit{medical brain drain rate})_{it-1}$$
$$+ u_{1it} \quad (i = 1, \dots, N; t = 2, 3, 4, 5)$$

Transforming the variables into logs is generally useful; the initial observations on medical brain drain rate as well as the lagged dependent variables are treated as endogenous variable in the system (that is, correlated with the errors u_{1it}) (Anderson and Hsiao 1981; Bhargava and Sargan 1983). The HIV prevalence rate is treated as an endogenous variable, and exogeneity assumptions are tested using likelihood ratio tests.

A dynamic model for HIV prevalence rate is estimated to investigate possible reverse causality (the possibility that higher medical brain drain rates predict higher HIV prevalence). Four-year averages at three time points are used in specification 4 to assess the robustness of the results from analyzing three-year averages. The errors, u_{1it}'s, can be decomposed in a random effects fashion as

$$(3) \qquad\qquad\qquad u_{1it} = \delta_i + v_{it}$$

where the δ_i terms are country-specific random effects that are distributed with zero mean and finite variance, and the v_{it} terms are independently distributed random variables with zero mean and finite variance. Equation (3) is a special case of the assumption invoked in the empirical modeling that the variance-covariance matrix of u_{1it} is an unrestricted positive definite matrix (Bhargava and Sargan 1983).

The empirical models for the number of adult deaths due to AIDS and for life expectancy are similar. The model for the number of deaths due to AIDS is written as

$$\ln(\textit{number of adult deaths due to AIDS})_{it} = b_0 + b_1 \ln(\textit{population})_i$$
$$+ b_2 \ln(\textit{proportion of labor force with secondary or tertiary education})_{it}$$
$$(4) \quad + b_3 \ln(\textit{GDP per capita})_{it} + b_4 \ln(\textit{HIV prevalence rate})_{it}$$
$$+ b_5 \ln(\textit{medical brain drain rate})_{it} + b_6 \ln(\textit{deaths due to AIDS})_{it-1}$$
$$+ u_{2it} \qquad (i = 1, \dots, N; t = 2, 3, 4, 5).$$

In equation (4), the country's entire population is accounted for, and medical brain drain and HIV prevalence rates are treated as endogenous variables. An interaction term between HIV prevalence and medical brain drain rates is included in an extended version of the model (specification 2, table 3). In view of the lags between contracting HIV and the onset of AIDS, lagged HIV prevalence rate is used instead of the current rate in specification 3 (table 3). In specification 4, the medical brain drain rate is replaced by the number of physicians in the home country (per 1,000 people) and the number working abroad. Although nonlinearities with respect to these variables are taken into account, it is difficult to simultaneously address all possible nonlinearities given the limited number of countries and time periods in the sample. In the model for life expectancy (table 4), specification 2 replaces the medical brain drain rate with the number of physicians in the home country and the number working abroad. Net percentages of school enrollment in primary education is a regressor in the model for life expectancy, because even a few years of primary education are associated with lower child mortality in Sub-Saharan Africa (Bhargava and Yu 1997).

IV. The Econometric Framework

The methodology used to estimate dynamic random effects models in which some explanatory variables are endogenous was developed by Bhargava and Sargan (1983) and Bhargava (1991). Let the dynamic model be given by

$$(5) \qquad y_{it} = \sum_{j=1}^{m} z_{ij}\gamma_j + \sum_{j=1}^{n_1} x_{1ijt}\beta_j + \sum_{j=n_1+1}^{n} x_{2ijt}\beta_j + \alpha y_{it-1} + u_{it}$$

where the z terms are time-invariant variables, the x_1 terms are exogenous time-varying variables, and the x_2 terms are endogenous time-varying variables. In the model for medical brain drain rates, for example, the HIV prevalence rate is likely to be an endogenous time-varying variable (that is, is included in the x_{2ijt}'s); unobserved factors affecting HIV prevalence rates may be influenced by country-specific random effects (δ_i) in equation (2).

It is useful to rewrite the dynamic model in a simultaneous equations framework by defining a reduced form for the initial observations and a system of ($T-1$) structural equations for the remaining time periods (Bhargava and Sargan 1983):

$$(6) \qquad y_{i1} = \sum_{j=0}^{m} z_{ij}\zeta_j + \sum_{j=1}^{T}\sum_{k=1}^{T} v_{jk}x_{ijk} + u_{i1}(i=1,\ldots,N)$$

and

$$
(7) \quad \underset{(T-1)\times T}{B} \cdot \underset{T\times N}{Y'} + \underset{(T-1)\times(m+1)}{C_z} \cdot \underset{(m+1)\times N}{Z'} + \underset{(T-1)\times nT}{C_x} \cdot \underset{nT\times N}{X'} = \underset{(T-1)\times N}{U'}
$$

In equations (6) and (7), Y, Z, and X are matrices containing observations on the dependent, time-invariant, and time-varying explanatory variables; dimensions of the matrices are written below the respective symbols. B is a $(T-1) \times T$ lower triangular matrix of coefficients:

$$
(8) \quad B_{ii} = \alpha, B_{i,I+1} = -1, B_{ij} = 0 \quad \text{otherwise} \quad (i = 1, \ldots, T-1; j = 1, \ldots, T).
$$

Matrix C_z contains coefficients of time-invariant regressors. Matrix C_x contains coefficients of time-varying regressors. Matrix U contains the error terms.

The profile log-likelihood functions of the model in equation (2) can be optimized using a numerical scheme such as E04 JBF from Numerical Algorithm Group (1991). Assuming that the number of countries is large but the number of time observations is fixed, asymptotic standard errors of the parameters are obtained by approximating second derivatives of the function at the maximum. The random effects decomposition in equation (3) can be tested in this framework using likelihood ratio tests. Given five time observations, under the null hypothesis of the random effects decomposition, the likelihood ratio statistic is distributed for a large N as a chi-squared variable with 12 degrees of freedom.

In short panels, it is reasonable to assume that a variable such as the HIV prevalence rates in equation (2) is correlated with the country-specific random effects, δ_i:

$$
(9) \qquad\qquad\qquad x_{2ijt} = \lambda_j \delta_i + x^*_{2ijt}
$$

where x^*_{2ijt} is uncorrelated with δ_i, and the δ_i terms are randomly distributed variables with zero mean and finite variance, as in equation (3). This correlation pattern was invoked by Bhargava and Sargan (1983); the advantage in employing equation (9) is that deviations of the x_{2ijt}'s from their time means that

$$
(10) \qquad x^+_{2ijt} = x_{2ijt} - \bar{x}_{2ij} \qquad (t = 2, \ldots, T; j = n_1 + 1, \ldots, n; i = 1, \ldots, N)
$$

where

$$
(11) \qquad \bar{x}_{2ij} = \sum_{t=1}^{T} x_{2ijt}/T \qquad (j = n_1 + 1, \ldots n; i = 1, \ldots, N)
$$

can be used as $[(T - 1)n_2]$ additional instrumental variables to facilitate identification and estimation of the parameters. While more general correlation patterns between the explanatory variables and errors are possible (Bhargava 1991), the identification conditions may not be met in dynamic models estimated from short panels. Exogeneity hypotheses can be tested using likelihood ratio tests. Given five time observations, the likelihood ratio statistic for testing correlation between the random effects and time means of HIV prevalence rates in equation (2) is distributed as a chi-squared variable with five degrees of freedom. Fixed-effects estimators (with dummy variables for countries) can be used to circumvent certain endogeneity problems, but the rise in the number of parameters that accompanies the increase in sample size leads to the problem of incidental parameters (Neyman and Scott 1948).

V. Empirical Results

This section presents the results from estimating random effects models for medical brain drain rates, the number of adult deaths due to AIDS, and life expectancy for Sub-Saharan African countries at five time points over the period 1990–2004.

Effect on Medical Brain Drain Rates

Specification 1 of the model in equation (2) treats the HIV prevalence rate as an exogenous variable; it does not restrict the variance-covariance matrix of the errors in the simple random effects fashion (equation 3) (table 2). The variance-covariance matrix is restricted to be that of the random effects model in specification 2, and the between/within variance ratio and the within variance are estimated. Specification 3 treats HIV prevalence rates as an endogenous variable that is correlated with the country-specific random effects. Specification 4 analyzes four-year averages.

These specifications yield several important findings. First, the likelihood ratio statistic for testing specification 2 against specification 1 is 50.96; it is distributed as a chi-squared variable with 12 degrees of freedom. The test rejects the null hypothesis of simple random effects decomposition (the 5 percent critical level is 21.0). The likelihood ratio statistic for testing the exogeneity of the HIV prevalence rate assumes the value 12.04, which is close to the 5 percent critical limit of 11.10. Thus the null hypothesis is rejected. In view of the closeness of the statistic to its critical limit, it is perhaps not surprising that the results from specifications 1 and 3 are close. The discussion focuses on the results from specification 3.

Second, the coefficient of the variable measuring physicians' wages in Sub-Saharan African countries as a percentage of physicians' wages in the United States is significant at the 5 percent level: countries with higher physician wages have lower emigration rates. These findings are consistent with the responses of health-care personnel who want to emigrate, among whom large

TABLE 2. Maximum Likelihood Estimates for Medical Brain Drain from Sub-Saharan African Countries Explained by Socioeconomic Variables and HIV Prevalence Rates, 1990–2004

Explanatory variable	Specification 1		Specification 2		Specification 3[a]		Specification 4[b]	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Constant	−0.465**	0.146	−0.815**	0.224	−0.485**	0.151	−1.055**	0.201
ln(physicians' wages in home country/physicians' wages in United States)	−0.036**	0.019	−0.057**	0.030	−0.035**	0.019	−0.066**	0.032
ln(percent school enrollment secondary)	0.115**	0.045	0.236**	0.069	0.124**	0.048	0.248**	0.081
ln(GDP per capita)	−0.035	0.033	−0.102**	0.057	−0.037	0.034	−0.085	0.066
ln(HIV prevalence)	0.079**	0.017	0.082**	0.022	0.071**	0.021	0.107**	0.036
Lagged dependent variable	0.921**	0.028	0.811**	0.053	0.914**	0.031	0.769**	0.035
Between/within variance ratio			0.445	0.354				
Within variance			0.051					
2 (maximized log-likelihood function)	512.75		461.79		570.55		300.45	
Chi-squared test for random effects decomposition (12 degrees of freedom)			50.96**					
Chi-squared test for exogeneity of HIV prevalence rate (5 degrees of freedom)					12.04**			

Dependent variable: Logistic *(medical brain drain rate)*.

**Significant at the 5 percent level.

Note: Data on 39 countries, with five time observations at 3-year intervals.

[a]HIV prevalence rate is treated as an endogenous variable.

[b]Specification uses 4-year averages at three time points.

Source: Authors' estimation results.

majorities—68 percent in Cameroon, 77 percent in Zimbabwe, 78 percent in South Africa, 84 percent in Uganda, 85 percent in Ghana—report a desire to earn more as a motivation (Awases and others 2003).

Third, net enrollment in secondary education is a positive and significant predictor of medical brain drain, with an estimated short-run elasticity of 0.12. This result is not surprising, as higher enrollments in secondary education entail greater expenditures on education; physicians educated in such environments are likely to have better emigration prospects.

Fourth, the HIV prevalence rate is a significant predictor of medical brain drain, with the short-run elasticity (0.07) robust across the first three specifications. Moreover, the coefficient of the lagged dependent variable is estimated at 0.91 in specification 3, indicating that the long-run impacts of the explanatory variables are about 11 times greater than the short-run coefficients. Thus the long-run elasticity of the medical brain drain rate with respect to HIV prevalence is about 0.8. This means that a doubling of the HIV prevalence rate implies an 80 percent increase in the medical brain drain rate in the long run. This is a large effect, with important policy implications, especially given that the average number of physicians per 1,000 people is only 0.15 in Sub-Saharan Africa (see table 1). Furthermore, higher ratios of physicians to the population will be needed as more people with HIV develop AIDS.

Fifth, the large estimated coefficients of the lagged dependent variable suggest that emigration patterns in Sub-Saharan African countries are becoming well established, presumably as a result of stable demand from OECD countries for physicians trained in specific countries. While specification 2 is rejected in favor of specification 1 by the likelihood ratio test, the estimated between/within variance ratio is not significant in specification 2, possibly because of the relatively small number of countries in the sample.

Sixth, the results from specification 4 employing four-year averages are similar to those from specification 1, indicating robustness of the results from three-year averages at five time points. In fact, use of four-year averages entails a loss of information, because the data from 2002, 2003, and 2004 cannot be used in specification 4.

Finally, a model similar to equation 2 was estimated for HIV prevalence rates with medical brain drain rate as an explanatory variable to investigate possible reverse causality. Coefficients of medical brain drain rate were not significant in any of the specifications, thereby supporting the model formulation and exogeneity assumptions.

Effect on Numbers of Adult Deaths Due to AIDS and on Life Expectancy

The results for the numbers of adult deaths due to AIDS are shown for four specifications (table 3). Specification 2 includes an interaction term between medical brain drain and HIV prevalence rates; this term is included because higher HIV prevalence rates can exacerbate the effects of medical brain drain on adult deaths caused by AIDS. In specification 3, the current HIV prevalence

TABLE 3. Maximum Likelihood Estimates for Number of Adult Deaths due to AIDS, 1990–2004

Explanatory variable	Specification 1[a] Coefficient	Specification 1[a] Standard error	Specification 2[a] Coefficient	Specification 2[a] Standard error	Specification 3[b] Coefficient	Specification 3[b] Standard error	Specification 4[c] Coefficient	Specification 4[c] Standard error
Constant	−2.833**	0.422	−3.412**	0.165	−5.011**	0.203	3.553**	0.481
ln(population)	0.450**	0.017	0.480**	0.006	0.635**	0.008		
ln(proportion of population with secondary or tertiary education)	−0.045	0.027	−0.005	0.026	0.024	0.016	−0.114	0.085
ln(GDP per capita)	−0.014	0.025	−0.072**	0.025	−0.008	0.006	−0.093**	0.047
ln(HIV prevalence)	0.542**	0.039	0.778**	0.038			0.660**	0.052
ln(HIV prevalence)$_{-1}$					0.836**	0.016		
ln(medical brain drain rate)	0.095**	0.019	−0.059**	0.023	−0.080**	0.020		
ln(number of physicians in home country/1,000 people)							−0.007	0.173
ln(number of physicians abroad)							0.113**	0.035
[ln(number of physicians in home country/1,000 people)]^2							0.001	0.028
[ln(number of physicians abroad)]^2							0.014**	0.006
ln(medical brain drain) × ln(HIV prevalence)			0.049**	0.008				
ln(medical brain drain) × ln(HIV prevalence)$_{-1}$					0.061**	0.005		

	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
ln(*HIV prevalence*) × ln (*physicians abroad*)							−0.032**	0.011
Lagged dependent variable	0.510**	0.012	0.498**	0.001	0.357**	0.008	0.535**	0.015
2 × (maximized log-likelihood function)	1,243.35		1,258.20		1,282.88		835.39	
Chi-squared test for exogeneity of medical brain drain and HIV prevalence rates (10 degrees of freedom)	104.03**		116.63**		96.94**			

Dependent variable: ln(*number of adult deaths caused by AIDS*).

**Significant at the 5 percent level.

Note: Data on 39 countries with five time observations at 3-year intervals are used in the estimation.

[a]Medical brain drain rate (ratio of number of physicians from Sub-Saharan Africa working in 16 OECD countries to the number of physicians working in Sub-Saharan and OECD countries) and HIV prevalence rates are treated as endogenous variables.

[b]HIV prevalence rate is lagged one period.

[c]Physicians in home country per 1,000 people and physicians abroad are entered as separate variables.

Source: Authors' estimation results.

TABLE 4. Maximum Likelihood Estimates from Dynamic Random Effects Models for Life Expectancy in Sub-Saharan African Countries Explained by Socioeconomic Variables and HIV Prevalence and Medical Brain Drain Rates, 1990–2004

	Specification 1		Specification 2	
Explanatory variable	Coefficient	Standard error	Coefficient	Standard error
Constant	1.117**	0.005	1.012**	0.004
ln(government health expenditure as percent of GDP)	0.040**	0.003	0.029**	0.003
ln(percent primary school enrollment)	0.022**	0.009	0.022**	0.002
ln(GDP per capita)	0.001	0.007	0.009**	0.004
ln(HIV prevalence)	−0.027**	0.003	−0.026**	0.004
ln(medical brain drain rate)	−0.005	0.003		
ln(physicians in home country, per 1,000 people)			−0.0001	0.003
ln(physicians abroad)			0.0001	0.002
Lagged dependent variable	0.662**	0.002	0.700**	0.001
2 × (maximized log-likelihood function)	1558.80		1588.15	
Chi-squared test for exogeneity of medical brain drain rate or physicians abroad (5 degrees of freedom)	16.64**		24.46**	

Dependent variable: ln(life expectancy).

**Significant at the 5 percent level.

Note: Data on 39 countries with five time observations at 3-year intervals in are used in the estimation. The medical brain drain rate (ratio of number of physicians from Sub-Saharan Africa working in 16 OECD countries to the number of physicians working in Sub-Saharan and OECD countries) is treated as an endogenous variable.

Source: Authors' estimation results.

rate is replaced by the value lagged one period (that is, three years). Specification 4 replaces the medical brain drain rate with the number of physicians in the home country and those working abroad, taking into account some nonlinearities. Constraints on the variance-covariance matrix implied by the random effects decomposition are rejected in all specifications.

The results in table 3 show the importance of HIV prevalence rates for the number of adult deaths due to AIDS. Although the short-run elasticity from specification 1 is just 0.54, the long-run elasticity is close to 1.0. In view of the lack of antiretroviral treatment for the vast majority of people with HIV during the sample period, it is not surprising to find large long-run effects of HIV prevalence rates on adult deaths.

In specification 1, the medical brain drain rate is a significant predictor of the number of deaths due to AIDS, with a short-run elasticity of 0.1 and a long-run elasticity of 0.2. Even where antiretroviral treatment is not provided, reduced access to physicians who provide basic care, such as treatment of

tuberculosis and other illnesses that accompany HIV, is likely to exacerbate the progression of the illness. Thus at the aggregate country level, medical brain drain appears to reduce the survival time of people with HIV.

In specifications 2 and 3, the interaction term medical brain drain and HIV prevalence rates is significant, indicating that the effects of medical brain drain are likely to depend on national HIV prevalence rates. The results from specifications 2 and 3 are close, indicating small differences between using current and lagged HIV prevalence rates. While the effects of medical brain drain rate on the number of deaths due to AIDS become positive after HIV prevalence rates cross the 3 percent threshold, it is difficult to precisely estimate the standard errors in nonlinear models with modest sample sizes. Nevertheless, the results in table 3 suggest the importance of devising economic and other incentives for reducing medical brain drain as HIV prevalence rates cross the 3 percent threshold (average HIV prevalence rates after 1991 in table 1 are above this threshold).

The results from specification 4 show that the number of physicians abroad is a significant predictor of the number of adult deaths due to AIDS, while the number of physicians in the home country is not statistically significant. It is difficult to precisely estimate the coefficients in the presence of nonlinearities among explanatory variables using only five time observations on 39 countries.

The results on life expectancy are presented for two specifications (table 4). Specification 2 replaces the medical brain drain rate with the numbers of physicians in the home country and those working abroad. The exogeneity hypotheses for the medical brain drain rate in specification 1 and for the number of physicians abroad in specification 2 are rejected. Because specification 1 is a special case of specification 2, another likelihood ratio test can be applied to choose between these two specifications. Specification 2 is the preferred model in table 4, because the likelihood ratio statistic assumes the value 29.35 and is asymptotically distributed as a chi-squared variable with one degree of freedom.

Government health expenditures expressed as a percentage of GDP have positive coefficients in both specifications and are statistically significant. The data on this variable were compiled annually from 1996; the estimation method treated average expenditures as a time-invariant variable. The net enrollment in primary education is also estimated with positive and significant coefficients in both specifications in table 4. This result is not surprising, as school enrollment rates are very low, especially in rural areas of Sub-Saharan African countries, and child survival and life expectancy are likely to increase with education. The fact that the coefficient of the GDP variable is insignificant in specification 1 may seem surprising in view of the relation between GDP levels and life expectancy (Preston 1976). However, the coefficient of GDP is statistically significant in specification 2, which is a more general formulation. It is also likely that the HIV/AIDS pandemic is driving recent variations in life expectancy in Sub-Saharan Africa.

The estimated coefficients of HIV prevalence rates are significant and negative in both specifications shown in table 4. The short-run elasticity of life expectancy with respect to the HIV prevalence rate is 0.02, and the long-run elasticity is 0.07. Thus a doubling of the HIV prevalence rate would reduce life expectancy to 7 percent in the long run. The estimated coefficient for the medical brain drain rate is negative in specification 1, but not statistically significant. In the long run, the increase in the number of AIDS deaths predicted by the medical brain drain rate also implies lower life expectancy. However, the coefficients of the numbers of physicians in the home country and abroad are statistically insignificant in specification 2 after controlling for population size.

The coefficients of the lagged dependent variables are significant in both specifications but lower than those obtained for medical brain drain rates in table 2. This result may be due to the differences in availability of antiretroviral treatment across Sub-Saharan Africa; longitudinal data on the percentages of people with HIV receiving antiretroviral treatment are not available in the databases. Overall, the empirical models for adult deaths due to AIDS and for life expectancy exhibit a triangular form, although the medical brain drain rate is not a significant predictor of life expectancy. It is plausible that a combined variable for the emigration of physicians and nurses (on whom data were not available) might have had greater explanatory power because of the role of nurses in administering treatments.

VI. Conclusions

What effect has Sub-Saharan Africa's AIDS pandemic had on the emigration of physicians? How has the medical brain drain affected life expectancy and the number of adult deaths due to AIDS?

Lower physicians' wages and higher HIV prevalence increase emigration of physicians from Sub-Saharan Africa. Empirical analyses at the country level that take into account country characteristics and unobserved heterogeneity are important for policy formulation. In particular, as is evident from the results here, physicians' wages and living standards need to be improved to better reflect the workload and risks in high HIV-prevalence environments. Some observers have suggested schemes in which developed countries would compensate Sub-Saharan African countries for the loss of their healthcare staff (Bhargava 2005a). Whether or not such schemes are adopted, it is imperative that international agencies implement pilot programs for reducing the emigration of physicians from Sub-Saharan Africa. Such policies would stop the vicious circle of higher HIV prevalence rates leading to increased emigration of physicians, which in turn lowers the quality and quantity of care for AIDS patients, increasing death rates and the numbers of orphaned children.

The medical brain drain rate is associated with an increase in the number of adult deaths after HIV prevalence rates cross the 3 percent mark. A doubling of the medical brain drain rate is associated with a 20 percent increase in adult

deaths from AIDS. These findings underscore the importance of retaining physicians in Sub-Saharan African countries, especially as antiretroviral treatment becomes more widely available.

The results from the model for life expectancy show the importance of lower HIV prevalence rates and to some extent higher government expenditures on health. Increasing life expectancy is beneficial for the welfare of children who might otherwise be orphaned by AIDS and for investments in education and training (Bhargava 2005b; Bell, Devarajan, and Gersbach 2006). The results presented here do not find a significant association between medical brain drain and life expectancy. Analyses that consider a longer-time frame may find evidence for such effects.

It would be helpful if OECD countries would compile information on the immigration of nurses, who play an important role in providing care to HIV patients. Inclusion of such data could affect some of the results presented here.

References

Anderson, T.W., and C. Hsiao. 1981. "Estimation of Some Dynamic Models with Error Components." *Journal of American Statistical Association* 76(375):598–606.

Awases, M., A. Gbary, J. Nyoni, and R. Chatora. 2003. *Migration of Health Professionals in Six Countries: A Synthesis Report.* World Health Organization, Regional Office for Africa, Brazzaville, Republic of Congo.

Barro, R.J., and J.W. Lee. 2001. "International Data on Educational Attainment: Updates and Implications." *Oxford Economic Papers* 53(3):541–63.

Bell, C., S. Devarajan, and H. Gersbach. 2006. "The Long-Run Economic Costs of AIDS: A Model with Applications to South Africa." *World Bank Economic Review* 20(1):55–89.

Belperio, P., and D. Rhew. 2004. "Prevalence and Outcomes of Anemia in Individuals with Human Immunodeficiency Virus: A Systematic Review of the Literature." *American Journal of Medicine* 116(Supplement 1):27–43.

Bhargava, A. 1991. "Identification and Panel Data Models with Endogenous Regressors." *Review of Economic Studies* 58(1):129–40.

———. 1997. "Nutritional Status and the Allocation of Time in Rwandese Households." *Journal of Econometrics* 77(1):277–95.

———. 2001. "Stochastic Specification and the International GDP Series." *Econometrics Journal* 4(2):274–87.

———. 2005a. "AIDS Epidemic and Health Care Infrastructure Inadequacies in Africa: A Socioeconomic Perspective." *Journal of AIDS* 40(2):241–42.

———. 2005b. "AIDS Epidemic and the Psychological Well-Being and School Participation of Ethiopian Orphans." *Psychology, Health and Medicine* 10(3):263–75.

Bhargava, A., and J.D. Sargan. 1983. "Estimating Dynamic Random Effects Models from Panel Data Covering Short-Time Periods." *Econometrica* 51(6):1635–60.

Bhargava, A., and J. Yu. 1997. "A Longitudinal Analysis of Infant and Child Mortality Rates in Developing Countries." *Indian Economic Review* 32(2):141–53.

Bhargava, A., D. Jamison, L. Lau, and C. Murray. 2001. "Modeling the Effects of Health on Economic Growth." *Journal of Health Economics* 20(3):423–40.

Caral, M., and K. Holmes. 2001. "The Multicentre Study of Factors Determining the Different Prevalences of HIV in Sub-Saharan Africa." *AIDS* 15(Supplement 4):S1–S132.

Carrington, W., and E. Detragiache. 1999. "How Extensive Is the Brain Drain? *Finance and Development* 36(2):46–49.

Cohen, D., and M. Soto. 2007. "Growth and Human Capital: Good Data, Good Results." *Journal of Economic Growth* 12(1):51–76.

Docquier, F., and A. Bhargava. 2007. "The Medical Brain Drain: A New Panel Data Set on Physicians' Emigration Rates (1991–2004)." (http://www.ires.ucl.ac.be/csssp/home_pa_pers/docquier/filepdf/mbd1_description.pdf).

Docquier, F., and A. Marfouk. 2006. "Measuring International Migration by Educational Attainment in 1990–2000, Release 1.1." In C. Ozden, and M. Schiff, eds., *International Migration, Remittances, and the Brain Drain*. Washington, D.C.: World Bank.

Gutierrez, J., B. Johns, T. Adam, S. Bertozzi, T. Edejer, R. Greener, C. Hankins, and D. Evans. 2004. "Achieving the WHO/UNAIDS Antiretroviral Treatment 3 by 5 Goal: What Will It Cost?" *Lancet* 364(9428):63–64.

ILO (International Labour Organization). 2005. "Laborsta Internet: Occupational Wages and Hours of Work." Geneva. (http://laborsta.ilo.org).

Jaffar, S., A. Grant, J. Whitworth, P. Smith, and H. Whittle. 2004. "The Natural History of HIV-1 and HIV-2 Infections in Adults in Africa: A Literature Review." *Bulletin of World Health Organization* 82(66):462–69.

Neyman, J., and E. Scott. 1948. "Consistent Estimates Based on Partially Consistent Observations." *Econometrica* 16(1):1–32.

Numerical Algorithm Group. 1991. *Numerical Algorithm Group Mark 13*. Oxford, United Kingdom: University of Oxford.

Ozden, C., and M. Schiff. 2006. *International Migration, Remittances, and the Brain Drain*. Washington, D.C.: World Bank.

Physicians for Human Rights. 2004. *An Action Plan to Prevent Brain Drain: Building Equitable Health Systems in Africa*. Boston, Mass.: Physicians for Human Rights. (http://physiciansforhumanrights.org/library/report-2004-july.html).

Preston, S. 1976. *Mortality Patterns in National Populations*. New York: Academic Press.

Subbarao, K., and D. Coury. 2004. *Reaching Out to Africa's Orphans: A Framework for Publication*. Africa Region Human Development Series. Washington, D.C.: World Bank.

UNAIDS (Joint United Nations Programme on HIV/AIDS). 2006. HIV Data. Geneva. (www.unaids.org/en/hiv_data/default.asp).

UNESCO (United Nations Educational, Scientific, and Cultural Organization). 2004. Statistics for Downloading. Institute for Statistics, Montreal. (http://www.uis.unesco.org).

USAID, UNICEF, and UNAIDS (U.S. Agency for International Development, United Nations Children's Fund, and Joint United Nations Programme on HIV/AIDS). 2003. *Children on the Brink 2002*. New York: United Nations Children's Fund.

Vujicic, M., P. Zurn, K. Diallo, O. Adams, and M. Dal Poz. 2004. "The Role of Wages in the Migration of Health Care Professionals from Developing Countries." *Human Resources for Health* 2 (3) doi:10.1186/1478-4491. (http://www.human-resources-health.com/content/2/1/3).

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

———. 2006. EDSTATS Database. Washington, D.C. (http://devdata.worldbank.org/edstats/).

WHO (World Health Organization). 2006. "Core Health Indicators." Geneva. (http://www3.who.int/whosis/core/core_select.cfm).

Foreign Direct Investment, Access to Finance, and Innovation Activity in Chinese Enterprises

Sourafel Girma, Yundan Gong, and Holger Görg

A recent, comprehensive database is used to investigate the link between inward foreign direct investment (FDI) and innovation activity in China. The results of the analysis suggest that private and collectively owned firms with foreign capital participation and those with good access to domestic bank loans innovate more than other firms do. Among enterprises not owned by the state, inward FDI at the sectoral level is positively associated with domestic innovative activity only among firms that engage in their own research and development or that have good access to domestic finance. At the sector level the effect of inward FDI into technology transfer is distinguished from the effect on domestic credit opportunities. FDI affecting credit is of little significance for state-owned enterprises and is independent of their access to finance. In contrast, better access to credit is an important channel through which FDI affects the innovation of domestic private and collectively owned enterprises. JEL codes: O31, F23, G32

Since undergoing economic reforms beginning in 1979, China has emerged as a rapidly growing manufacturer and exporter. For economic development to continue apace, technological progress is crucial. In this regard, two features of the economy are notable. First, China has been an important location for research and development (R&D) and innovative activity by domestic enterprises since at least the mid-1990s (Jefferson and others 2006). Second, China

has become an important host country for inward foreign direct investment (FDI). Indeed, according to the Organisation for Economic Co-operation and Development (OECD 2004), in 2003 China overtook the United States as the world's largest recipient of FDI.[1]

This article examines whether there is a link between increased levels of inward FDI and innovation activity by Chinese domestic enterprises. It draws on a particularly rich and recent firm-level data set for 1999–2005 provided by the China National Bureau of Statistics (see China National Bureau of Statistics 2006, for a description) paying particular attention to the importance of domestic access to finance.

China is an interesting case to study because of the juxtaposition of an impressive record in attracting FDI with a highly inefficient and state-dominated domestic financial system. The Chinese financial system is widely regarded as inefficient and skewed toward providing financial resources for (also largely inefficient) state-owned enterprises (Huang 2003). It is therefore of considerable policy interest to investigate whether this skewed allocation of financing is related to the ability of domestic firms to benefit (or not) from the increased influx of FDI into the economy.

Increased FDI could affect product innovation by domestic enterprises through many channels. First, a domestic firm receiving an injection of foreign capital faces looser financial constraints—constraints that could hinder innovation (Harrison and McMillan 2003). Second, FDI may bring an inflow of technology. Multinational enterprises are assumed to have better technology than domestic firms (Markusen 2002); foreign capital inflow through acquisition, joint venture, or some other form of capital transfer may thus lead to the installation of foreign technology in the domestic firm. The superior knowledge injected into the economy through FDI may leak to domestic firms (through, for example, worker mobility and imitation), in the same way that productivity improvements spill over (Görg and Greenaway 2004). These firms may then be able to engage in more innovation activity.

A rich panel data set of 239,085 domestic enterprises in Chinese manufacturing industries for 1999–2005 is used to investigate these issues empirically. The analysis contributes to the literature on innovation and technology adoption in developing and transition economies. This literature, reviewed by Keller (2004) and Pack (2006), has amassed a large body of evidence on a number of countries. Econometric work, however, particularly on China, remains limited, despite the immense interest by policymakers and academics.

Using firm-level data for 1997–99 on some 20,000 enterprises, Jefferson and others (2006) model a knowledge production function to estimate

---

1. See Amiti and Javorcik (2008), Wei (2003), and Lemoine (2000) for analyses of the increased inflow of FDI and its determinants.

the determinants of innovation activity in Chinese enterprises.[2] They find that firm size and own R&D expenditure are important determinants of innovation.

This article expands on their analysis in a number of ways. First, it investigates the impact of inward FDI and the role of firm characteristics on innovation activity at the firm and industry level. This issue, which is not covered by Jefferson and others (2006), is of considerable policy interest. Second, the article addresses the role of access to finance for innovation and the link between such access and a firm's ability to benefit from positive spillovers from inward FDI. This issue has not, to the authors' knowledge, been investigated elsewhere. Third, unlike Jefferson and others (2006), who include dummy variables for different types of ownership, this article investigates whether the determinants of innovation activity, particularly the effect of inward FDI, differ across firms by type of ownership (privately owned, collectively owned, or state owned). Fourth, the data set used here covers more firms and a more recent time period than that used by Jefferson and others (2006).

The results show, not unexpectedly, that firms with foreign capital participation and good access to finance innovate more than others do and that inward FDI at the sectoral level is positively associated with domestic innovative activity only if firms engage in their own R&D activities or have good access to domestic finance. Following up on the suggestion of a possible adverse effect of domestic credit constraints on firms' ability to benefit from inward FDI, further analysis shows that finance constraints adversely affects private and collectively owned firms but not state-owned firms, which enjoy preferential access to domestic financial resources. Again, when the effects of sector-level inward FDI into technology transfer and FDI affecting domestic credit opportunities are distinguished, FDI affecting domestic credit opportunities is found to be of little significance for state enterprises and is independent of their access to finance. By contrast, it is an important channel through which FDI affects domestic private and collectively owned enterprises.

The article is structured as follows. Section I describes the empirical approach, and section II introduces the data set and provides some summary statistics. Section III discusses the econometric results. The final section provides some concluding comments.

---

2. Other, more loosely related articles should be noted. Hu, Jefferson, and Qian (2005) estimate the determinants of firm-level productivity using a data set similar to that used by Jefferson and others (2006). They focus on the role of own-firm R&D as well as technology purchased from foreign or domestic sources on firm-level productivity. Girma and Gong (forthcoming) use a Chinese data set to investigate the link between inward FDI spillovers and the productivity of state-owned enterprises. Berthelemy and Demurger (2000) use regional data to investigate the effect of inward FDI on regional growth in China.

I. Empirical Approach

In the empirical model a domestic firm (indexed by i) either innovates at time t ($S_{it} > 0$) or it does not ($S_{it} = 0$). A Tobit model is formulated in terms of a latent variable model to determine the relation between FDI and the rate of product innovation:

$$
(1) \quad
\begin{aligned}
S_{it}^* ={}& \alpha_1 FC_{it-1} + \alpha_2 FDI_{jt-1} + \alpha_3 FIN_{it-1} + \alpha_4 \mathbf{X}_{it-1} \\
& + \alpha_5 FIN_{it-1}{}^* \mathbf{FDI}_{jt-1} + \alpha_6 RD_{it-1}{}^* FDI_{jt-1} + D_r + D_j + D_t + \varepsilon_{it}
\end{aligned}
$$

$$
\begin{aligned}
S_{it} &= 0 \quad && if\ S_{it}^* \leq 0 \\
S_{it} &= S_{it}^* \quad && if\ S_{it}^* > 0
\end{aligned}
$$

where the dependent variable S is defined as the share of innovation output products involving the use of new process innovation or novel technology in total output.[3] This variable, which measures the output of the innovation process, is a more suitable measure than R&D, which is an input into the innovation process (see Criscuolo, Haskel, and Slaughter 2005). The D variables in equation (1) are full sets of regional (r), industry (j), and time (t) dummy variables.

$\mathbf{X}$ is a vector of firm-level determinants of innovation. It includes R&D intensity, the ratio of employee training expenditure to the total wage bill, export intensity, subsidies, age, and the firm's market share within its three-digit industry. The choice of these firm-level covariates is guided by theoretical considerations as well as evidence from the empirical literature. R&D is an important input into the innovation process and thus is included in the model. Human capital is also an important determinant of innovation. One proxy for human capital is the amount of training provided by a firm, which is included in the empirical analysis. Criscuolo, Haskel, and Slaughter (2005) provide evidence that firms that are active on export markets are more innovative than others. The model here captures this notion by controlling for firms' export intensities. Because subsidies can help firms engage more in innovation (see Görg and Strobl 2007), a measure of the level of production-related grants is also included in the model. As Jefferson and others (2006) argue, the age of a firm may also be important in explaining innovation activity (as a proxy for a firm's experience) and hence the possibility for learning effects. Their approach is adopted here by including firm age in the equation. Aghion and others (2005) discuss the role of competition for innovation; Aitken and Harrison (1999) show that multinationals may affect the competitive landscape in the domestic economy, leading to an increase in competition for domestic firms.

---

3. Definitions of all variables, plus summary statistics, are provided in table 1, discussed in the next section.

To take account of these findings, the model includes a firm's market share as an indicator of its competitive position.

FIN is a measure of a firm's access to finance (measured by its ability to obtain loans from domestic banks). Financial constraints are a serious impediment to innovation activity (Hall 2002). This effect may be particularly pronounced in China, where the financial sector is highly regulated and inefficient, and lending is skewed toward inefficient state-owned enterprises (Huang 2003).

FC is a measure of foreign capital participation in firm *i*. It captures the central concern of this article—the impact of FDI on innovative activity in Chinese domestic firms. *FC* is included to allow for the fact that firms with some share of foreign capital may be more innovative than other firms, for the reasons discussed above.

FDI is a vector of industry-region-specific FDI indices. It captures the potential spillover or crowding-out effects of FDI at the industry level. The effect of **FDI** is allowed to vary based on a firm's R&D activity and access to finance by including two interaction terms in the empirical estimation of equation (1), namely **FDI** and R&D intensity and **FDI** and *FIN*. The interaction of **FDI** and R&D intensity captures the notion that firms with higher absorptive capacity are better able to benefit from the technology transferred by incoming FDI.[4] The interaction of **FDI** and *FIN* allows firms with better access to finance to benefit more from inward FDI; because they are less financially constrained, they may be better able to implement the new technology and less affected by reductions in the availability of domestic finance caused by demand for loans by foreign firms.

All covariates in the empirical model are lagged by one period to mitigate potential endogeneity concerns. Nevertheless, some firm-level variables in the specification may be endogenous. One is R&D intensity, which is a major input into the product innovation process. The choice of this input is likely to be correlated with factors that determine the firm's decision to innovate. Similar arguments can be made regarding the potential endogeneity of the other firm-level variables. To deal with this possible problem, all lagged firm-level variables except age are considered potentially endogenous. The instrumental variables technique for Tobit models developed by Smith and Blundell (1986) is used to estimate this model.[5]

---

4. See Girma (2005) for a discussion of the importance of absorptive capacity and an empirical illustration using firm-level data for the United Kingdom.

5. The estimation of Tobit models with endogenous regressors involves two steps. The first is to generate residual terms from linear regressions of each endogenous variable on the instrumental variables and all other exogenous regressors. The second is to estimate a standard Tobit model by including the residual terms from the first step in the list of covariates. The standard errors are bootstrapped to take account of the fact that residual terms are generated regressors. The residual terms are correction terms for the endogeneity problem; jointly statistically significant coefficients can be taken as evidence in favor of the hypothesis that instrumented variables are indeed endogenous. A one-step variant of this estimator involving stronger distributional assumptions is also available (Newey 1987). However, it fails to attain convergence in the data used here. This type of convergence problem is frequently encountered when there is more than one endogenous regressor.

Twice-lagged values of the potentially endogenous variables are used as instruments. The assumption is that conditional on the regressors, these variables are asymptotically uncorrelated with the error term of the model. Ultimately, of course, this is an empirical issue, tested using the Sargan–Hansen test for the validity of instrumental variables.

Additional instruments are also used. They include the share of state-owned enterprises in a region or industry, the share of loss-making state-owned enterprises in a region or industry, the level of regional financial development (bank loans to the private sector as a share of total loans), and whether the firm is politically affiliated with local, regional, or central governments. These instruments are designed to account for the endogeneity of sector-level FDI and access to finance. The share of the state sector, for example, is a proxy for state dominance in the region or industry; to the extent that access to finance is different for state-dominated sectors and regions, this is a reasonable instrument for firm-level access to finance. Similar arguments can be made for the share of loss-making state-owned enterprises and the level of regional financial development.

A large number of enterprises in China are affiliated with some level of government administration. The function of the relevant government body is to offer credit guarantees and political protection to the affiliated firms. This political affiliation variable is strongly related to firms' access to finance, because China's financial system remains dominated by the four large state banks. Different levels of political affiliation are used as instruments to reflect the realistic assumption that the main effect of political affiliation on innovation comes through its effects on finance. Ultimately, however, the relevance of the instruments is an empirical issue that is tested for in the estimation below.

II. DESCRIPTION OF THE DATABASE AND CONSTRUCTION OF VARIABLES

The econometric analysis draws on confidential micro data that underlie the *Annual Reports of Industrial Enterprise Statistics*, compiled by the China National Bureau of Statistics. The reports cover all firms with annual turnover of more than 5 million yuan (about $600,000). The firms in the data set account for an estimated 85–90 percent of total output in most industries.

The data set includes information on firm ownership structure, industry affiliation, geographic location, establishment year, employment, gross output, product innovation, R&D, value added, net fixed assets, exports, and employee training expenditures.[6] The whole sample (1.3 million observations from about 446,000 firms) is used to construct the variables of interest (such as the share of foreign firms in an industry or region or firms' market share).

6. Nominal values are deflated using industry-specific ex-factory price indices obtained from the *China Statistical Yearbook 2006* (China National Bureau of Statistics 2007).

The econometric work is confined to domestic-owned enterprises, the focus of this article.

The China National Bureau of Statistics assigns a categorical variable to each firm in the database indicating its ownership status. It is also possible to construct a continuous measure of ownership composition from the database by looking at the fraction of paid-in capital contributed by the state and by private (domestic and foreign) investors. This measure of ownership is used here. Firms are defined as state-owned, collectively owned, or privately owned based on majority ownership of the firm. The information necessary for the econometric estimation is available for 239,085 domestic firms (630,900 total observations).

The data set provides information on the extent of foreign capital participation at the level of the firm. This makes it possible to calculate the share of foreign ownership in the domestic enterprise and identify the direct effects of FDI on domestic firms' innovative activity. A different method is used to estimate the indirect (spillover) effect of FDI at the industry level. For each of the 171 three-digit industries and 31 provinces, the proportion of output accounted for by companies with foreign ownership in the industry and region is calculated.[7] Alternative measures of industry and region FDI are the proportion of new products accounted for by multinational companies (labeled FDI innovation), and the share of domestic bank loans extended to foreign multinationals (FDI loan).

The data reveal no substantial relation between firm ownership on the one hand and innovation activity or the level of R&D on the other (table 1). As expected, on average state-owned enterprises receive higher shares of bank loans and larger subsidies from the government. They are less export intensive and receive more modest inflows of foreign capital than privately or collectively owned firms.

The pattern of product innovation by state-owned enterprises across industries at the two-digit reveals three noteworthy points (table 2). First, the proportion of innovating firms rose over time in most sectors. In contrast, the share of new product sales in total sales, while generally significant, declined slightly in most sectors. Second, labor-intensive sectors (such as food manufacturing and paper products) have the lowest proportion of innovators. In contrast, export-competing labor-intensive sectors (such as textiles) exhibit a

7. Officially, foreign-owned multinationals are defined as enterprises with at least a 25 percent share of foreign capital. Domestically owned enterprises that have foreign capital participation of less than 25 percent are not considered in this definition. The richness of the data set is exploited by weighing the output of firms with foreign capital by the extent of their foreign participation, measured by the share of foreign capital at the firm level. Under this definition of sectoral FDI, firms classified as domestic but that have some foreign capital also contribute (proportionally) to the aggregate output of the foreign sector. The recent literature on productivity spillovers from FDI notes that domestic firms may benefit not only from horizontal but also from vertical spillovers through customer–supplier linkages (see Javorcik 2004). Vertical measures (backward and forward spillovers) were calculated by the authors but found not to be consistently statistically significant. They are therefore not included in the analysis that follows.

TABLE 1. Variable Definitions and Summary Statistics

Variable	Definition	State-owned enterprise		Private enterprise		Collective enterprise	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Product innovation	Share of output involving new process or product innovation	0.041	0.150	0.034	0.151	0.021	0.116
R&D	Restricted sample of firms with nonzero product innovation	0.319	0.295	0.390	0.350	0.369	0.331
	R&D expenditure divided by sales	0.002	0.021	0.002	0.013	0.001	0.007
	Restricted sample of firms with nonzero R&D expenditure	0.013	0.046	0.012	0.337	0.008	0.018
Labor training	Employee training expenditure per employee	0.007	0.027	0.008	0.035	0.008	0.037
	Restricted sample of firms with nonzero labor training outlay	0.015	0.037	0.021	0.053	0.037	0.053
Export intensity	Share of exports in total sales	0.043	0.164	0.127	0.298	0.110	0.281
	Restricted sample of exporters	0.307	0.328	0.580	0.380	0.594	0.374
Market shares	Firm's share of sales in total three-digit industry region sales	0.044	0.134	0.022	0.082	0.021	0.073
Domestic finance	Domestic bank loans normalized by total assets	1.806	2.519	0.825	1.874	0.876	1.886
Subsidy	Log of production subsidy from local and central governments	0.983	2.335	0.642	1.843	0.796	2.052
Age	Log of years since establishment	3.147	0.914	1.848	0.935	2.569	0.806
Foreign capital	Share of foreign multinationals capital in firm's total capital	0.002	0.033	0.004	0.050	0.006	0.060
FDI	Share of foreign multinationals' sales in three-digit industry-region total sales	0.143	0.181	0.194	0.191	0.175	0.182
FDI innovation	Share of multinationals' innovative output in *three*-digit industry-region total innovation	0.097	0.196	0.140	0.212	0.127	0.206
FDI loan	Share of multinationals' domestic bank loans over total domestic bank loans	0.099	0.161	0.168	0.198	0.148	0.184
Number of firms		239,085		148,694		55,842	
Number of observations		630,900		316,461		189,082	

Source: Authors' analysis based on data from China National Bureau of Statistics.

TABLE 2. Sectoral and Temporal Pattern of Product Innovation by
State-Owned Enterprises (percent)

Two-digit industry classification	Share of innovators		New product sales as share of total sales	
	1999	2005	1999	2005
13: Food processing	2.0	10.1	32.3	16.6
14: Food production	4.3	11.6	29.2	23.9
15: Beverage industry	6.0	12.0	27.2	25.1
16: Tobacco processing	12.3	21.1	14.9	15.2
17: Textile Industry	17.3	17.2	30.7	29.6
18: Garments and other fiber products	3.5	6.5	45.0	45.3
19: Leather, furs, down and related products	4.1	8.1	49.4	39.7
20: Timber processing	2.8	6.8	46.2	23.0
21: Furniture manufacturing	4.2	10.0	36.0	21.4
22: Papermaking and paper products	4.0	7.2	37.1	19.0
23: Printing and record medium reproduction	1.8	5.9	37.5	35.0
24: Cultural, educational, and sports goods	9.4	9.2	33.5	38.9
25: Petroleum refining and coking	5.0	6.4	28.9	20.9
26: Raw chemical materials and chemical products	9.2	10.7	31.3	33.2
27: Medical and pharmaceutical products	24	25.2	35.8	37.2
28: Chemical fiber	14.0	10.4	26.7	39.4
29: Rubber products	12	9.8	32.0	30.5
30: Plastic products	9.1	10.2	38.2	33.9
31: Nonmetal mineral products	3.7	10.7	38.1	23.0
32: Smelting and pressing of ferrous metals	5.8	6.9	29.6	24.8
33: Smelting and pressing of nonferrous metals	6.0	9.7	32.9	335
34: Metal products	6.1	7.9	33.4	31.1
35: Ordinary machinery	14.2	13.2	29.5	320
36: Special purpose equipment	17.8	17.2	34.8	37.3
37: Transport equipment	14.1	15.5	35.5	34.7
39: Other electronic equipment	14.8	14.0	36.1	41.8
40: Electric equipment and machinery	26.8	23.2	47.6	53.3
41: Electronic and telecommunications	25.7	25.7	35.3	46.0
42: Instruments and meters	5.7	7.0	39.2	33.0

Source: Authors' analysis based on data from China National Bureau of Statistics.

relatively large number of innovators. Third, the intensity of product inno-
vation is remarkably similar across labor-intensive, capital-intensive, and
technology-intensive sectors.

III. Discussion of the Results

The benchmark Tobit model controls for firm heterogeneity by allowing for
firm random effects (table 3, column 1). The model also includes two
additional dummy variables for private and collectively owned firms.

   The estimation shows that R&D intensity exerts a positive and significant
influence on the rate of product innovation. This is as expected, given that

TABLE 3. Innovation Spillovers from FDI and Access to Finance: Results from Alternative Estimators

Variable	(1) Random effects Tobit model	(2) Tobit instrumental variables model	(3) Linear generalized method of moment model
R&D	2.312*** (30.4)	4.323*** (19.0)	2.118*** (10.8)
Labor training	0.481*** (11.5)	0.862*** (7.22)	0.700*** (3.65)
Export intensity	0.252*** (39.3)	0.238*** (24.4)	0.217*** (18.5)
Market share	0.576*** (41.6)	0.600*** (29.7)	0.901*** (21.0)
Finance	0.0413*** (46.9)	0.0696*** (35.6)	0.0684*** (23.8)
Subsidy	0.0271*** (38.8)	0.0313*** (24.7)	0.0399*** (18.3)
Age	0.0505*** (27.4)	0.0487*** (19.2)	0.00197*** (6.69)
Foreign capital	0.135*** (6.85)	0.217*** (5.54)	0.168*** (3.65)
FDI	−0.643*** (−45.8)	−0.797*** (−35.3)	−0.500*** (−23.5)
FDI × R&D	1.428*** (88.1)	1.881*** (54.2)	1.241*** (26.1)
FDI × finance	0.153*** (8.22)	0.0922** (2.21)	−0.00929 (−1.51)
Private firms	0.0463*** (9.39)	0.0634*** (10.8)	0.0389*** (12.4)
Collective firms	−0.110*** (−21.9)	−0.0686*** (−11.5)	−0.0800*** (−7.15)
Number of observations	630,900	390,352	390,352
p-value from Hansen's test of overidentification			0.179

***Significant at the 1 percent level; **significant at the 5 percent level; *significant at the 10 percent level.

Note: All specifications include time, regional, and industry fixed effects. Numbers in parentheses are t-statistics. All regressors are lagged by one period in all regressions. The use of twice-lagged variables as instruments in the Tobit instrumental variables and linear generalized method of moments (GMM) models sharply reduced the number of observations in the estimations. State-owned firms form the base group in all regressions.

Source: Authors analysis based on data from China National Bureau of Statistics.

R&D intensity is a major input in the product innovation process. Firms that invest in employee training have higher innovation intensity, suggesting possible complementarity between human capital investment and innovation, as suggested by Redding (1996). Production innovation and exporting are positively correlated, and firms that enjoy higher market shares in their industry are more likely to engage in product innovation activity. Firms' receipt of subsidies is also positively related to innovation. These findings are in line with the literature. The results also suggest that older firms are more likely to engage in product innovation than their younger counterparts, possibly indicating the importance of experience in the innovation process.[8]

The results indicate that firms with some foreign capital participation are more likely to engage in product innovation. This effect may reflect the influx

8. In an analysis of a smaller sample of Chinese firms, Jefferson and others (2006) find no statistically significant relation between age and innovation. The sample used here is much larger and the estimation controls for a much larger number of covariates than the earlier study did.

of new technology or the reduction in financial constraints associated with the capital injection (the data do not allow the two hypotheses to be tested separately).[9] Access to finance is also positively associated with innovation. The magnitude of the effect of firm-level foreign capital is economically significant. A doubling of the foreign ownership share, for example, is associated with a 13.5 percentage point change in the share of new products in total output, all else equal. The economic significance of access to domestic finance is noteworthy: a one standard deviation (2.5 from table 1) increase in this variable leads to a 10 percentage point increase in innovation intensity.

Firms benefit from inward FDI only if they are actively engaged in R&D or have good access to domestic finance in the form of bank loans. If the point estimates are taken at face value, firms with R&D intensities of at least 0.45 (and no domestic bank loans) and firms with a ratio of bank loans to assets of at least 4.2 (and no R&D) benefit from FDI. These two threshold values are well beyond the mean values in the sample (reported in table 1).

The results in column 1 treat all variables as exogenous. This may not be too heroic an assumption, considering that all covariates are lagged one period, dummy variables for private and collectively owned firms are added, and unobserved firm-level heterogeneity is allowed for by including firm random effects. There may still be a lingering concern about endogeneity, however. This problem is addressed more formally by estimating the model using the endogenous Tobit model (column 2 in table 3). Of course, the reliability of the endogenous Tobit hinges on the validity of the instruments used. To the authors' knowledge, there are no formal tests of the validity of instruments within the context of these endogenous Tobit specifications. The appropriateness of the instruments is gauged by estimating a linear instrumental variables model (using the same set of instruments as in the endogenous Tobit) and obtaining a Sargan test statistic of the implied overidentifying restrictions. The test result, reported in column 3, suggests the validity of the instrumental variables. Reassuringly, the results of the estimations in both columns 2 and 3 are very much in line with the baseline equation in column 1, in terms of both the statistical significance and the magnitude of the coefficients. In the rest of the discussion the focus is therefore on estimations using the Tobit estimator allowing for firm random effects (column 1).

This estimation allows for firm-level heterogeneity, but it constrains the coefficients of the independent variables to be the same for all types of firms. This may be an unrealistic assumption given the large performance differences between state-owned and other types of enterprises in China. The ownership dimension of the data set is exploited to address the sample, which is broken into state-owned, private, and collectively owned enterprises (table 4).

---

9. Another reason why foreign-owned firms may be more innovative is that they tend to employ more highly skilled workers. The estimation partly controls for this possibility by controlling for the quality of the workforce using a variable on labor training.

TABLE 4. Estimates of FDI Spillovers and Access to Finance from Panel Tobit Models with Firm-Specific Heterogeneity

Variable	(1) State-owned firms	(2) Private firms	(3) Collectively owned firms	(4) Loss-making firms	(5) Profit-making firms
R&D	1.100*** (12.3)	2.857*** (24.3)	6.591*** (19.5)	1.034*** (8.27)	3.213*** (32.8)
Labor training	0.552*** (6.40)	0.464*** (7.84)	0.415*** (4.56)	0.345*** (2.99)	0.449*** (9.90)
Export intensity	0.363*** (25.5)	0.232*** (26.2)	0.274*** (19.2)	0.267*** (15.6)	0.252*** (36.4)
Market share	0.487*** (29.7)	0.535*** (22.8)	0.497*** (11.5)	0.435*** (14.8)	0.625*** (39.7)
Finance	0.0085*** (31.7)	0.0384*** (26.8)	0.0406*** (18.9)	0.0372 (0.66)	0.0470*** (46.7)
Subsidy	0.0235*** (24.3)	0.0335*** (29.8)	0.0163*** (9.69)	0.0243*** (14.7)	0.0268*** (34.9)
Age	0.0478*** (15.0)	0.0465*** (17.9)	0.0806*** (17.1)	0.0218*** (5.93)	0.0409*** (21.6)
Foreign capital	0.333*** (7.73)	0.0978*** (3.35)	0.135*** (3.45)	0.0816* (1.68)	0.123*** (5.68)
FDI	−0.502*** (−19.4)	−0.631*** (−32.0)	−0.838*** (−25.1)	−0.358*** (−10.9)	−0.684*** (−43.9)
FDI × R&D	1.106*** (42.8)	1.474*** (62.6)	1.560*** (38.5)	1.178*** (30.7)	1.478*** (82.2)
FDI × finance	0.0975 (1.42)	0.0575* (1.93)	0.250*** (5.50)	0.128 (1.21)	0.112*** (5.29)
Number of observations	125,357	316,461	189,082	117,001	513,899

***Significant at the 1 percent level; **significant at the 5 percent level; *significant at the 10 percent level.

Note: All specifications include time, regional, and industry fixed effects. Numbers in parentheses are t-statistics. All regressors are lagged by one period in all regressions.

Source: Authors' analysis based on data from China National Bureau of Statistics.

TABLE 5. Access to Finance by FDI Firms and Innovation Spillovers to Domestic Firms

Variable	(1) State-owned firms	(2) Private firms	(3) Collectively owned firms	(4) Loss-making firms	(5) Profit-making firms
R&D	0.823*** (7.99)	2.914*** (36.1)	5.197*** (24.1)	1.134*** (13.2)	2.591*** (42.0)
Labor training	0.250** (2.32)	0.338*** (8.32)	0.300*** (5.20)	0.294*** (3.71)	0.309*** (10.6)
Export intensity	0.247*** (13.3)	0.105*** (17.8)	0.133*** (15.5)	0.144*** (13.0)	0.119*** (27.0)
Market share	0.350*** (16.0)	0.460*** (27.4)	0.282*** (10.4)	0.347*** (17.1)	0.431*** (40.1)
Finance	0.0011*** (19.1)	0.0289*** (27.8)	0.0281*** (19.6)	0.028 (1.21)	0.0311*** (44.9)
Subsidy	0.0172*** (14.0)	0.0234*** (29.6)	0.0117*** (10.8)	0.0159*** (14.0)	0.0181*** (35.5)
Age	0.0309*** (7.51)	0.0404*** (21.6)	0.0590*** (19.2)	0.0383*** (13.6)	0.0437*** (30.8)
Foreign capital	0.205*** (3.23)	0.0542*** (2.64)	0.0737*** (2.92)	0.0698** (2.09)	0.0757*** (5.26)
FDI innovation	−0.0345** (−2.35)	−0.110*** (−9.01)	−0.0452*** (−2.72)	−0.091* (−1.82)	−0.034*** (−10.4)
FDI innovation × R&D	0.384*** (8.74)	0.518*** (24.3)	0.314*** (9.82)	0.409*** (11.7)	0.443*** (30.1)
FDI innovation × finance	−0.0191 (−0.91)	0.0205** (2.22)	0.0490** (2.75)	−0.0067 (−1.19)	0.0677** (2.46)
FDI loan	−0.190 (−1.49)	−0.493*** (−35.2)	−0.475*** (−22.3)	−0.277*** (−12.0)	−0.433*** (−42.1)
FDI loan × R&D	0.653*** (11.2)	0.828*** (36.3)	0.864*** (24.5)	0.735*** (17.8)	0.764*** (47.5)
FDI loan × finance	0.0026 (0.90)	0.0115** (2.03)	0.0378** (2.47)	−0.00185 (−0.26)	0.059** (2.57)
Number of observations	125,357	316,461	189,082	117,001	513,899

***Significant at the 1 percent level; **significant at the 5 percent level; *significant at the 10 percent level.

Note: All specifications include time, regional, and industry fixed effects. Numbers in parentheses are *t*-statistics. All regressors are lagged by one period in all regressions.

Source: Authors' analysis based on data from China National Bureau of Statistics.

Several striking differences are apparent across the three ownership types. First, the relation between access to finance and innovation is strongest among private and collectively owned firms, which receive less favorable treatment from China's financial system than state-owned enterprises do. Second, the coefficient on foreign capital is largest for state enterprises, suggesting that injections of foreign capital are associated with the highest positive impact on innovation for this type of firm. This may reflect the fact these state-owned firms are more inefficient than other firms and therefore offer the greatest opportunities for improvement as a result of the influx of foreign capital.[10] Third, and perhaps most striking, the interaction term of FDI and access to finance is positive for private and collectively owned firms but statistically insignificant for state-owned enterprises. Access to domestic finance plays no role in generating spillovers to state-owned enterprises, which are largely inefficient but enjoy preferential access to domestic financial resources.

Profit-making firms can be distinguished from loss-making firms, most of which are owned by the state (columns 4 and 5 in table 4). The results are in line with expectations: access to finance has no effect on innovation in loss-making enterprises, and also does not matter for indirect effects from sector-level inward FDI.[11]

Sector-level FDI can affect domestic innovation by transferring technology to or creating credit opportunities for domestic firms. The next step of the analysis tries to distinguish these two channels by calculating two different FDI measures. The first measure is aggregate innovation by foreign multinationals, calculated as innovation output by foreign multinationals in a sector or region divided by total innovation output. The second measure is aggregate borrowing by foreign multinationals, calculated as the share of domestic bank loans in total bank loans in the sector or region.

The results in columns 1–3 of table 5 show that the effects of the two variables are broadly similar to those for private and collectively owned firms. FDI has a positive effect only if the firm is active in R&D and has access to bank loans. State-owned enterprises that invest in their own R&D also benefit more from technology transfer by multinationals than those that do not, but the firms' financial position does not mitigate the effect of FDI technology. The effect of FDI on credit opportunities has no statistically significant relation with state enterprises' ability to innovate. This result suggests that preferential access to domestic financial resources means that finance is not a constraint for state enterprises.

In alternative estimations in columns 4 and 5, the data are broken into loss-making and profit-making enterprises. Results for loss-makers closely resemble

---

10. This result is in line with the work by Bartel and Harrison (2005) that shows state-owned enterprises in Indonesia benefit greatly from foreign ownership in the enterprise.

11. As a robustness check, table 4 was reestimated using the endogenous Tobit estimator. The results are very similar and hence not reported here.

those for state-owned enterprises, suggesting that access to finance has little effect on innovation for such firms.[12]

IV. Conclusions

The econometric analysis conducted here shows that access to finance is an important issue for firms' innovation activity and their ability to benefit from inward FDI. This is mainly the case for private and collectively owned firms, however. It is far less important for state-owned firms, which receive preferential treatment under the current domestic financial system.

Firms with foreign capital participation and those with good access to domestic bank loans—that is, firms with less binding financial constraints—innovate more than others. Inward FDI at the sectoral level is positively associated with domestic innovative activity only if firms engage in their own R&D activities (that is, have some absorptive capacity) or have good access to domestic finance. This finding points to the possible adverse effect of domestic credit constraints on firms' ability to benefit from inward FDI. Grouping firms by ownership type reveals that access to finance plays a role only among firms that are not state owned. Although state-owned enterprises are largely inefficient, they enjoy preferential access to domestic financial resources; access to finance thus provides no bottleneck for them.

Sector-level inward FDI has two effects. It transfers technology and may increase domestic credit opportunities. The effect on credit is of very little significance for state-owned enterprises and is independent of their access to finance. In contrast, it is an important channel through which FDI affects the innovation of domestic private and collectively owned enterprises.

References

Aghion, Philippe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120(2):701–28.

Aitken, Brian J., and Ann E. Harrison. 1999. "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela." *American Economic Review* 89(3):605–18.

Amiti, Mary, and Beata S. Javorcik. 2008. "Trade Costs and Location of Foreign Firms in China." *Journal of Development Economics* 85(1–2):129–49.

Bartel, Ann P., and Ann E. Harrison. 2005. "Ownership versus Environment: Disentangling the Sources of Public-Sector Inefficiency." *Review of Economics and Statistics* 87(1):135–47.

Berthelemy, Jean-Claude, and Sylvie Demurger. 2000. "Foreign Direct Investment and Economic Growth: Theory and Application to China." *Review of Development Economics* 4(2):140–55.

China National Bureau of Statistics. 2006. *Annual Reports of Industrial Enterprise Statistics*. Beijing.

12. One other robustness check was conduced, which included squared terms of foreign capital, FDI innovation, and FDI loans in the estimations. The results, which are not reported here, are robust to this slight modification of the estimating equation.

———. 2007. *China Statistical Yearbook 2006*. Beijing: China Statistics Press.

Criscuolo, Chiara, Jonathan E. Haskel, and Matthew J. Slaughter. 2005. *Global Engagement and the Innovation Activities of Firms*. NBER Working Paper 11479. Cambridge, Mass.: National Bureau of Economic Research.

Girma, Sourafel. 2005. "Absorptive Capacity and Productivity Spillovers from FDI: A Threshold Regression Analysis." *Oxford Bulletin of Economics and Statistics* 67(3):281–306.

Girma, Sourafel, and Yundan Gong. Forthcoming. "FDI, Linkages and the Efficiency of State-Owned Enterprises in China." *Journal of Development Studies*.

Görg, Holger, and David Greenaway. 2004. "Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment?" *World Bank Research Observer* 19(2):171–97.

Görg, Holger, and Eric Strobl. 2007. "The Effect of R&D Subsidies on Private R&D." *Economica* 74(294):215–34.

Hall, Bronwyn H. 2002. "The Financing of Research and Development." *Oxford Review of Economic Policy* 18(1):35–51.

Harrison, Ann E., and Margaret S. McMillan. 2003. "Does Direct Foreign Investment Affect Domestic Credit Constraints?" *Journal of International Economics* 61(1):73–100.

Hu, Albert G. Z., Gary H. Jefferson, and Jinchang Qian. 2005. "R&D and Technology Transfer: Firm-Level Evidence from Chinese Industry." *Review of Economics and Statistics* 87(4):780–86.

Huang, Yasheng. 2003. *Selling China: Foreign Direct Investment during the Reform Era*. New York: Cambridge University Press.

Javorcik, Beata S. 2004. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages." *American Economic Review* 94(3):605–27.

Jefferson, Gary H., Bai Huamao, Guan Xiaojing, and Yu Xiaoyun. 2006. "R&D Performance in Chinese Industry." *Economics of Innovation and New Technology* 15(4–5):345–36.

Keller, Wolfgang. 2004. "International Technology Diffusion." *Journal of Economic Literature* 42(3):752–82.

Lemoine, Françoise. 2000. "FDI and the Opening Up of China's Economy." CEPII Working Paper 2000-11. Centres d'Etudes Prospectives et d'Informations Internationales, Paris.

Markusen, James R. 2002. *Multinational Firms and the Theory of International Trade*. Cambridge, Mass.: MIT Press.

Newey, Whitney K. 1987. "Efficient Estimation of Limited Dependent Variables Models with Endogenous Explanatory Variables." *Journal of Econometrics* 36(3):231–50.

OECD (Organisation for Economic Co-operation and Development). 2004. *Trends and Recent Developments in Foreign Direct Investment*. Directorate for Financial, Fiscal and Enterprise Affairs, Paris.

Pack, Howard. 2006. "Econometric vs. Case Study Approaches to Technology Transfer." In Bernard M. Hoekman, and Beata S. Javorcik, eds., *Global Integration and Technology Transfer*. London, United Kingdom: Palgrave Macmillan and World Bank.

Redding, Stephen. 1996. "The Low-Skill, Low-Quality Trap: Strategic Complementarities between Human Capital and R&D." *Economic Journal* 106(March):458–70.

Smith, Richard J., and Richard W. Blundell. 1986. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." *Econometrica* 54(4):679–86.

Wei, Yingqi. 2003. "Foreign Direct Investment in China." Working Paper 2003/002. Lancaster University Management School, United Kingdom.