80525

THE WORLD BANK

Research Observer

Volume 27 • Number 2 • August 2012

OXFORD
UNIVERSITY PRESS

The World Bank Research Observer is intended for anyone who has a professional interest in development. *Observer* articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the *Observer* considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the Editor. Please direct all editorial correspondence to the Editor, *The World Bank Research Observer*, 1818 H Street, NW, Washington, DC 20433, USA.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

For more information, please visit the Web sites of the *Research Observer* at www.wbro.oxfordjournals.org, the World Bank at www.worldbank.org, and Oxford University Press at www.oxfordjournals.org.

The Odds of Achieving the MDGs

Delfin S. Go • José Alejandro Quijada*

Three questions are frequently raised about the attainment of the Millennium Development Goals (MDGs). Where do developing countries stand? What factors affect their rate of progress? Can lagging countries achieve these goals in the few years remaining until 2015? This paper examines these questions and takes a closer look at the variation in the rate of progress among developing countries. We argue that answers from the available data are surprisingly positive. In particular, three-quarters of developing countries are on target or close to being on target for all of the MDGs. Among the countries that are falling short, the average gap for the top half is about 10 percent. For those that are on target, or close to it, solid economic growth, policies, and institutions have been the key factors in their success. With improved policies and stronger growth, many countries that are close to being on target could achieve these targets by 2015 or soon after. JEL codes: F55, O19, O43

One puzzle about the Millennium Development Goals (MDGs) befuddles greatly. Why has the overall progress toward the MDGs been so varied when the economic performance of developing countries has been observed to be markedly better for the more than 15 years since the mid-1990s? Until the recent economic crisis, the external environment was favorable: trade was expanding, export prices were buoyant, and both foreign aid and debt relief were increasing. Moreover, for a remarkably broad range of developing countries, economic growth was accelerating because of better policies and institutions. This situation was encouraging because it was true not only for large, middle-income countries, such as China and India, but also for poor countries in sub-Saharan Africa.[1] Because of improved policies and institutions, the recent crisis was different for low-income countries, which did relatively well. There was no widespread failure in domestic policy, growth remained positive, and the poor were protected by increased

spending on social safety nets.[2] Therefore, the following question begs answers: Where did all of the economic progress go, and what did it buy for the MDGs?

The answers to the above questions lie beyond the global numbers themselves. Solving this puzzle provides some answers to three key questions that are frequently raised about the MDGs: (1) Where do developing countries stand? (2) What factors affect the rate of progress of developing countries? and (3) Will lagging countries achieve the goals in the few years remaining until 2015? This paper examines these questions and several related issues. In the process, we argue that answers from available information are surprisingly helpful and hopeful.

The global numbers tell a familiar story in two ways.[3] In terms of the remaining distance toward the 2015 targets (figure 1a), the latest information confirms that progress remains strong on the reduction of both extreme poverty and hunger, access to safe drinking water, and gender parity in primary and secondary education. In terms of the distance to the trajectory required to be on target (figure 1b), according to current trends (or historical growth rates), the developing world is on track to reach the global target of reducing extreme poverty and the proportion of people without safe drinking water by half by 2015.[4] Rapid growth in China, East Asia, and the Pacific Region has already cut extreme poverty by half. Developing countries will likely achieve the MDGs for gender parity in primary and secondary education as well as in access to safe drinking water, and they will be close to reducing hunger and to the primary education completion rate. However, by either yardstick, the distance to the goals or the distance to being on track, progress continues to lag in health-related development outcomes, such as reductions in child mortality and maternal mortality and access to sanitation. New data and methodologies indicate much more progress than previously thought in reducing maternal mortality, but this MDG continues to have the greatest lag (Hogan et al. 2010). Considering current trends, the world is likely to miss these three targets by 2015. Moreover, low-income countries, particularly fragile states and those in sub-Saharan Africa, lag behind because of a combination of low starting points and difficult circumstances (Easterly 2009, Clemens et al. 2007, World Bank and IMF 2010).

Behind these aggregate numbers, however, there is wide variation in performance across indicators, countries, and groups of countries that requires further analysis. Bourguignon et al. (2010), Leo and Barmeier (2010), and ODI (2010) showed that progress has been more heterogeneous than shown by the aggregate figures. Although the MDGs were conceived as global targets to spur development efforts and support poor countries, it is necessary to measure and describe progress at the country or other level to understand the reasons for both the advances and the remaining gaps.[5] As a prelude to the analysis in the paper and although there are variations and complications, a key point is the fundamental distinction between growth and development, which has a clear resonance in the

Figure 1. Current Global Distance to the MDGs
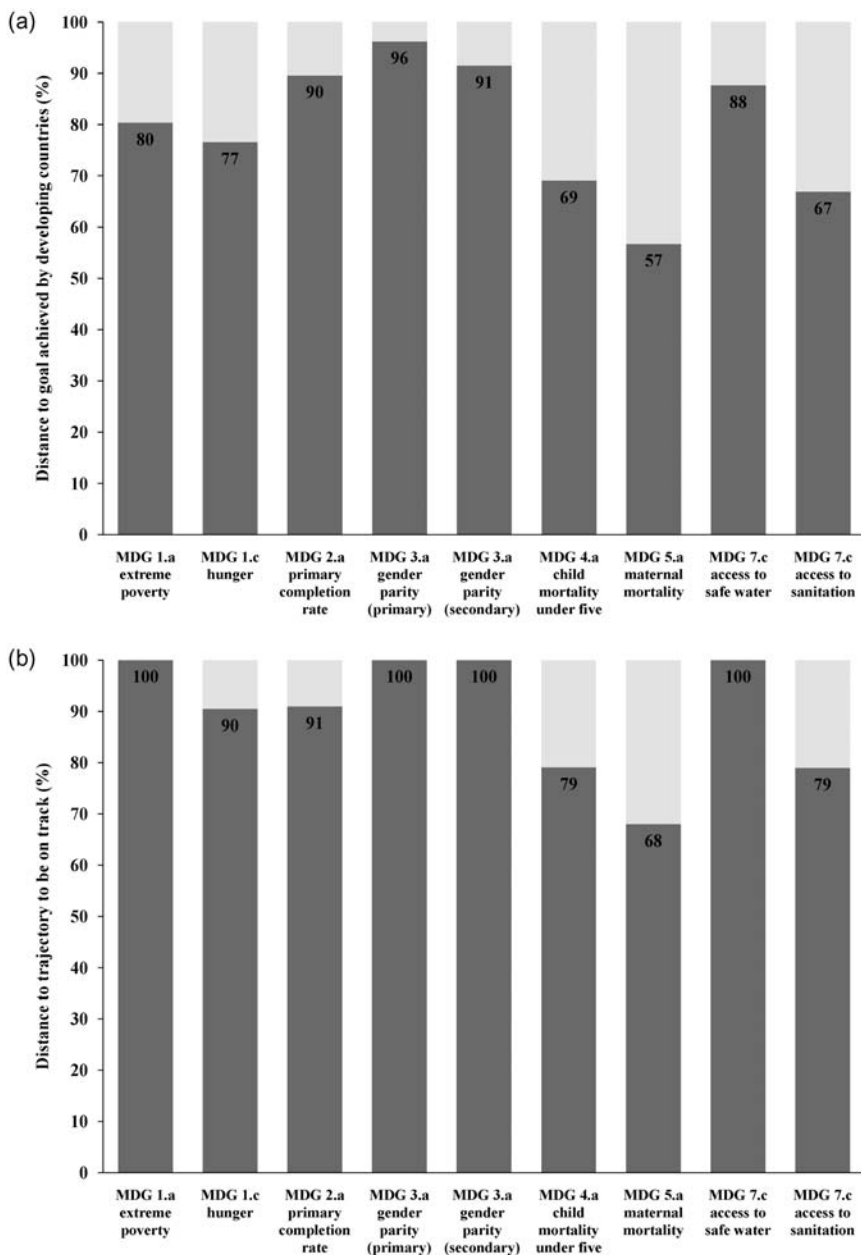
a. Distance of latest indicators to 2015 goals

b. Distance to the trajectory to be on track to achieve the goals by 2015

Note: Distance to goal achieved in this graph is a weighted average of the latest indicators, using population weights in 2009.

Source: Authors' calculations based on data from the World Development Indicators database.

main findings of the study. Improving developing outcomes will require not only increases in GDP per capita but also system-wide changes in policy and institutions to bring about more inclusive growth or broad-based development in order to improve the living conditions, opportunities, and quality of life of all individuals, groups, and nations in the world. Separating the aggregates provides further support for this point. Indeed, global and regional summaries typically amass data for countries of dissimilar development and types—fragile, low-income, and middle-income countries. For example, the Europe and Central Asia region covers middle-income countries, such as Albania and Bulgaria, and low-income countries, such as Tajikistan and Uzbekistan. Among the developing countries in Sub-Saharan Africa, some are middle-income countries (such as Mauritius and South Africa). Some lower-middle-income countries (such as Angola and the Democratic Republic of Congo) are resource rich, but their levels of development may be closer to those of low-income countries.

To illuminate these issues and untangle the aggregate numbers, we use the three basic questions raised above to examine individual country performance and to structure the paper. Section II investigates where individual countries stand and presents our MDG performance measurement and assessment. We introduce a simple but reasonable approach to measure and categorize MDG progress and to assess the likelihood of developing countries reaching the MDG goals. Our approach characterizes MDG progress by country performance in terms of countries that are on track to achieve the targets and by the distance or "closeness" of lagging countries to being on track to achieve the targets. Section III examines factors that affect the progress of countries. We examine the importance of different typologies in the variations in progress toward reaching the MDG targets by 2015. Examples of such factors are initial income and policy-institutional conditions, subsequent growth and policy-institutional achievement, the poorest of the developing countries versus the other countries, and the level of fragility (broadly following Collier and O'Connell 2006). Section IV attempts to determine whether lagging countries are likely to achieve their goals in the few years remaining before 2015. This question is not easy to answer, but we attempt to identify some answers within the limitations of the data. The final section summarizes our key findings and provides valuable insights for future research and policy changes.

Where Do Countries Stand with Respect to Attaining the MDGs?

The Definition of MDG Performance

The MDGs are typically defined in terms of the number or percentage of people (e.g., reducing the number of poor people by one-half or achieving 100 percent

access to primary education). Although data are collected on a country basis, the influence of each country on the global average depends on the size of its population. When large countries, such as China and India, are doing well, as in the MDG related to the reduction of poverty, their progress will be reflected in the global average, which will also hide the progress (or lack of progress) in smaller countries. To examine how poor countries are doing, data are presented in terms of progress in individual countries. This approach does not replace the standard approach (e.g., figure 1), but it provides additional information.

To examine country progress, we distinguish countries that are on target and countries that are off-target or lagging. We further differentiate lagging countries that are "close" to being on target from those that are "far" from being on track, forming three broad categories of performance.[6] Although there are alternative ways to describe progress, these three broad categories are intuitively appealing, and further refinement is likely to diminish the number of observations for each group because of data constraints (see below).

Illustration 1 How We Measure MDG Performance
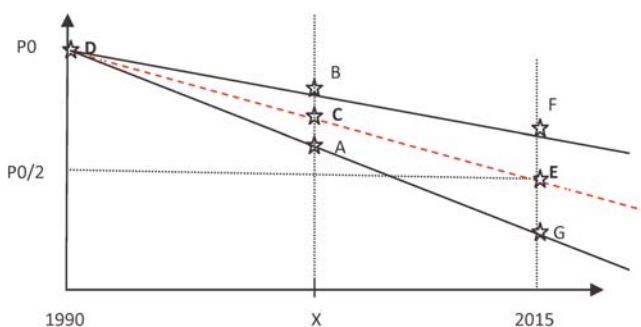
For example, a 50 percent reduction in poverty.
Source: Authors' description.

MDG performance in this paper is measured by deviations of the latest data from the trajectory required to reach the MDGs (similar to the idea in figure 1b but applied to individual countries). Different starting points imply a unique trajectory for each country to reach a specific MDG. Hence, comparing the slope or growth rate of the country's actual historical path with the required path (to meet the MDGs on time) is a good way to assess progress. The reference year for measuring progress is officially set as 1990. For each country and MDG indicator, we calculate the linear annualized rate of improvement required to reach the indicator's 2015 goal from the reference year. The illustration above shows how we measure MDG performance for a 50 percent reduction in extreme poverty.

A country is classified as on target if the latest actual or observed MDG performance, point A, meets or exceeds a point, such as C, that is suggested by the correct trajectory or trend to meet the goals by 2015. A country's annual rate of progress or slope between the reference year and the latest data implies an achievement path that will land the country at point G by 2015, which is more than enough to reduce poverty by 50 percent, as point E shows. An example is China. Since 1990, China has reduced its poverty rate by more than 70 percent, far above the 2015 target of reducing poverty. A country is considered off-target or lagging if its latest MDG performance, such as point B, falls short of this path. An example is Mali, where, instead of decreasing, the poverty rate increased by more than 25 percent from 1989 to 2006. Segment BC uses the country's most recent data to measure and illustrate its gap to becoming on target by 2015.

Within the off-target group, we consider two ways to further separate those countries that are close from those that are far from the target. In the main approach of the paper, the group's average distance to be on target for each MDG serves as a convenient or natural cut-off point to divide the lagging countries into two subgroups: off-target and above average and off-target and below average. We argue that lagging countries in the top half of the off- target and above average category are indeed "close to the target," whereas lagging countries in the bottom half of the off-target and below average category are "far from the target." The computed mean gaps are more conservative than the cut-off points used in Leo and Barmeier (2010), which defines lagging countries as close to target if their trajectory is within 50 percent of the required progress to reach the goals, earning half of a full score. In our methodology, we do not use an arbitrary cutoff point of 50 percent. Moreover, the mean gaps are all less than 50 percent across the MDGs, and they provide data-specific cutoff points to split the off-target countries.

Because mean gaps may conceivably be affected by outliers or spurious factors not addressed by the data, we also employ two absolute levels of closeness as alternatives: countries that are within 10 and 20 percent of becoming on target.

Detailed historical data on MDG performance are required to calculate the achievement path for each country to meet each of the MDGs. Unfortunately, such data are not available in many countries for 1990, although estimates for recent years tend to be more complete. If no country data are available for 1990, we use the closest available information in the late 1980s or early 1990s as substitutes for the starting point and then calculate the rate of progress required from that point to meet the MDG. This approach may be inaccurate if the data for the available starting point are significantly different from the level of MDG performance in 1990 or the sample period does not capture the latest progress. The latter is a particularly important issue now because data generally are yet not available for 2009, the year of the recent global economic crisis. In addition, for countries without at least two data points, progress cannot be measured even if

data are available for a recent year. Nevertheless, the approach allows us to include more countries than if we relied only on data from 1990 and 2008.

We restrict our attention to six MDGs and nine development targets with explicit and quantifiable 2015 goals (United Nations 2008). The following are the selected development targets:

- MDG 1.a: reduce by one-half, between 1990 and 2015, the proportion of people whose income is less than $1.25 a day (poverty headcount ratio at $1.25 a day, PPP, percent of the population).
- MDG 1c: reduce by one-half, between 1990 and 2015, the proportion of people who suffer from hunger (malnutrition prevalence, weight for age, percent of children under 5).
- MDG 2.a: ensure that by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling (primary completion rate, total, percent).
- MDG 3.a: eliminate gender disparity in primary and secondary education, preferably by 2005, and at all levels of education no later than 2015 (ratio of females to males in primary and secondary enrollment).
- MDG 4.a: reduce by two-thirds, between 1990 and 2015, the under-five mortality rate (mortality rate, under five, per 1,000).
- MDG 5.a: reduce by three-quarters, between 1990 and 2015, the maternal mortality ratio (maternal mortality ratio, per 100,000 live births).
- MDG 7.c: reduce by one-half, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation (improved water source and sanitation facilities, percent of population without access).

In what follows, we take a close look at MDG performance in developing countries with a particular focus on those countries facing larger gaps in terms of MDG achievement.

Variation in country performance

Appendix table S1.1 (supplemental appendix available at http:/wber. oxfordjournals.org/) summarizes the location of each developing country with respect to each of the six MDGs, according to the definition of performance above and where data are available. Figure 2 shows the distribution of countries according to the three groups: countries that are on target, close to the target, and far from the target for each MDG.

Although more developing countries are off-track than on track to achieve the targets, about three-quarters of developing countries are, on average, on target or close to being on target because of more than a decade of better policy and growth, as will be shown later in the paper. Of the three groups, about 45 percent of countries are now on target across the MDGs, and roughly another 30 percent

Figure 2. The Pattern of Country Performance by MDG
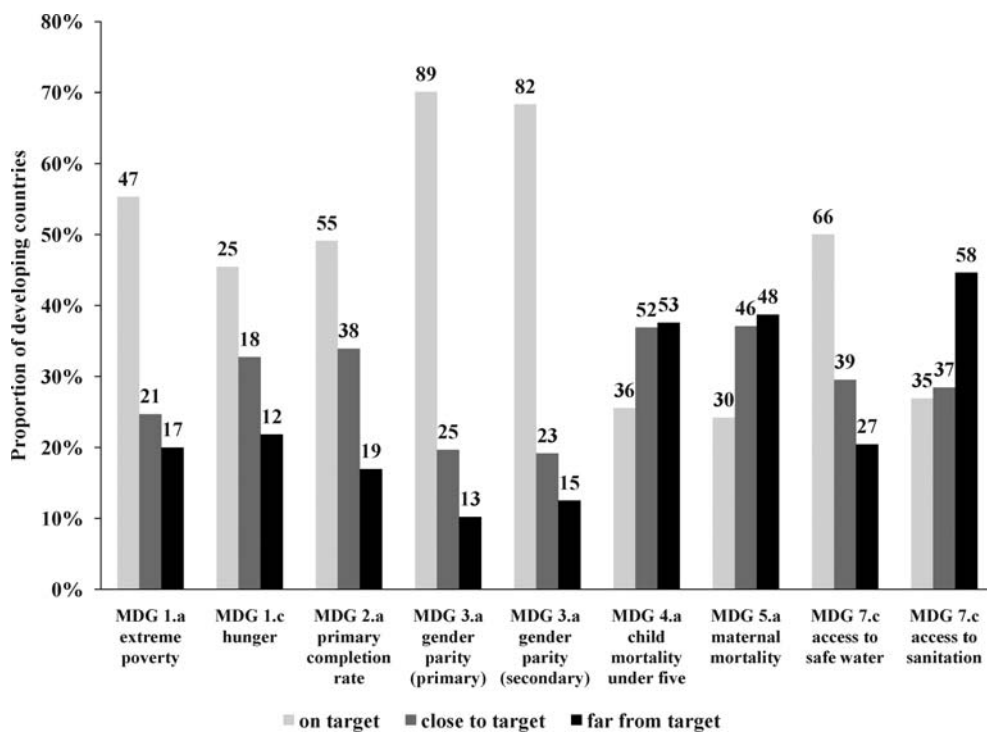
Distribution of countries by level of progress toward MDGs

Note: The number above each bar is the number of countries. A country is "close to the target" if its distance to getting on target (that is, its gap of trajectory) is smaller than the average gap of all lagging countries. Otherwise, it is "far from the target" (that is, its distance is greater than the average gap).

Source: Authors' calculations based on data from the World Development Indicators database.

are close to being on target. Countries that are far from being on target constitute the smallest group, at about 25 percent. Nevertheless, this group represents a significant percentage and concern.

For gender parity in primary education, 89 of the countries (or 70 percent) are on target, and another 25 (20 percent) are close to being on target. Regarding gender parity in secondary education, 82 of the countries (68 percent) are already on track, whereas 23 countries (19 percent) are close. For access to safe drinking water, 66 (50 percent) are on target, and another 39 countries (30 percent) are getting close to being on target. The primary completion rate in developing countries also shows encouraging signs: 55 countries (49 percent) are now on target, whereas 38 countries (34 percent) are close. With regard to reducing extreme poverty, 47 countries (55 percent) are on track, and another 21 (25 percent) are close.

In the next subsection, we argue that countries that are close to being on target are actually "close." Hence, if we take the performance of the first two groups to signify substantial progress, the picture from the variation in country performance is hopeful and not at all grim. For instance, the share of countries that are on target or close to being on target is very high for several MDGs— about 90 percent for gender parity in primary and secondary education and roughly 80 percent for the primary education completion rate, extreme poverty, access to safe drinking water, and reduction in hunger.

Progress is mixed or poor for access to sanitation, maternal mortality, and child mortality. Fewer countries are on track for these MDGs: 30 countries (or 24 percent) for maternal mortality, 35 countries (27 percent) for sanitation, and 36 countries (25 percent) for child mortality. In contrast, relatively more countries are far from being on target for these health-related indicators, ranging from 48 to 58 countries (about 37 to 45 percent). A silver lining that somewhat counter-balances the negative pattern that comes from the middle group (close to being on target). The number of countries in this category is substantial, ranging from 37 to 55 countries (about 29 to 38 percent).

Many middle-income countries are on target across the MDG indicators (table S1.1). In addition to being on target on several indicators, a number of these countries showed great achievement by having no single MDG classified as far from target. Examples include Albania, Armenia, Brazil, Chile, Ecuador, Egypt, Honduras, Lithuania, Iran, Macedonia, Malaysia, Nicaragua, and Sri Lanka. Of the large countries, China is on target for all MDGs except sanitation (which is close to being on target); India is on track on four indicators and close on another three; and Indonesia is on target or close to being on target for all MDGs, but information is lacking on its poverty rate (reference year, 1990). In the next section, we examine the role of initial incomes on MDG progress.

Like many middle-income countries, several low-income countries are doing well, but the pattern is not defined. Table 1 lists these countries by MDG, confirming that progress in many African and poor countries was strong.

Distance to being on Target among Lagging Countries

Although the variation among lagging countries is large, the average gap is not. Lagging countries are, on average, only 23 percent away from being on track to achieve all of the MDGs (table 2). They are especially close to the targets for gender parity in primary education (average gap of 7 percent), gender parity in secondary education (16 percent gap), reduction of hunger (19 percent gap), primary education completion (20 percent gap), and, to some extent, under-five mortality (23 percent gap). However, for each target, there are countries where progress has been scant or limited. For example, although the global goal will be

Table 1. Low-Income Countries that are Achieving the MDGs

Selected Millennium Development Goal	Low-income countries that have achieved the goal	Low-income countries that are on track to achieve the goal
Poverty	· Cambodia · Kenya · Mauritania	· Central African Republic · Ethiopia · Ghana
Universal primary education	· Myanmar · Tajikistan · Tanzania	None
Gender parity in primary education	· Bangladesh · Gambia, The · Ghana · Haiti · Kenya · Kyrgyz Republic · Madagascar · Malawi · Mauritania · Myanmar · Rwanda · Tanzania · Uganda · Zambia · Zimbabwe	· Benin · Burkina Faso · Burundi · Cambodia · Comoros · Ethiopia · Guinea · Nepal · Sierra Leone · Solomon Islands · Togo
Gender parity in secondary education	· Bangladesh · Kyrgyz Republic · Myanmar	· Gambia, The · Malawi · Mauritania · Nepal · Rwanda
Under-five mortality rate	None	· Bangladesh · Eritrea · Lao PDR · Madagascar · Nepal
Access to safe drinking water	· Afghanistan · Burkina Faso · Comoros · Gambia, The · Ghana · Korea, Democratic People's Republic of · Kyrgyz Republic · Malawi · Nepal	· Benin · Cambodia · Guinea · Uganda
Access to sanitation	· Lao PDR · Myanmar · Tajikistan	· Rwanda

Note: List of low-income countries is based on fiscal year 2011 World Bank classification; see table A1.13 in World Bank and IMF (2011a).

Source: Authors' calculation based on World Development Indicators database (as of March 2011).

Table 2. Average Gaps of Lagging Countries to Getting on Target

| | Average distance to getting on target (gaps, %) | | |
| | | Countries that are | |
	All off target countries	close to the target	far from the target
MDG 1.a Extreme poverty	39 (96)	17	67
MDG 1.c Hunger	19 (60)	9	35
MDG 2.a Primary education completion	20 (96)	9	40
MDG 3.a Gender parity in primary education	7 (22)	4	14
MDG 3.a Gender parity in secondary education	16 (52)	8	29
MDG 4.a Child mortality under five	23 (59)	8	38
MDG 5.a Maternal mortality	32 (80)	11	51
MDG 7.c Access to safe drinking water	25 (76)	14	41
MDG 7.c Access to sanitation	27 (50)	16	34
Simple average	23	11	39

Note: A country is "close to the target" if its distance to getting on target (that is, its gap of trajectory) is smaller than the average gap of all lagging countries. Otherwise, it is "far from the target" (that is, its distance is greater than the average gap). Figures in parentheses indicate the range of variation (Maximum value − Minimum value) of countries off target, by MDG. Averages and numbers of countries cover only those with data and that may vary by MDG.

Source: Authors' calculations based on data from the World Development Indicators database.

reached by 2015, several countries are far from reducing their extreme poverty by one-half.

The mean gaps of lagging countries are relatively larger for indicators such as access to safe drinking water, access to sanitation, maternal mortality, and reduction of extreme poverty. Nevertheless, these mean gaps are all noticeably less than 50 percent: access to safe drinking water, 25 percent; access to sanitation, 27 percent; maternal mortality, 32 percent; and reduction of extreme poverty, 39 percent.

More important, among countries that are off-track, the top half are, on average, only about 11 percent away from being on target. The mean distance of this subgroup is only 4−9 percent for gender parity in primary and secondary education, child mortality, primary education completion, and reduction of hunger. Indeed, countries that are close to the target need to increase primary education completion by only 9.2 percent (or 1.5 percent per year), on average, to be on track to reach the 2015 target.

Table 3 provides the proportion of countries within 10 percent or 20 percent of being on target. From another perspective, table 4 lists countries that are within 10 percent of being on target by MDG. In other words, this table shows that many lagging countries are already within striking distance of being on target. Although more arbitrary, these alternative and absolute levels of closeness are less

Table 3. Developing Countries that are within 10–20 Percent of being on Target

	Distribution of lagging countries			
	Gap ≤ 10 percent		Gap ≤ 20 percent	
	Number of countries	Proportion of countries (%)	Number of countries	Proportion of countries (%)
MDG 1.a Extreme poverty	9	24	13	34
MDG 1.c Hunger	10	33	18	60
MDG 2.a Primary education completion	23	40	39	68
MDG 3.a Gender parity in primary education	28	74	36	95
MDG 3.a Gender parity in secondary education	16	42	23	61
MDG 4.a Child mortality under five	33	31	48	46
MDG 5.a Maternal mortality	20	21	37	39
MDG 7.c Access to safe drinking water	10	15	32	48
MDG 7.c Access to sanitation	6	6	25	26
Simple average	17	32	30	53

Source: Authors' calculations based on data from the World Development Indicators database.

affected by outliers relative to mean gaps. By these measures, the closeness of lagging countries to being on target is also confirmed. One-third of off-target countries have, on average, a gap of 10 percent or less from being on target across the MDGs. Countries such as Bangladesh (reduction in extreme poverty, hunger, and maternal mortality), Indonesia (reduction in hunger, child and maternal mortality, access to safe drinking water), and Mali (gender parity in primary education and access to safe drinking water) are in this category. It is encouraging that more than half of these countries have a gap of 20 percent or less. Of the countries that are within 20 percent of target, the best results are for gender parity in primary education, primary education completion, gender parity in secondary education, and reduction of hunger. The worst results are for access to sanitation, reduction of extreme poverty, and reduction of maternal mortality, with access to safe drinking water and under-five mortality in the middle.

Country Patterns versus the Global Picture

The reference unit matters in a number of ways. Simple country averages that give equal importance to each country qualify the global story, which uses weighted averages that give more importance (i.e., a statistical bias) to countries with large populations. This can work in both directions:

Table 4. List of Lagging Countries that are within 10 Percent of being on Target

MDG 1.a Extreme poverty	MDG 1.c Hunger	MDG 2.a Primary education completion	MDG 3.a Gender parity in primary education	MDG 3.a Gender parity in secondary education	MDG 4.a Child mortality under five	MDG 5.a Maternal mortality	MDG 7.c Access to safe drinking water	MDG 7.c Access to sanitation
Bangladesh	Bangladesh	Bhutan	Belize	Bulgaria	Algeria	Algeria	Azerbaijan	Botswana
Burkina Faso	Bolivia	Cambodia	Cape Verde	Congo, Rep.	Antigua and Barbuda	Bangladesh	Colombia	Brazil
El Salvador	Egypt, Arab Rep.	Comoros	Chile	Georgia	Argentina	Brazil	Eritrea	Dominican Republic
Guinea	Indonesia	Cuba	Congo, Dem. Rep.	Grenada	Belarus	Cambodia	Haiti	Morocco
India	Jordan	El Salvador	Congo, Rep.	Guatemala	Bhutan	Cape Verde	Indonesia	Peru
Lao PDR	Kenya	Gambia, The	Djibouti	Macedonia, FYR	Cape Verde	Dominican Republic	Iran, Islamic Rep.	Turkey
Lesotho	Nigeria	Ghana	El Salvador	Madagascar	Colombia	Egypt, Arab Rep.	Kiribati	
Philippines	Pakistan	Guatemala	Grenada	Morocco	Dominican Republic	Ethiopia	Mali	
Uganda	Rwanda	Honduras	Guatemala	Pakistan	Ecuador	Haiti	Myanmar	
	Zambia	Iraq	Guinea-Bissau	Russian Federation	Ethiopia	India	Venezuela, RB	
		Jamaica	Jamaica	Senegal	Guatemala	Indonesia		
		Kyrgyz Republic	Lao PDR	Solomon Islands	Honduras	Lao PDR		
		Lebanon	Lebanon	Sudan	Indonesia	Mongolia		
		Lithuania	Maldives	Swaziland	Kazakhstan	Morocco		
		Macedonia, FYR	Mali	Vanuatu	Kiribati	Nepal		
		Mauritius	Mozambique	Zimbabwe	Kyrgyz Republic	Peru		
		Moldova	Nigeria		Liberia	Rwanda		
		Morocco	Paraguay		Libya	Syrian Arab Republic		
		Nepal	South Africa		Malawi	Tunisia		

Continued

Table 4. *Continued*

MDG 1.a Extreme poverty	MDG 1.c Hunger	MDG 2.a Primary education completion	MDG 3.a Gender parity in primary education	MDG 3.a Gender parity in secondary education	MDG 4.a Child mortality under five	MDG 5.a Maternal mortality	MDG 7.c Access to safe drinking water	MDG 7.c Access to sanitation
		Philippines	St. Vincent and the Grenadines		Moldova	Yemen, Rep.		
		South Africa	Sudan		Montenegro			
		Tanzania	Suriname		Niger			
		Turkey	Swaziland		Paraguay			
			Tajikistan		Russian Federation			
			Tonga		Samoa			
			Uruguay		Sri Lanka			
			Vanuatu		St. Vincent and the Grenadines			
			Venezuela, RB		Suriname			
					Syrian Arab Republic			
					Tajikistan			
					Turkmenistan			
					Uzbekistan			
					Yemen, Rep.			

Source: Authors' calculations based on data from the World Development Indicators database.

Country variation in performance generally softens the gloomier global picture. As shown earlier (figure 1, table 2), the average gap of lagging countries, especially in the top half, is small across the MDGs. Moreover, the percentage of countries that are on track or close to be on track is high when they are combined (75 percent). The statistics are remarkable, revealing progress that is much more varied and much more hopeful than the recent pessimism about achieving the MDGs. That pessimism was likely colored by the gaps at the global level, the difficult circumstances of poor countries, the potentially negative impact of the recent global crisis, and the lack of available data to assess outcomes. For example, although only 27 percent of low-income countries are on track to achieve or have achieved the target of reducing extreme poverty, almost 90 percent of these countries are in the top half of the lagging group and, therefore, have the goal of reducing extreme poverty within their reach. Similarly, about 40 percent of low-income countries are close to the primary education completion goal, although only 7 percent of the countries in this income group are on target.

That said, there are serious concerns arising from country variation. Although the proportion of countries that are in the bottom half of the off-target countries is lower (25 percent) than the other groups, these countries are disproportionately far from the targets, especially for the reduction of extreme poverty (67 percent on average) and maternal mortality (51 percent). This disproportionately higher distance for the bottom half of the off-target countries marks all MDGs except gender parity in primary education (table 2), pointing to the rather uneven distribution that affects MDG indicators. The range of variation is considerably large among off-target countries. For the reduction of extreme poverty and primary education completion, the gap between the countries that are closest to and farthest from being on target is 96 percent, a fact that clearly illustrates the variation in performance. This is the case for El Salvador and Uzbekistan for extreme poverty reduction and for Bhutan and Djibouti for primary completion rates. Clearly, the benefits of growth (if any) and, more important, of broad-based development are not reaching this last group of countries.

Looking at specific MDGs, the progress in reducing world poverty and meeting that goal is essentially the result of rapid advances by China and India, with the absolute number of poor people decreasing rapidly in China (although the absolute number of poor will still be large because of the size of the population). Despite the progress on extreme poverty, the average shortfall of lagging countries, at 39 percent, remains the largest among the MDGs. Among lagging countries in the bottom half, extreme poverty also has an average distance to being on target at a very serious, if not alarming, 67 percent. These observations underscore the importance of inclusive or development-based growth that raises everyone's quality of life and standard of living versus simple growth that raises only the average income. Take a related measurement called poverty gap at $1.25 a day,

which is the mean shortfall from the poverty line (counting the nonpoor as having zero shortfall) expressed as a percentage of the poverty line; it reflects the depth of poverty and its incidence. From 2006 to 2010, China had a poverty gap of 3.2 percent; India, 7.5 percent. Thus, the poor in these countries were already close to the international poverty line for extreme poverty, and growth would easily bring them across the threshold. For many countries in sub-Saharan Africa, however, the poverty gaps are high—for example, Madagascar's poverty gap is 43.3 percent; Zambia, 37 percent; Central African Republic, 31.3 percent; Malawi, 32.3 percent; and Tanzania, 28.1 percent (data are from the World Development Indicators). These African countries would need not only higher growth but also broad-based development to lift many individuals out of extreme poverty.

Regarding under-five mortality, the average distance to being on target is only 23 percent for lagging countries, somewhat less daunting than the global distance derived from the population of all under-five children. Moreover, the top half of lagging countries is only 8 percent from being on target. However, the distance of the countries far from the target is high, at 38 percent on average.

Although the progress of maternal mortality, an outcome-oriented goal, lags the most at the global level, there are hopeful signs at the country level. The average distance to being on target of the top half of lagging countries is only 11 percent. However, the average gap for all lagging countries is still disturbingly high, at 32 percent, second only to extreme poverty. The gap of the bottom half of lagging countries, at an alarming 51 percent, is the second highest.

The patterns at the aggregate and country levels generally support one another in the progress toward some of the more output-oriented goals—improving the primary education completion rate, reducing hunger, achieving gender parity in primary and secondary education, and providing access to safe drinking water. The lack of progress in sanitation is somewhat similar.

MDGs provide powerful benchmarks for measuring progress on key development outcomes, and one immediate impact is the effort to increase developing countries' statistical capacity to generate the related indicators. Although much progress has been made, there are still gaps within the existing indicators (see Appendix table S1.2). The above analysis demonstrates that these efforts need to extend to other outcome indicators across countries (such as learning outcomes in education versus completion rates) and to variation within countries (urban versus rural, attainment by income group) in order to better gauge how progress is distributed.

What Factors Affect the Rate of Progress?

Why are some countries on target, whereas others are not? Of the lagging countries, why are some close to target and others far away? The development factors

or driving forces often cited as the keys to attaining MDG-related development outcomes include economic growth as well as sound policies and institutions that are fundamental to effective service delivery to the poor. (see, for example, World Bank 2004). Although frequently cited and conceptually appealing as part of a natural working hypothesis, it is difficult to provide empirical documentation of their impact on achieving the MDGs. We pursue this approach further by examining whether the initial conditions of these factors or subsequent growth, policy and institutions improve the odds of reaching the goals. The analysis examines these elements in two ways: (1) using prima facie evidence from graphical associations and patterns, which point to these elements' likely association with the diverse progress of countries; and (2) in the next section, providing some simple statistical correlations and links in an attempt to answer the question of whether lagging countries can meet the MDGs by 2015.

Initial Conditions

Initial conditions count in MDG performance. In most cases, countries that are doing better (those that are on or close to the target) exhibited favorable starting conditions around 1990 (the reference year). A higher per capita GDP in 1990 is generally associated with better MDG performance (figure 3a).

Although there is no perfect indicator of the overall quality of policy and institutions in developing countries, the World Bank's annual Country Policy and Institutional Assessment (CPIA) provides a broadly consistent framework for assessing country performance on 16 items grouped in four clusters: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. The score is from one (low) to six (high) for each policy, which covers a wide range of economic and noneconomic issues, such as macroeconomic and fiscal policy, debt policy, trade, human development policy in education and health, gender equality, social protection, environmental policy, budgetary and financial management, and corruption in the public sector.[7] The index focuses on policies and institutional arrangements, the key elements that are within the country's control, rather than on actual outcomes (such as growth rates) that are influenced by elements outside of the country's control. Using the 1996 CPIA, the earliest available index with comparable scales and criteria,[8] suggests that countries that begin with good policy and institutions tend to do better in the MDGs (figure 3b). There are clearly other measures of policy and institutions, governance, and government effectiveness. We consider state capacity and fragility in this graphical section as well as other measures when we examine their statistical associations with MDG performance.[9]

Starting points—inherited initial conditions—explain why middle-income countries generally do better than low-income countries. Having grown earlier,

a. MDG performance and initial income conditions

b. MDG performance and initial institutional conditions

Note: A country is "close to the target" if its distance to getting on target (that is, its gap of trajectory) is smaller than the average gap of all lagging countries. Otherwise, it is "far from the target" (that is, its distance is greater than the average gap).

Source: Authors' calculations based on data from the World Development Indicators database.

they also tend to have implemented earlier a better set of policies and institutions. The link between the two factors is apparent in the following way—higher income brings greater resources to bear on a country's development problems while better policy and institutions ensure that those resources are allocated and used effectively to achieve better development outcomes. Hence, the initial levels of income and institutional capacity for good policy matter. However, there are variations. For extreme poverty and gender parity in primary education, countries with the fastest progress are those that experienced medium poverty and female-to-male primary enrollment ratios in the 1990s. The latter results draw attention to the challenges of poverty reduction in the proportionate way that MDGs are defined at low-income and middle-income levels. For poor countries, the distance to the goal is long; for middle-income countries, halving existing low poverty rates is difficult.

Growth and Policy

Although starting points (given their inherited nature) do not say much about what countries can or should do, they need not predetermine outcomes. The good news is that economic growth and policy performance after the initial year appear to count significantly, if not more than the starting points. The growth of income and the quality of that growth to elevate development—as manifested by the recent state of policies and institutions (2009)—appear to jointly move with MDG performance (table 5). Countries that are on target or close to being on target tend to have faster growth and better level of policies and institutions than countries that are far from the target. To help interpret the CPIA scores in table 5, a small variation in the overall score, such as a 0.1 increase, implies a significant improvement in development policy and institutions that is defined by construction (see also the next section for further discussion). Indeed, over time,

Table 5. Growth and CPIA Scores Are Higher in Countries that are on Track or Close to being on Track

Average values across MDGs (weighted by the number of countries in each MDG category)

	On target	Close to the target	Far from the target
Average GDP per capita growth (1990-2009)	2.4	1.8	1.2
Country Policy and Institutional Assessment Index (2009)	3.7	3.5	3.3

Note: The pairwise correlation between average GDP per capita growth and the CPIA index is 0.32 (significant at 0.01 level). GDP per capita, purchasing power parity constant 2005 international dollars. A country is "close to the target" if its distance to getting on target (that is, its gap of trajectory) is smaller than the average gap of all lagging countries. Otherwise, it is "far from the target" (that is, its distance is greater than the average gap).

Source: Authors' calculations based on data from the World Development Indicators database.

good policies and institutions are expected to lead to stronger future growth and better development outcomes such as poverty reduction, notwithstanding possible yearly fluctuations caused by external factors (World Bank 2007).

Low State Capacity or Government Failure

State capacity, fragility, or government failures, as the opposite of sound policy and institutions, are relevant. A state's ability to raise revenue, allocate and spend the revenue, and deliver critical public services to all of its citizens are important factors in the progress of MDGs. Besley and Persson (2011) recently developed a state capacity index. Figure 4a shows that progress is more pronounced at higher levels of state capacity.

Government failures, as reflected by the frequency and severity of conflicts, disastrous policy and institutional environments, and growth collapses, are also significant factors. Recent studies (World Bank and IMF 2010, Harttgen and Klasen 2010, Arbache et al. 2008) have drawn attention to the disproportionately negative effects of these factors on MDG performance and on human development indicators, such as child mortality, women's life expectancy, and education for girls. In broad terms, the data show that the proportion of on-target countries tends to rise with declining state fragility (figure 4b). In the graph, fragility is the index from the Center for Global Policy, which ranges from 0 (no fragility) to 25 (high fragility), divided into four categories ranging from little to extreme fragility (Marshall and Cole 2010).

All of these factors—especially state capacity, fragility, the initial conditions and subsequent growth, policy, and institutions—indicate why the MDGs are such significant challenges for the world's 79 poorest countries serviced by the World Bank's International Development Association (IDA) (figure 4c). IDA countries had a threshold per capita gross national income of $1,165 for fiscal year 2011, with average per capita growth and recent institutional performances that are well below average. Half of the IDA countries are in sub-Saharan Africa.[10]

Will Lagging Countries Achieve the Goals?

This question is not easy to answer because of the limitations of statistical analysis and tests on MDG performance. We briefly review the issues, selectively drawing findings from other studies. Because of the constraints, we resort to using simple approaches, such as pairwise correlations, to obtain some indication of the statistical strength of associations suggested by the graphs above. However, to identify answers to the broad question, we also attempt to examine countries' probability of falling into one of the three categories of success previously described in the second section, on target, close to target, or far from target, using a

Figure 4. MDG Performance, State Capacity and Fragility
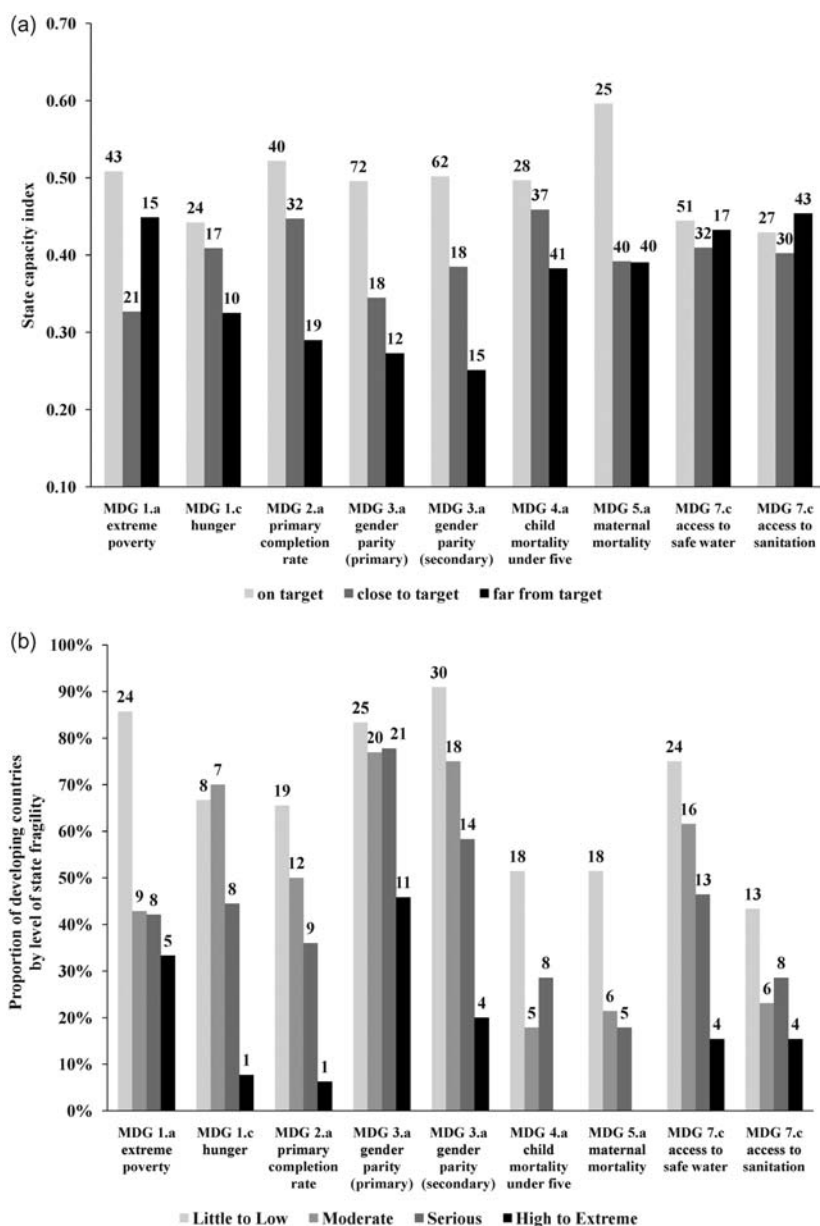
Continued.

set of development drivers defined below. This approach is limited by the descriptions of the MDG data and the set of factors employed. Hence, the results may not generalize and are preliminary and suggestive.

Figure 4. Continued

(c)
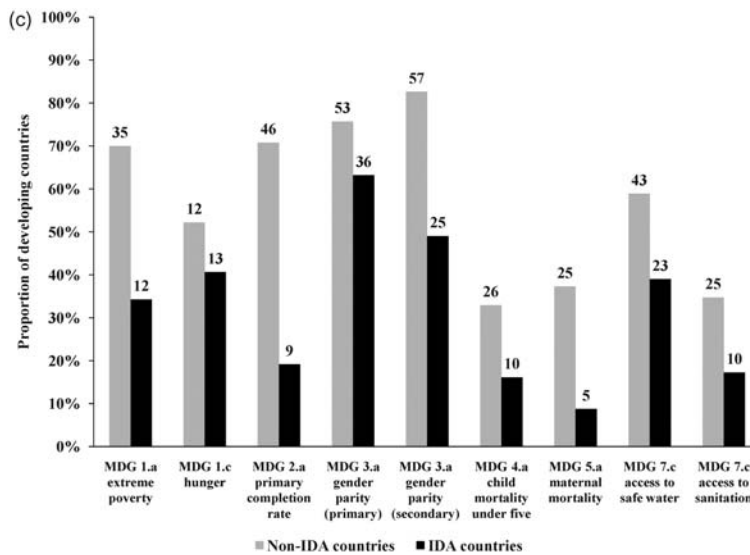
a. MDG performance and state capacity
b. Countries on target to achieve the MDGs by level of state fragility
c. MDG performance in IDA countries: countries on target
Note: Figures above or beside each bar indicate the number of countries
Source: Authors' calculations based on data from the World Development Indicators database, Besley and Persson 2011 and Marshall and Cole 2010.

Difficulty of International Comparison

Noting that development outcomes such as the reduction of both extreme poverty and child mortality are often not measured with high frequency, the first caveat is the quality of the available data. Information about MDG progress is even more restrictive than the underlying indicators. By definition, MDGs are concerned with the variation of the underlying indicators from the reference year (1990) to the latest year with the available data, which are then compared with the required change to reach the targets by 2015 (see the second section). Where data are available, there is essentially only one point of observation for each country and MDG, with no time series to improve estimation.[11] In low-income countries, the capacity to conduct household surveys and collect other relevant data is also weak. Appendix table S1.2 reports the status of data availability by MDG, income level, and region for this study.

Second, specific factors are suggested by recent micro studies and impact evaluations to be strongly relevant for MDGs, particularly those affecting human

development. These factors tend to be complex and wide ranging and to have great specificity of policy in their respective areas. Factors that affect child health issues, for example, may not be relevant for others. Evidence from 172 micro data sets from Demographic and Health Surveys for more than 70 countries in Günther and Fink (2010) found that child mortality and the incidence of diarrhea benefit considerably from access to certain facilities (such as a latrine, flush toilet, pump, well, spring, and piped water), while controlling for mother's education and age and other fixed effects, such as electricity and indications of household assets and wealth (such as radio, refrigeration, bikes, and urban location). The findings from more than 70 impact evaluations of projects (World Bank and IMF 2011) indicate that, in addition to the effectiveness of service delivery and supply factors, demand factors as well as the accountability and incentives of service providers and clients are significant in improving health and education outcomes. A key lesson suggests that policy interventions should go beyond supply-side improvements, budget allocation, or input provision. The list of policies is not only wide ranging but also context specific. Examples include addressing uptake issues of health or education services by households, facility-based care and services, community-based interventions and support, such as training and group sessions, cost sharing of services, cash transfer programs to target poor people, information to change behavior and increase accountability, and pay-for-performance programs targeting specific health or education workers. These findings are consistent with those in Devarajan and Reinikka (2004) and World Bank (2004).

The scope of policy interventions suggested by micro studies contrasts with more clearly defined reform measures and variables that are available for traditional macroeconomic, fiscal, or trade policy. To obtain international comparisons of aggregate MDG indicators, authors have utilized broad or multidimensional indices, such as proxies and summaries of the range of policy interventions, the quality of policy and institutions, corruption levels, the degree of country fragility, and the level of state capacity. For example, Wagstaff and Claeson (2004) used the CPIA for policy and institutions to better explain the effects of public health spending on health-related MDGs. In an earlier study, Filmer et al. (2000) concluded that the links between public spending and health results are weak if poorly functioning facilities, demand-side factors, and other factors in the chain are not considered. Rajkumar and Swaroop (2008) found that corruption and bureaucratic quality mattered significantly in terms of the effects of public health spending on different health outcomes. Baldacci et al. (2008) found that public spending and health outcomes appear stronger if the analysis accounts for governance. For water and sanitation, however, it is not sufficient to include the central government's spending on infrastructure; the more relevant indicator is the capital spending of local governments or the local public entities in urban

centers that are responsible for providing water and sanitation. However, 84 percent of the 884 million people who lacked access to safe water in 2010 were in rural areas. Hence, public spending for water wells and storage tanks is important. Furthermore, private income levels and spending on food clearly affect poverty, nutrition, hunger, and various health and education indicators.

Other issues exist. The direction of the impact between broad development outcomes and broad development drivers and institutional factors is likely to go both ways, and the drivers are likely to be correlated. Unlike per capita income, policy and institutional variables are generally not comparable or available over time. To account for the many factors that are not readily measurable or available, studies have employed various fixed effects models to reflect varying initial conditions. However, these fixed effects tend to be specific to certain MDGs. For a survey of additional literature and issues, see Lay (2010) and Lofgren (2011).

Overall, comprehensive micro data sets, such as those in Günther and Fink (2010) for child mortality, would be ideal for investigating the determinants of the evolution of the other MDGs. Each MDG would likely require a separate set of micro data and a separate study. As additional impact evaluations are undertaken, additional lessons may be generalized from case studies of different MDG areas. These approaches, however, are outside the scope of this paper. Moreover, it seems important that specific policy interventions in particular context and circumstances add up to system-wide improvement in policy and institutions conducive to inclusive growth and sustainable development. Acemoglu and Robinson (2012) argued that policy and institutions in the broadest sense matter; inclusive political and economic institutions explain a large part of nations' long-term economic prosperity and successful development outcomes. There are, of course, many ways to measure policy and institutions for the empirical analysis of MDGs, which we consider in the next section.

In view of the constraints involved in a cross-country analysis of aggregate MDG performance, we employ simple approaches or methodologies that focus on certain aspects of the study. First, we establish the strength of correlations among the broad factors. Next, we limit further analysis to probability functions of the different categories of MDG performance through the multinomial logit model (versus using the underlying indicators). For comparability across MDGs in the probability functions, we use a common set of potentially independent variables. This analysis is not a substitute for in-depth studies of each underlying MDG indicator or its associated development outcomes.

Simple Correlations

Are the MDGs and the broad development factors suggested by the graphs correlated statistically? For the MDGs, we examine the underlying indicators in terms

of their levels. The list of development factors, which is not meant to be exhaustive, includes economic growth, income levels and alternative measures of policy and institutions. The lack of a perfect indicator for policy and institutions means that several alternative definitions and possibilities are available. For our purpose, we examine more than a dozen policy and institution-related variables covering a wide range of issues, such as conflict, corruption, state capacity, violence and political rights, and the CPIA. Table 6 presents the list of institutional variables as well as their pairwise correlations. One observation stands out immediately: there is a high number of significant associations and observed correct signs, which suggests some consistency among the different measurements of policies and institutions, making these measures substitutes for one another. Because these measures of policy and institutions are highly correlated despite some differences in their definition and measurement, it is difficult to include more than one of them in the same empirical relationship without encountering the statistical problem of multicollinearity.

Table 7 shows the pairwise correlations between the levels of the MDG indicators and the various factors in the list. Almost all of the correlation coefficients have the right sign of association, and the number of coefficients that are significant at 10 percent level or better practically fill the matrix. Thus, the results provide broad empirical support for the intuitive argument of the graphs: if growth, income level, and various policies and institutions continue to improve in developing countries, the underlying indicator of each MDG will likely also improve. The direction of effects almost certainly goes both ways. These observations are likely the minimum one can safely identify about the associations examined.[12]

Among the list of factors, the level of income (GDP per capita) and the state fragility index from Marshall and Cole generally have the highest correlation values across the board. It is interesting to note that it is the level of income, which is a positive effect and reflection of economic growth, rather than growth itself, that has a higher correlation value. Relatively high correlation can also be found in factors such as the indexes of prosperity and state capacity in Besley and Persson (2011), the World Bank's CPIA (for both 1996 and 2009), the control of corruption, government effectiveness, rule of law and regulatory quality from Kaufmann et al. (2009), good governance in Knack and Kugler (2002), and the functioning of government of the Economist Intelligence Unit (2007). In the case of the CPIA, one of its major components relates to a country's economic management, which was likely to be more affected by the global economic crisis in 2009 than other indicators. However, its correlation coefficients (in tables 6 and 7) appear stable and comparable to a pre-crisis CPIA index in 2006. Besley and Persson's state capacity is relatively new and is a component of their Prosperity Index. It is defined as the government's ability to levy an income tax (i.e., a share of income to

Table 6. Pairwise Correlations Between Institutional Variables

	CPIA 2009	CPIA 2006	Prosperity index	Peacefulness	State capacity	State fragility index	Management performance	Functioning of government (EIU)	Functioning of government (Freedom House)	Good governance	Voice and accountability	Political stability -no violence	Government effectiveness	Regulatory quality	Rule of law	Control of corruption
CPIA 2009	1.00															
CPIA 2006	0.96	1.00														
Prosperity index (Besley and Persson 2011)	0.37	0.38	1.00													
Peacefulness (Besley and Persson 2011)			0.75	1.00												
State capacity (Besley and Persson 2011)	0.28	0.30	0.68		1.00											
State fragility index (Marshall and Cole 2010)	-0.63	-0.64	-0.69	-0.36	-0.34	1.00										
Management performance (Bertelsmann Transformation Index 2006)	0.77	0.80	0.33	0.22		-0.59	1.00									
Functioning of government (Economist Intelligence Unit 2007)	0.62	0.66	0.26			-0.61	0.74	1.00								
Functioning of government (Freedom House)	0.50	0.54	0.20			-0.47	0.85	0.73	1.00							
Good governance (Knack and Kugler 2002)	0.45	0.47	0.29			-0.52	0.40	0.50	0.37	1.00						

Voice and accountability (Kaufmann et al. 2009)	0.46	0.49	0.24	0.21		-0.55	0.87	0.75	0.92	0.31	1.00					
Political stability -no violence- (Kaufmann et al. 2009)	0.35	0.33	0.56	0.44	0.24	-0.64	0.62	0.52	0.53	0.32	0.59	1.00				
Government effectiveness (Kaufmann et al. 2009)	0.79	0.80	0.47		0.31	-0.71	0.82	0.71	0.65	0.57	0.65	0.60	1.00			
Regulatory quality (Kaufmann et al. 2009)	0.79	0.83	0.38	0.17	0.23	-0.63	0.84	0.66	0.66	0.45	0.71	0.51	0.88	1.00		
Rule of law (Kaufmann et al. 2009)	0.61	0.60	0.43	0.21	0.24	-0.66	0.79	0.69	0.68	0.47	0.71	0.76	0.85	0.77	1.00	
Control of corruption (Kaufmann et al. 2009)	0.60	0.64	0.45	0.23	0.26	-0.65	0.75	0.63	0.67	0.52	0.67	0.67	0.86	0.76	0.86	1.00

Note: All presented correlations are significant at 10% level or better.
Source: Authors' calculations.

Table 7. Pairwise Correlations of MDG Indicators in Levels and Various Development Factors (c.2009)

	Poverty	Malnutrition prevalence	Primary completion rate	Ratio of girls to boys in primary education	Ratio of girls to boys in secondary education	Under 5 mortality	Maternal mortality	People without access to safe water	People without access to sanitation facilities
Average growth in GDP per capita (1990-2009), 2005 I$PPP	-0.22		0.26	0.15	0.22	-0.25	-0.21	-0.21	-0.24
GDP per capita 2009, 2005 I$PPP	-0.73	-0.66	0.58	0.34	0.49	-0.64	-0.62	-0.59	-0.67
CPIA 2009	-0.41	-0.43	0.39	0.33	0.41	-0.44	-0.48	-0.41	-0.43
CPIA 2006	-0.43	-0.42	0.38	0.36	0.43	-0.45	-0.50	-0.43	-0.42
Prosperity index (Besley and Persson 2011)	-0.54	-0.65	0.46	0.44	0.54	-0.59	-0.58	-0.53	-0.49
Peacefulness (Besley and Persson 2011)		-0.32			0.17			-0.17	
State capacity (Besley and Persson 2011)	-0.42	-0.48	0.47	0.42	0.40	-0.47	-0.46	-0.33	-0.38
State fragility index (Marshall and Cole 2010)	0.74	0.72	-0.60	-0.49	-0.70	0.80	0.76	0.72	0.66
Management performance (Bertelsmann Transformation Index 2006)	-0.26	-0.31	0.26	0.30	0.38	-0.26	-0.25	-0.30	-0.18
Functioning of government (Economist Intelligence Unit 2007)	-0.35	-0.21	0.38	0.37	0.51	-0.43	-0.41	-0.36	-0.22
Functioning of government (Freedom House)	-0.23	-0.23	0.26	0.21	0.34	-0.26	-0.19	-0.27	

Good governance (Knack and Kugler 2002)	-0.39	-0.26	0.33	0.33	0.38	-0.49	-0.50	-0.33	-0.34
Voice and accountability (Kaufmann et al. 2009)	-0.28	-0.30	0.33	0.25	0.39	-0.32	-0.25	-0.34	-0.15
Political stability -no violence- (Kaufmann et al. 2009)	-0.26	-0.32	0.36	0.36	0.43	-0.38	-0.37	-0.34	-0.23
Government effectiveness (Kaufmann et al. 2009)	-0.51	-0.43	0.49	0.43	0.55	-0.55	-0.51	-0.54	-0.46
Regulatory quality (Kaufmann et al. 2009)	-0.43	-0.41	0.32	0.32	0.39	-0.42	-0.43	-0.44	-0.32
Rule of law (Kaufmann et al. 2009)	-0.36	-0.33	0.43	0.43	0.50	-0.50	-0.48	-0.49	-0.35
Control of corruption (Kaufmann et al. 2009)	-0.38	-0.38	0.37	0.35	0.42	-0.46	-0.43	-0.45	-0.32

Note: All presented correlations are significant at 10% level or better.

Source: Authors' calculations.

generate government revenue to cover government expenditures), interpreted as the fiscal constraint to choose levels of redistributive transfers and provisions of public goods and services. This index and the various indicators of governance clearly influence the effective delivery of public services emphasized by the various micro studies.

Likelihood of Success

Beyond broad correlations, is it possible to make statements about the effect of these factors on the MDG gaps? We limit the investigation to the probability of MDG success using previously defined MDG performance categories and probability functions, such as the multinomial logit model. The results are clearly dependent on how MDG progress is described in this study and how the MDG targets are established in the first place (versus more general development outcomes) and may therefore not generalize to other ways of describing or explaining success and development outcomes, more in-depth analysis of each underlying MDG indicator, or other approaches. Nonetheless, by using categorical or discrete values of MDG performance, we hope to minimize the two-way interactions between the dependent and independent variables. We also use a common format and set of initial conditions to avoid overly fine tuning the relationship.

The multinomial logit model is intuitive and well suited for assessing the likelihood of a country falling into one of the three defined categories (on target = 1; close to target = 2; and far from target = 3), linking performance to the various development drivers. This type of model is typically employed to model individual discrete choices, such as the occupational choice of households in micro simulations or the demand for modes of transportation. Go and Quijada (2011) discuss various methodological issues associated with this approach.[13] Our baseline representation takes the following form:

MDG performance=f {GDP per capita, policy and institutions, initial conditions},

where the initial conditions include the GDP per capita and the level of MDG indicators circa 1990. We estimate 17 different specifications for each of the nine MDG targets under analysis. The reference category is "far from target." Models 1 to 16 differ in the variable for policy and institutions, which is combined pairwise with the level of income in each nonlinear regression. Model 17 considers only initial conditions and GDP per capita growth between 1990 and 2009 as independent variables. Unlike income level in the other models, economic growth is a change variable, which is generally easier to assess and project in the few years remaining until 2015.[14]

Table 8 summarizes the main findings. We notice a high degree of significance for many of the development drivers. GDP per capita in combination with one of

Table 8. Multinomial Logit Estimates: Likelihood of MDG Success as Explained by Level of Development and Institutions

	MDG 1.a extreme poverty		MDG 1.c hunger		MDG 2.a primary completion rate		MDG 3.a gender parity (primary)		MDG 3.a gender parity (secondary)		MDG 4.a child mortality under five		MDG 5.a maternal mortality		MDG 7.c access to safe water		MDG 7.c access to sanitation	
	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target
1 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+
CPIA 2009		+	+	+	+	+	+			+								
2 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+		+	+	+	+			+	+	+	+
CPIA 2006			+	+			+	+	+	+	+	+	+	+				
3 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Prosperity index (Besley and Persson 2011)			+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4 GDP per capita 2009 (thousands, 2005 I$PPP)		+	+	+	+	+	+	+	+	+	−		−		−		−	−
Peacefulness (Besley and Persson 2011)		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
State capacity (Besley and Persson 2011)		+	+	+			+	+	+	+			+					
6 GDP per capita 2009 (thousands, 2005 I$PPP)			+		+	+	+	+	+	+	+				+		+	+
State fragility index (Marshall and Cole 2010)	−	−		−	−	−	−	−	−	−	−	−				−		

Continued

Table 8. Continued

	MDG 1.a extreme poverty		MDG 1.c hunger		MDG 2.a primary completion rate		MDG 3.a gender parity (primary)		MDG 3.a gender parity (secondary)		MDG 4.a child mortality under five		MDG 5.a maternal mortality		MDG 7.c access to safe water		MDG 7.c access to sanitation	
	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target	Close to target	On Target
7 GDP per capita 2009 (thousands, 2005 I$PPP)		+	+	+	+	+		+	+	+					+	+	+	+
Management performance (Bertelsmann Transformation Index 2006)	+		+	+	+		+									+		
8 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+			+	+	+	+			+	+	+	+
Functioning of government (Economist Intelligence Unit 2007)			+	+	+	+				+					+	+	+	
9 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+			+	+	+	+	+		+	+	+	+
Functioning of government (Freedom House)			+	+	+	+												–
10 GDP per capita 2009 (thousands, 2005 I$PPP)		+	+	+	+	+			+	+	+	+	+		+	+	+	+
Good governance (Knack and Kugler 2002)	+				–	–						+						
11 GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+			+	+	+	+	+		+	+	+	+
Voice and accountability (Kaufmann et al. 2009)	+		+		+	+									–	–		–

#	Variable														
12	GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+
	Political stability –no violence– (Kaufmann et al. 2009)	+	+	–	+					–	+	+	+	–
13	GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+
	Government effectiveness (Kaufmann et al. 2009)	+	+	+	+				+	–	+	+	+	
14	GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+
	Regulatory quality (Kaufmann et al. 2009)	+	+	–	+	+	+		+		+	+		+
15	GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+
	Rule of law (Kaufmann et al. 2009)	+	+	+	+	+	+		+	–	+	+		+
16	GDP per capita 2009 (thousands, 2005 I$PPP)	+	+	+	+	+	+	+	+	+	+	+	+	+
	Control of corruption (Kaufmann et al. 2009)	+	+	+	+				+	–	+	+		+
17	Average growth in GDP per capita (1990–2009), 2005 I$PPP		+	+	+	+	+	+	+	+	+	+	+	+

Notes: "Far from target" is the reference category. Initial conditions not showed. (+) denotes positive and significant coefficients at 10% level or better. (–) denotes negative and significant coefficients at 10% level or better. Blank cells indicate nonsignificant coefficients.

Source: Authors' calculations.

the policy and institutional variables seems to be significantly related to better MDG performance. Economic growth by itself appears to be a good factor (model 17), because it is correlated with the other variables, such as income level and policy and institutions. Although growth may be affected by favorable external shocks, the ability of countries to benefit from favorable events will likely improve with better policy and institutions.

The level of income is generally important across specifications and development goals. However, in two cases, gender parity in primary education and reduction in maternal mortality, GDP per capita seems less important. In the case of gender parity in primary education, this finding may not be surprising given that most countries are already on target to reach this goal by 2015 (figures 1 and 2). In the case of maternal mortality, there may be several factors at work. As a system-based outcome, it may require many improvements in the health system, including incentives and the accountability of all players, as the micro studies suggested. As the MDG with the slowest global progress, it also has the largest gaps, which can be closed only partially by higher income. In any case, higher levels of income, as well as economic growth in the last model, are significantly and positively associated with the likelihood of a country being "close to target" versus "far from target" in almost three-fifths of our model specifications.

The linkages between MDG achievement and better policy and institutional frameworks also appear strong, although the results are more varied. The lack of a single ideal institutional indicator at the aggregate level is a plausible reason for the varied findings. Each version of the institutional variable may not capture all of the specific policy interventions considered important in micro data and case studies, particularly when considering child and maternal mortality. Some of the institutional variables may be correlated with the level of income. Hence, their coefficients do not always register with the expected signs (see, for example, access to safe water and sanitation). Similarly, for the income variable, we do not find significant linkages for gender parity in primary education and maternal mortality, and the reasons are likely the same. Nevertheless, it is important to highlight that most of the policy and institutional variables in our specifications are significantly and positively linked to income- and education-related MDGs, such as reduction in extreme poverty, hunger, and cross-gender completion of primary school. Within the limitations of the variables and methodology, the clear winners, in terms of being more consistently significant across MDG performance, are the pre-crisis CPIA (year 2006) and the state fragility index. Both variables are significantly linked to better MDG performance for at least seven of the goals under analysis, including health-related development targets. It is not surprising to find that the CPIA index is significantly correlated with health-related MDGs given that its components include policies for human development, social inclusion and equity. In examining the marginal effects of growth and CPIA, Go and

Quijada (2011) confirmed that growth has a significant effect on the progress of MDGs, whereas good policies and institutions are especially crucial for system- and outcome-based MDGs, such as the reduction of child and maternal mortality.

Finally, we turn to the question of whether higher income and better policy and institutions will improve the likelihood of better MDG results among lagging countries. As an illustration, we take model 2, which is as good as any in table 8. We consider a quarter standard-deviation increase in GDP per capita and in the quality of policy and institutional assessments for the period from 2009 to 2015.[15]

The results from this simulation show that higher income and better policy and institutions can jump-start lagging countries (figure 5). Many more developing countries can get on track, particularly for those MDGs with the greatest lag. A quarter-standard-deviation rise in both per capita income and the CPIA would mean that as many as 26 more developing countries can get on track for the

Figure 5. The Number of Countries Becoming on Track with Higher Income and Better Policy and Institutions
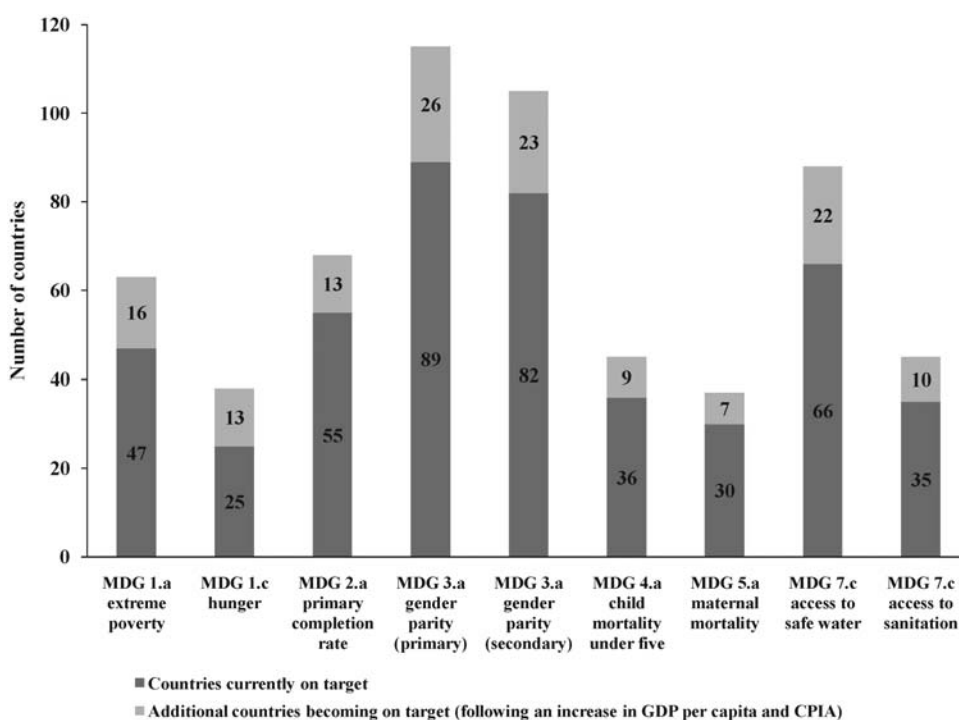
Note: The results are for a quarter standard-deviation increase in GDP per capita and CPIA index in model 2.
Source: Authors' calculations.

MDGs—an average increase of 31 percent in the number of on-track countries. This forecast is based on a greater than 50-percent probability of each country getting on target. Statistically, the probability of lagging countries can only reach 100 percent as an upper (asymptotic) limit, but a 95-percent confidence interval of a 50-percent increase will generally cover that upper limit. The percentage increase in the number of countries getting on track generally increases most for targets such as reduction of hunger (52 percent), reduction of extreme poverty (34 percent) and access to safe drinking water (33 percent). For the other MDGs (primary education completion, gender equality in primary and secondary education, and reduction in both child and maternal mortality), the increase in the number of countries is above 20 percent, which is still substantial. Individual countries that are good candidates to get on track are those that are currently very close—that is, within 10 percent of being on track (table 3).

How achievable or feasible are these gains? Recent history suggests they may be attainable or close to attainable, but prospects look uncertain or less likely given a weak global economy since the Great Recession of 2008–09. Achieving a quarter-standard-deviation gain in income level means that per capita GDP growth in developing countries will need to increase by 3 percent per year from 2009 to 2015, 1.6 times its historical rate of 1.9 percent a year. That kind of growth performance was achieved by developing countries, including those in the two lagging groups, during the boom period from 2003 to 2007 (table 9). However, world economic and trade conditions have since become much less favorable.[16] In addition, aid flows from donor countries may decline as a result of weaker fiscal conditions in those countries.[17]

Because serious warnings are now attached to growth prospects and aid flows, reforming policy and institutions becomes both important and necessary to ensure that domestic revenues and efforts can offset the risks of such prospects and flows shrinking when they need to expand to help the developing countries either meet or come close to their MDGs. Such reforms are likely to help these countries to avoid growth collapses or government failures because fragility has a negative impact on the progress of MDGs (figure 4b, model 6 in table 8, and World Bank and IMF 2010.) In the illustration above, a quarter-standard-deviation gain in the CPIA is about a 0.1 improvement in the overall CPIA score and represents a significant policy improvement for a country; it is half of the difference between the CPIA for on-target countries and for countries close to the target (see table 5). From 2006 to 2009, 55 countries (43 percent of developing countries for which scores are available) experienced an improvement of 0.1 points or better. These countries include Georgia, Nigeria, Djibouti, and Peru. For better results in the MDGs, additional policy improvements will continue to be needed. In this regard, a 0.2- or 0.3-point increase in the CPIA represents a substantial policy shift or regime change, which is rare for any country in a given

Table 9. Recent Growth Performance in Developing Countries

Growth periods	Years covered	Growth of per capita GDP of developing countries under alternative MDG performance		
		on target	*close to the target*	*far from the target*
I. Reference period	**1990–2009**	2.42	1.77	1.22
II. Recent growth accelerations				
Modern trend-break	**1995–2007**	3.46	2.61	2.01
New millennium	**2000–2007**	3.97	2.90	2.25
Boom years	**2003–2007**	4.82	3.65	3.07
III. Recent global economic crisis				
Crisis years	**2008–2009**	1.48	1.79	1.48
Peak crisis	**2009**	-1.09	0.28	0.65

Source: Authors' calculations based on data from the World Development Indicators database.

year but conceivable and likely over time. Because policy reforms take time to be designed and implemented and to bear fruit, they should be undertaken as soon as possible.

Two final caveats need to be noted. First, MDG performance in lagging countries close to being on target will need to accelerate soon for them to reach their MDG targets by 2015. This mathematical constraint is reflected in the following way. If these countries simply continue on their historical growth rates, however decent, the gap will widen by 2015 (segment FE versus BC in illustration 1). With only a few years left for developing countries to meet the MDGs by 2015, depending on how recent the data are for each country, the problem of actually meeting the MDGs will become crucial.

The second caveat concerns missing observations that may affect the robustness of the results. Such missing observations seem unlikely, however, on the basis of indirect evidence. The "missing countries" by MDG are generally not the "basket cases" with respect to the two explanatory factors in the models used—growth and policy; nor are they the exceptional cases (i.e., the averages tend to the middle). Hence, missing observations are unlikely to tilt the results in either direction (see Go and Quijada (2011) for more details).

Final Remarks

In this paper, we show that three-quarters of developing countries are on target or close to being on target for all of the MDGs, which is unexpectedly encouraging. Moreover, among the countries that are falling short, the average gap for the top half is about 10 percent. For those that are on target, or close to it, solid economic growth, policies, and institutions have been the key factors in their success. Improving developing outcomes further will require not only increases in

GDP per capita but also system-wide improvements in policy and institutions that bring inclusive growth or broad-based development in order to improve the living conditions, opportunities, and quality of life of all individuals, groups, and nations in the world. Although there are variations and complications, this vital distinction between growth and development has a clear resonance in the main findings of the study. With some simplification, growth (which brings more money and resources) tends to improve the more output-oriented goals such as primary completion rate, access to drinking water, and gender equality in terms of ratios of girls to boys in primary and secondary schools. However, the more outcome-oriented goals in the health sector such as maternal and child mortality tend to require system-wide improvement in the quality of policies and institutions. This is also especially true for the 25 percent of developing countries that are lagging the most across the MDGs, where the remaining gaps are disproportionately high. The same distinction between growth and its quality also partly explains two opposing results in the income-based measure of poverty—that rapid growth in many developing countries has ensured that the goal on extreme poverty will be scaled at the global level; however, the gaps in lagging countries are still the largest among the MDGs.

By examining country-level figures rather than global figures, recent historical data indicate that developing countries are clearly doing better. Lagging countries, on average, are very close to their MDG targets, and their odds of getting on track can improve dramatically with stronger growth and sounder policy and institutions (i.e., development that benefits also the most vulnerable and truly needy people and that undoes unfavorable conditions that limit their quality of life). The implications are clear. With 2015 less than a few years away, stronger growth in developing countries must be stimulated to a higher plane strategically and quickly, a rapid—but sustainable in the long run—way of moving more countries toward the MDGs and preventing them from subsequently slipping. This goal will not be easy, however, if global economic and trade conditions continue to be unfavorable and donor support continues to deteriorate. This situation is unfortunate because growth was accelerating before 2008, and progress on the MDGs was evident in many countries.

As developing countries face a less friendly global economy and a dangerous period of increasing economic vulnerability, the challenge will be to continue improving policy and institutions to maintain progress and to avoid both growth collapses and government failures, which tend to have very negative effects on the MDGs. Further improvement in policy and institutions is especially necessary not only because of the short time left to 2015 but also because of the more difficult challenges in both the MDGs and countries that are lagging the most. Improved policies and institutions are crucial to improve not only the income aspect of growth but also its quality and effects on the poor. For countries close to the

target and where growth has already taken place, further gains in development outcomes will also require further improvements in policy and institutions. Even the middle-income countries on track to attain the MDGs are home to indigenous and socially excluded groups that are still very poor and often well behind in many development outcomes (World Bank and IMF 2011).

How to bring about stronger (i.e., true development-based economic) growth and what constitutes "good" policies and institutions in developing countries are complex issues that cover a wide range of areas, problems, and concerns. These issues are not limited to economic areas such as macroeconomic and fiscal policy, debt policy, and trade but include broader issues such as human development policy in education and health, gender equality, social protection, environmental policy, budgetary and financial management, and corruption in the public sector. Policy-based interventions should be not only broad and wide-ranging in order to foster sustained development but, as micro studies have shown, also appropriately specific to needy groups as well as local circumstances and problems. Although these complex issues are clearly beyond the scope of this work, we hope that this paper has provided further insights to the central challenges of development.

Notes

1. This observation has been widely documented. For Africa, see Arbache, Go, and Page (2008) and Ndulu (2008).

2. World Bank and IMF (2010) discussed the impact of the recent global economic crisis on the MDGs.

3. For more details, see World Bank (2011) and United Nations (2008).

4. Data used in this paper were those available during the drafting of World Bank and IMF (2011). More recent data and trends compiled in World Bank and IMF (2012) indicate that the goals for poverty and safe drinking water would have been reached in 2010.

5. Fukuda-Parr and Greenstein (2010) state that development goals are not "hard-planning targets" but rather guidelines "meant to encourage countries to strive for accelerated progress." Their approach consists of comparing rates of change in development indicators before and after 2001, the year the United Nations outlined its strategy for achieving the MDGs, assuming that progress should be measured against the moment MDGs were adopted. Moreover, measuring broad development outcomes through specific indicators is never precise, so the variation in MDG performance is partly the result of indicator or measurement issues. However, we do not examine these issues here. For discussion of some of the issues in measuring broad development outcomes through the Millennium Development Goals, see box 1.2 of World Bank (2011).

6. In what follows, the terms "on target" and "on track" are used interchangeably.

7. The scores are available in the World Development Indicators database.

8. An earlier version of the CPIA goes back to the 1970s but uses a different scale and criteria. For example, the assessment of governance issues was not included in the earlier CPIA.

9. We also looked at several dimensions of trade—export sophistication and shipping connectivity, commodity versus noncommodity exporters as well as landlocked versus other countries. These associations are presented in detail in World Bank and IMF (2011). Export sophistication and shipping connectivity are likely to be correlated with a country's level of development, growth performance, infrastructure, and policies and institutions for trade, private sector development, and doing business.

10. The average GDP per capita growth in IDA countries (1990–2009) is 1.36, one point below the average growth in non-IDA countries (2.38). The CPIA index in 2009 was, on average, 3.26 in IDA countries versus 3.69 in non-IDA countries. Fragile or conflict-affected countries (one or more years, 2006–09) exhibit average per capita GDP growth (1990–2009) close to 1.03 percent and a CPIA index of 3.00 in 2009. However, nonfragile states have grown, in per capita terms, at an average rate of 2.27 percent since 1990. The CPIA index for these countries was 3.68 in 2009.

11. The underlying indicators about development outcomes (like reduction in poverty) are also measured infrequently. For example, countries normally conduct household surveys of incomes and expenditures, the basis for measuring poverty, every three or five years, and in some cases, even ten years.

12. We also conducted pairwise correlations between the variation of the MDG-related indicators and the same list of factors. Although there were many good correlations (significant at the 10 percent level and correct signs), there were now more gaps in the matrix (for insignificant values or incorrect signs). This finding suggests that there are likely more factors that are associated with the variation of MDGs than could be accounted for by simple pairwise correlations, again confirming the general conclusions of various micro studies.

13. Go and Quijada (2011) discuss statistical issues relating to the estimation method, dependent variable, the independence of irrelevant alternatives, endogeneity and reverse causality, and multinomial versus ordered logit estimation.

14. We do not include variations of institutional variables in specification 17 because of the lack of time-series data going back to 1990. When such data are available, methodological inconsistency across periods is the major drawback.

15. For an alternative version, see Go and Quijada (2011), where growth, rather than income levels, is used as a development driver. The results are generally the same.

16. See World Bank (2012), for example, for a recent global outlook. World Bank and IMF (2010) also noted that developing countries generally did better than high-income countries from 2008 to 2009 and discussed the various reasons. However, developing countries are generally more vulnerable to an unfavorable outturn than they were in 2007. Although developing countries' fiscal positions and growth prospects are healthier than those of developed countries, they have generally less fiscal space (i.e., breadth, depth, and quality of economic resources) and weaker conditions than in 2007.

17. Dang, Knack, and Rogers (2009) found that aid flows from 1977 to 2007 fell by 20–25 percent on average from donor countries with banking crises, beyond any income-related effects.

References

Acemoglu, D. J. A. Robinson. 2012. *Why Nations Fail.* New York: Crown Publishers.

Arbache, J., D. Go, and J. Page. 2008. "Is Africa's Economy at a Turning Point?" *In Africa at a Turning Point? Growth, Aid and External Shocks*, ed. D. Page, and J. Go, 14–86. Washington, DC: World Bank.

Baldacci, E., B. Clements, S. Gupta, and Q. Cui. 2008. "Social Spending, Human Capital, and Growth in Developing Countries." *World Development* 36(8): 1317−41.

Bertelsmann Transformation Index. 2006. Available at http://bti2006.bertelsmann-transformation-index.de/.

Besley, T. T. Persson. 2011. *Pillars of Prosperity − The Political Economics of Development Clusters*. Princeton, NJ: Princeton University Press.

Bourguignon, F., A. Bénassy-Quéré, S. Dercon, A. Estache, J. W. Gunning, R. Kanbur, S. Kasen, S. Maxwell, and J. Platteau, and A. Spadaro. 2010. "The Millennium Development Goals: An Assessment." In R. Spence and M. Kanbur, eds., *Equity and Growth in a Globalizing World*. Commission on Growth and Development. Washington, DC: World Bank.

Clemens, M. A., C. J. Kenny, and T. J. Moss. 2007. "The Trouble with the MDGs: Confronting Expectations of Aid and Development Success." *World Development* 35(5): 735−51.

Collier, P., and S. A. O'Connell. 2006. "Opportunities and Choices." *Explaining African Economic Growth*, Chapter 2 of synthesis volume. African Economic Consortium, Nairobi, Kenya.

Dang, H. A., S. Knack, and H. Rogers. 2009. "International Aid and Financial Crises in Donor Countries." Policy Research Working Paper 5162. World Bank, Washington, DC.

Devarajan, S., and R. Reinikka. 2004. "Making Services Work for Poor People." *Journal of African Economies* 13(1): i142−i166.

Easterly, W. 2009. "How the Millennium Development Goals Are Unfair to Africa." *World Development* 37(1): 26−35.

Economist Intelligence Unit. 2007. Index of Democracy. Available at http://www.economist.com/media/pdf/DEMOCRACY_INDEX_2007_v3.pdf.

Filmer, D., J.S. Hammer, and L. Pritchett. 2000. "Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries." *World Bank Research Observer* 15(2): 199−224.

Freedom House. Available at http://www.freedomhouse.org.

Fukuda-Parr, S., and J. Greenstein. 2010. "How Should MDG Implementation Be Measured: Faster Progress or Meeting Targets?" Working Paper 63. International Policy Center for Inclusive Growth, Brasilia, Brazil.

Go, D., and J.A. Quijada. 2011. "Assessing the Odds of Achieving the MDGs." Policy Research Working Paper 5825. World Bank, Washington, D.C.

Günther, I. G. Fink. 2010. "Water, Sanitation and Children's Health: Evidence from 172 DHS Surveys." Policy Research Working Paper 5275. World Bank, Washington, DC.

Harttgen, K., and S. Klasen. 2010. "Fragility and MDG Progress: How Useful Is the Fragility Concept." Working paper 20. Robert Schuman Centre for Advanced Studies, European University Institute, Florence, Italy.

Hogan, M., K. Foreman, M. Naghavi, S. Ahn, M. Wang, S. Makela, A. Lopez, R. Luzana, and C. Murray. 2010. "Maternal Mortality for 181 Countries, 1980−2008: A Systematic Analysis of Progress towards Millennium Development Goal 5." *Lancet* 375(9726): 1609−23.

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2009. "Governance Matters VIII: Governance Indicators for 1996−2008." World Bank Policy Research June 2009, Washington, DC.

Knack, S., and M. Kugler. 2002. "Constructing an Index of Objective Indicators of Good Governance." PREM Public Sector Group, World Bank.

Lay, J. 2010. "MDG Achievements, Determinants and Resource Needs: What Has Been Learnt?" Policy Research Working Paper 5320, World Bank, Washington, DC.

Leo, B., and J. Barmeier. 2010. "Who Are the MDG Trailblazers? A New MDG Progress Index." Working Paper 222, Center for Global Development, Washington, DC.

Lofgren, H. 2010. "What Determines the Evolution of MDG Indicators? A Selective Review of the Literature." Unpublished manuscript, World Bank, Washington, DC.

Marshall, M., and B. Cole. 2010. *Global Report 2009: Conflict, Governance, and State Fragility.* Washington, DC: Center for Global Policy.

Ndulu, B. J. 2008. *The Political Economy of Economic Growth in Africa, 1960–2000.* Cambridge, U.K.: Cambridge University Press.

ODI (Overseas Development Institute). 2010. "Millennium Development Goals Report Card: Measuring Progress across Countries." London.

Rajkumar, A., and V. Swaroop. 2008. "Public Spending and Outcomes: Does Governance Matter?" *Journal of Development Economics* 86: 96–111.

United Nations. 2008. *Report of the Secretary-General on the Indicators for Monitoring the Millennium Development Goals.* E/CN.3/2008/29. New York.

Wagstaff, A., and M. Claeson 2004. The Millennium Development Goals for Health: Rising to the Challenges. Washington, DC: World Bank.

World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People.* Washington, DC: World Bank.

⸻. 2007. "Selectivity and Performance: IDA's Country Assessment and Development Effectiveness." An IDA 15 document prepapted by IDA and DECVP (International Development Association and Development Economics, Office of the Chief Economist). Washington, DC.

⸻. 2012. *Global Economic Prospects: Uncertainties and Vulnerabilities.* Vol. 4, January. Washington, DC: World Bank. http://www.worldbank.org/prospects.

World Bank and IMF. 2012. *Global Monitoring Report 2012: Food, Nutrition, and the Millennium Development Goals.* Washington, DC: World Bank. http://www.worldbank.org/gmr20112.

World Bank. 2011. *Global Monitoring Report 2011: Assessing the Odds of Achieving the MDGs.* Washington, DC: World Bank. http://www.worldbank.org/gmr2011.

⸻. 2010. *Global Monitoring Report 2010: The MDGs after the Crisis.* Washington, DC: World Bank. http://www.worldbank.org/gmr2010.

World Development Indicators, World Bank. Available at http://data.worldbank.org.

The Challenges of Bankruptcy Reform

Elena Cirmizi • Leora Klapper • Mahesh Uttamchandani

The 2008 financial crisis was followed by a global economic downturn, a credit crunch, and a reduction in cross-border lending, trade finance, and foreign direct investment, which adversely affected businesses around the world. The consequent increase in the number of firm insolvencies in the corporate sector highlights the need for commercial bankruptcy laws to liquidate efficiently unviable firms and reorganize viable ones, so as to maximize the total value of proceeds received by creditors, shareholders, employees, and other stakeholders. The authors summarize the theoretical and empirical literature on bankruptcy design, discuss the challenges of introducing and implementing bankruptcy reforms, and present examples of how policymakers are trying to take advantage of the current economic downturn as an opportunity to engage in meaningful reform of the bankruptcy process. They also review the main principles of efficient insolvency laws and bankruptcy procedures. JEL codes: G33, G38, G38, K40, O16

The 2008 global financial crisis led to an increased risk of insolvency among firms worldwide due to declining demand for goods and services, decreasing availability of external finance, declining investments, and reductions in remittances. As in previous financial crises in Russia, East Asia, and Argentina, policymakers have responded in part by shifting their attention to the effectiveness of current bankruptcy laws and their role as a key mechanism in addressing widespread firm-level financial distress (Claessens, Djankov, and Klapper 2003). Specifically policymakers have engaged in reform efforts to improve the structure of current reorganization and liquidation mandates, and the ability of existing court systems to enforce these laws in court.

In developing a legal framework for the efficient resolution of insolvency, the essential challenge is to incentivize (i) the reorganization of viable firms and (ii) the liquidation of unviable ones at a low cost. Effective bankruptcy laws recognize

© The Author 2011. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

that keeping viable businesses alive is the most efficient outcome for creditors, employees, and a firm's network of suppliers. Less successful companies, however, should ideally be taken over by more capable owners or liquidated through asset sales, so that only the most efficient users of economic resources continue to operate as active companies.

In its design features, effective insolvency laws also try to balance the interests of the parties involved to ensure an equitable resolution to the matter at hand without discouraging future risk-taking by investors and entrepreneurs. First, insolvency laws should include adequate protections of the rights of creditors and other stakeholders because, at the most basic level, the ability of creditors to provide start-up capital, working capital, and continued investment to entrepreneurs is essential to the dynamism of the market economy (see Hart 2000 and Stiglitz 2002 for reviews). Frequently the liquidation of assets and the distribution of capital among creditors develop into a collective action problem among managers, employees, creditors, and suppliers. For instance, when a debtor becomes insolvent, creditors have incentives to engage in a "run on assets," enforcing their individual claims, and possibly liquidation, as quickly as possible, even if the result is a reduction in the overall value obtained. To prevent this scenario from occurring, effective bankruptcy laws provide a mandatory and orderly mechanism for the coordination of the reallocation of assets of insolvent firms among stakeholders (Jackson 1982). Well-developed mechanisms for asset recovery also ensure that entrepreneurs will be willing to engage in new ventures and take risks in the future. Lastly a goal of designing bankruptcy laws that deliver an ex post efficient outcome is to increase the aggregate return to stakeholders, in the sense that the highest total value is obtained for the distressed firm.

The working of a country's judicial system plays a nontrivial role in balancing the interests of those involved in a bankruptcy proceeding. Beyond the clear enumeration of equitable legal rights, there is a need for an efficient judicial system to enforce these rights, or at least to serve as a credible threat (see Modigliani and Perotti 2000). In reality, however, courts and the judges often act as an impediment to the efficient resolution of insolvency and are frequently the focus of bankruptcy reforms.

More generally, the resolution of bankruptcy depends greatly on the broad institutional context within which firms in specific countries operate (Scott 1995). Despite the frequency of insolvency and firm closure, the use of legal procedures associated with in-court bankruptcy vary significantly around the world, due to differences in legal traditions, accounting standards, regulatory frameworks, and macroeconomic factors (Claessens and Klapper 2005). For instance, formal bankruptcies (within the courts) are less common in countries with concentrated banking systems and among firms with single banking relationships, and are more common in firms with more complex capital structures (Bebchuck 1988). Furthermore the

laws in some countries only allow for the liquidation, but not the restructuring, of insolvent firms and provide limited protection for entrepreneurs and managers of bankrupt firms. Therefore owners might be forced to liquidate their assets even when the firm is viable. Other countries have more bankruptcy options (such as reorganization and out-of-court mediation), though the effectiveness of these laws in practice varies across countries (reviewed by Lee, Peng, and Barney 2007).

We focus on only the legal framework for the resolution of nonfinancial firm insolvency because, in most developed countries, bank, insurance, and "financial firm" insolvencies are generally conducted under separate legal regimes. Indeed international best practice (reflected in the World Bank Principles for Effective Insolvency and Creditor Rights Systems [World Bank 2005] and the UNCITRAL Legislative Guide on Insolvency Law [UNCITRAL 2005]) indicates that the significant public policy and financial stability considerations inherent in such insolvencies often renders them best dealt with outside of the traditional commercial insolvency framework.

In this paper we summarize the key features of well-functioning bankruptcy laws and the importance of strong insolvency laws for private sector development and growth. We then review the impact of the recent financial crisis on the resolution of bankruptcy and analyze some recent bankruptcy law reforms from around the world.

Features of Well-functioning Insolvency Regimes

Effective insolvency regime promote economic growth and competition by allowing economies to take timely action in cases of debtor's default and nonperformance. In the aftermath of the financial crisis of the late 1990s and great recession, it is widely recognized that sound insolvency systems constitute one of the main areas integral for sound financial systems and financial stability. There is a number of comprehensive best practice features to be used as a guideline for creating sound insolvency regimes or as a benchmark for assessing existing ones.

Insolvency Law Design

In many countries, the existing bankruptcy laws do not efficiently address and resolve the issues brought forth by firm-level insolvency (Djankov and others 2008). A survey on debt enforcement of practitioners from 88 countries indicates that bankruptcy procedures are time-consuming, costly, and inefficient (that is, they are unable to preserve the business as a going concern). In only 36 percent of countries can an insolvent business be preserved as a going concern, and an average of 48 percent of an insolvent business's value is lost in debt enforcement. In a well-functioning bankruptcy system, the laws would ensure that the highest

total value is achieved for the distressed firm. In other words, whether the firm should be closed down, liquidated piecemeal, sold as a going concern, or reorganized should depend on which option maximizes the total value of proceeds received by creditors, shareholders, employees, and other stakeholders.

In addition to bankruptcy laws, speedier court resolutions can also reduce uncertainty for entrepreneurs, creditors, and management, and improve assets value and transparency. Actions that expedite court procedures include minimizing dependence on the courts (through the appointment of a receiver for distressed companies, as for example in Georgia), establishing special courts (for example in India, Thailand, Indonesia, and Uganda), and limiting appeals and introducing time limits (for example in Tajikistan and Lithuania).

There are several important principles that underlie the design of a good bankruptcy regime:

- Ensure equitable treatment of similarly situated creditors, recognize existing creditors' rights, and establish clear rules for ranking priority claims (UNCITRAL 2005).
- Maximize the value of assets and preserve the insolvency estate to allow equitable distribution to creditors. For instance, the recovery rate varies among economies from 4.4 cents on the dollar claimants in the Philippines to 92.5 cents in Japan (World Bank 2010).
- Preserve some portion of firm value for shareholders, even in bankruptcy. Otherwise shareholders may do anything to prevent bankruptcy, including undertaking high-risk projects when the corporation is under distress (Hart 2000).
- Provide for timely resolution of insolvency. For instance, Ireland provides the fastest bankruptcy procedure—less than four months—whereas in many developing countries the process takes many years, for example eight in Mauritania and seven in India (World Bank 2010).

At the same time, ensuring the right incentive structure is critical. A primary objective of insolvency laws is to prevent disorderly and discriminatory individual grabs by protecting creditors' equality and ensuring that the proceeds of the debtor's assets are divided between the creditors according to the bankruptcy law's hierarchy of payment. Well-functioning insolvency laws achieve this goal in a variety of ways, although a review of leading bankruptcy laws around the world yields a number of common elements:

- *Establishing a single, clear hierarchy of payments.* The order of various priorities should be precise, transparent, and easy to understand. This not only allows creditors to realize, with some degree of comfort, their relative priority, but

also allows a presiding court to determine clearly which parties' economic interests are truly at stake in an insolvency proceeding and, consequently, whose interests should be safeguarded.

- *Providing for the immediate transfer of a failed reorganization into liquidation.* While it is an accepted principle of insolvency law that the law should provide for a balance between liquidation and reorganization, one of the primary incentives for a creditor to participate in a reorganization process is the understanding that, should that process fail, no additional steps or processes will be required to ensure that the court or insolvency administrator remains in control of the insolvent estate and that the assets in question will remain under a continuing conservatorship for the benefit of the creditors.

- *Allowing creditors to play a large role in the insolvency, but not to manipulate the process.* In both dominant paradigms of modern insolvency laws (administrator-led and debtor-in-possession), creditors play a large role in the insolvency through the use of creditor committees or de facto control of the administrator. While the presiding court must ensure that the interests of all stakeholders are protected, one of the key incentives for creditors to participate in the collective process of bankruptcy (rather than rush to individual enforcement) is the comfort that they, as a group, will exert some control over that process. This includes having clear voting requirements for the approval of any plan of reorganization that appropriately divides creditors into classes, for the purposes of voting, based on shared economic interests.

- *Balancing certainty and flexibility.* This may be the most difficult element to achieve, but it is a key feature of leading legislation. For the reasons noted above, creditors will require a level of certainty to incentivize their participation in a reorganization process that could otherwise be regarded as needlessly delaying the enforcement of their rights. This will include, wherever practicable, preserving prebankruptcy rights and priorities inside of bankruptcy. At the same time, without a measure of flexibility (such as the ability to authorize post-commencement priority financing, often at high margins), it will be difficult to craft a workable reorganization plan that serves the broad interests of all stakeholders. This flexibility–certainty balance usually requires both a well-designed law and a highly competent cadre of judges who are able to determine where the law's flexibility can most appropriately be applied, without unduly compromising the rights of stakeholders.

To Allow Risk Taking. Historically the earlier introduction of bankruptcy codes in England and in the United States may have supported the more dynamic private sector entry and exit seen in those countries (Di Martino 2002). Conversely, in Italy and France, the commercial code introduced by Napoleon in 1807 reinforced the severity and the penal character of medieval legislation that discouraged firm

failures (Bignon and Sgard 2007). Qualitative evidence suggests that in England the devices and instruments provided by legislators were more effective than Italian equivalents in attracting a larger number and higher quality of new entrepreneurs. Furthermore quantitative evidence shows that English procedures assured creditors higher dividends and a shorter waiting-time than in Italy.

Since the ease of bankruptcy determines the maximum downside risk of a venture, only high-risk entrepreneurs will be willing to make significant investments in start-ups in countries with unfriendly bankruptcy regimes. Thus entrepreneurship is encouraged by limiting downside risks and increasing upside gains, leading to an increase in and the number and variety of people pursuing entrepreneurial activities (Lee, Peng, and Barney 2007). Indeed data from a Eurobarometer survey show that the fear of bankruptcy is one of the most important reasons given by individuals for not forming their own businesses, although the extent of the deterrent effect varies with the quality of bankruptcy laws and other features of the business environment across countries (Armour and Cumming 2007).

To Promote Macroeconomic Growth. The Schumpeterian theory of "creative destruction" contends that firm exit is a necessary condition for economic growth: when innovative activity in an industry increases, firms' overall survival rates often decrease, but those that do survive tend to be stronger. Research based on firm-level data supports this assertion that the continuous process of reallocation of resources plays an important role for aggregate productivity and output growth (for example Bartelsman, Haltiwanger, and Scarpetta 2009). Resource reallocation is driven by incumbent firms adapting to market and technological changes, but also by firm dynamics—the entry of new firms, their expansion in the initial years of life, and the exit of weak or obsolete firms. This important relationship between entry, exit, and growth has been examined in both the management and finance literature and supported empirically using firm-level data (for example Porter 1990; Audretsch 1991; Nickell 1996; Klapper, Laeven, and Rajan 2006).

For instance, longitudinal firm-level sector data in the United States shows a tremendous reallocation of activity across service-sector firms, which has been generated by firm turnover. For example, the exit of very low productivity plants was the primary contributor to the productivity growth of the automobile repair shop industry between 1987 and 1992 (Foster, Haltiwanger, and Krizian 1998). Moreover plant-level data for Colombia finds that market reforms are associated with rising overall productivity that is primarily driven by reallocation away from low- and toward high-productivity businesses (Eslava and others 2004). An efficient economy innovates quickly; but when the economy is unable to redeploy resources away from inefficient uses, technological adoption becomes sluggish and growth is reduced (Bergoeing, Loayza, and Piguillem 2010).

Yet the efficient reallocation of capital depends on strong insolvency laws that ensure a quick and low-cost resolution of financial distress. For instance, bankruptcy reforms in South Korea after the 1997 economic crisis were found to contribute to productivity growth by allowing inefficient firms to exit, encouraging new entries, and stimulating competition among surviving firms to become more efficient (Lim and Hahn 2003).

The Importance of Legal and Judicial Efficiency

It is important for countries to strike the right balance between the protection of creditor and shareholder rights. On the one hand they need to ensure that banks and other creditors receive the highest total value in the sale or liquidation of a distressed firm, and on the other hand they must protect shareholders' interests by identifying and reorganizing viable enterprises.

To encourage greater rehabilitation of distressed firms, some countries choose a more debtor friendly regime. However, debtor-friendly laws along with weak courts might incentivize each creditor to collect outstanding debt privately before other creditors, even though coordinated liquidation would maximize the total returns to the creditors as a group. For instance, a first-come, first-served ordering of creditors' claims might cause a firm to be sold in an ad hoc approach; such is the case in Egypt, and indeed many Middle East and North African countries, where, even while a company may be attempting to reorganize, there is no comprehensive stay of proceedings against creditors enforcing their rights, and it is extremely difficult for the debtor to seek additional financing during the restructuring. In contrast, under Chapter 11 of the U.S. Bankruptcy Code, not only is there such a stay on assets during bankruptcy proceedings, but the debtor remains in possession and control of its company throughout the restructuring process and can even seek super-priority additional financing, ahead of prior creditors (with court approval), to finance its restructuring.

In comparison, efficient bankruptcy courts give order to the sales and distribution of assets of insolvent firms and can positively affect loan terms (such as spreads, rates, and collateral requirements), leverage ratios, and bank recovery rates (Davydenko and Franks 2008; Acharya, Rangarajan, and Kose 2008). For instance, the introduction of Debt Recovery Tribunals in India reduced delinquency in loan repayment rates by between 3 and 11 percent and interest rates fell by up to 2 percentage points (Visaria 2009).

Additional Challenges for Insolvent Firms

In Middle-income Countries. As middle-income countries undertake reforms to "catch up" their insolvency systems with international best practice, they will

also have to consider addressing some of the specific challenges endemic to developing countries. For example: issues such as the treatment of corporate groups (where an enterprise consists of two or more legal entities), which even the most advanced insolvency laws do not contemplate; and the treatment of insolvencies that span two or more jurisdictions, thereby creating complex issues of asset recovery, jurisdiction, and regulatory oversight (Uttamchandani 2008). While these issues may be overly complex for small undeveloped economies, India, China, Turkey, and other middle-income countries wrestling with basic questions may also have to tackle these more challenging issues at the same time.

In particular, given the dramatic increase in foreign direct investment in these countries over the past few years, large insolvencies will increasingly have transnational dimensions. This will put the countries' insolvency systems into direct contact with systems from advanced countries, necessitating clear rules of engagement. In Europe, for example, the Parmalat 2003 bankruptcy case underscored the need for courts in multiple jurisdictions to coordinate efforts in winding up an insolvent estate and for insolvency administrators appointed in one jurisdiction to have clear guidelines under which they could seek recognition from local courts in other jurisdictions. Because of the increased integration that some middle-income countries have achieved with the global economy, they will no longer have the luxury of limiting their insolvency regimes to purely domestic considerations.

In Labor-intensive Firms. There has been little empirical work addressing the specific challenges of labor-intensive firms in the bankruptcy process, despite the growing importance of service-oriented firms. For instance, how does the relationship between a firm and its employees affect the choices made by an insolvent firm? Indeed bankrupt firms routinely cite employee retention as a critical concern (Berkovitch, Israel, and Zender 1997; Berk, Stanton, and Zechner 2010).

The literature suggests that the process of corporate bankruptcy varies by labor intensity (Wang 2009). First, labor-intensive firms increase their leverage more sharply prior to bankruptcy compared with capital intensive firms, relying on borrowing to finance firm growth instead of undertaking typical restructuring activities. Second, labor-intensive firms are more likely to be liquidated during the bankruptcy process. However, among firms that emerge and remain publicly listed, those with above-median human capital share are 14 percent less likely to refile for bankruptcy within five years of emergence (Wang 2009). These findings have implications for both the capital structure decisions of labor-intensive firms and the effectiveness of asset reallocation in bankruptcy.

The first explanation for this phenomenon may be the idea that labor-intensive firms are more vulnerable to the departure of valuable employees during bankruptcy; this might explain creditors' willingness to continue lending to human-

capital-intensive firms prior to bankruptcy. In addition, labor intensive firms are also highly redeployable, consisting in some cases of little beyond real estate and office equipment, thus they would be less likely to suffer from fire-sale discounts during liquidation.

In Small and Medium-sized Enterprises (SMEs). Most literature considers the importance of bankruptcy codes in addressing the needs of creditors that lend to large, capital intensive firms. However, good bankruptcy systems can also be important for smaller firms. For instance, 80 percent of U.S. firms that filed for bankruptcy reported assets under $1 million, and 88 percent reported having fewer than 20 employees (Warren and Westbrook 1999). In addition, SMEs are especially vulnerable to macroeconomic and financial shocks; for example, SME insolvencies in Denmark, Italy, Spain, and Ireland exceeded 25 percent between 2007 and 2008 (OECD 2009).

Reforms that reduce the time and cost of reorganization (relative to liquidation) appear to be particularly important for smaller firms, relative to larger firms. For example, a study of a reform in Belgium to encourage corporate reorganization and reduce liquidation rates finds a significant decline in micro- and small business failure rates (though the study does not examine large firms). The reform encouraged small firms to reorganize instead of liquidate; and the liquidation rate of partnerships in bankruptcy fell by an annual average of 8.4 percent (Dewaelheyns and Van Hulle 2006). A study of a reform in Brazil to simplify the reorganization of insolvent firms (with a necessary caveat that the paper does not prove causality) finds a relatively larger effect of the reform on the cost and access to debt by smaller firms (Funchal 2008). These studies suggest that creditors concerned about the cost of bankruptcy relative to the size of the estate are less likely to liquidate firms in countries with more efficient reorganization procedures.

Lenders to small businesses often require that the owner provides a personal guarantee to the loan, such as a second mortgage on his or her house (Berkowitz and White 2004; Djankov and others 2008). Guarantees of this sort put the personal assets of the firm's owner on the line and blur the distinction between the assets of the firm and those of the owner; in other words, the limited liability of the firm no longer applies to this particular loan. In addition, personal bankruptcy laws would apply to this firm in the case of default. A survey of a sample of individuals from the United States who filed for bankruptcy during the 1980s estimates that around 20 percent had debts from a failed business (Sullivan, Warren, and Westbrook 1989). While the personal guarantee of a firm's owner might encourage a level of financial discipline, in countries without a personal bankruptcy framework a single business failure could doom an owner to a lifetime of outstanding debt (Uttamchandani and Menezes 2010) and effectively prevent them from re-entering the market as seasoned entrepreneurs (Armour and Cumming 2005).

The Response of Insolvency Laws to Financial Crisis

The 2008 global financial crisis caused a sharp increase in the number of insolvencies around the world. For example, during 2009, the number of corporate bankruptcies in Japan was 13,306, up 4.9 percent from 12,681 in 2008 (Teikoku Databank 2010); in Great Britain 94,135, a 5.88 percent increase compared to 2008 (Ministry of Justice 2010); and in Germany the number of corporate bankruptcies was 32,687, which represents an 11 percent annual increase (Statistisches Jahrbuch 2009). In 2009, 60,837 businesses in the United States declared bankruptcy, representing a 40 percent increase in filings from 2008 (American Bankruptcy Institute 2010).

An important feature of bankruptcy laws is disentangling unviable firms that should be liquidated from those that are viable, but suffering from insufficient access to credit or temporary drops in demand. During a crisis, ensuring that viable companies can continue to operate as going concerns, and preserving jobs, becomes especially important. The bankruptcy processes—which are often already under strain during normal times—can be completely overwhelmed (Demirgüç-Kunt 2009). In response, policymakers are debating whether existing bankruptcy regimes adequately address current business demands.

The World Bank's Financial Crisis Survey shows that following the 2008 financial crisis, the use of bankruptcy procedures in five central European countries was less frequent than the use of state aid and debt restructuring. On average, 8.3 percent of European firms applied for state aid in the previous 12 months (as of July, 2009), whereas only 2 percent of all surveyed companies filed for bankruptcy, though the figure was 6 percent of firms with overdue payments (Correa and Iootty 2010). Many researchers have expressed concern that the costs of direct intervention by governments, such as giving assistance to individual companies, comes at a significant fiscal obligation to taxpayers. In addition, it prevents meaningful restructuring, encourages other private sector firms to expect similar assistance, gives incentive for imprudent risk-taking incentives, and paves the way for more frequent and costlier crises in the future (Caprio, Demirgüç-Kunt, and Kane 2008; Demirgüç-Kunt and Levine 2008; Demirgüç-Kunt and Servén 2009).

Policymakers have drawn important lessons from the 1997 East Asia Financial Crisis in which several countries reformed their corporate insolvency laws when existing bankruptcy systems did not allow the corporate sector to rehabilitate during the long term economic recession (Armour and Deakin 2001). When illiquidity spread across the region, South Korea, Malaysia, and Thailand were forced to modify their laws to favor the reorganization of distressed firms, as an alternative to liquidation, including provisions to the laws that added incentives for creditors and debtors to negotiate (Carruthers and Halliday 2007). Indonesia

and Thailand also introduced specialized courts to implement bankruptcy procedures (Claessens, Djankov, and Xu 2000).

For example, the revisions to Indonesia's bankruptcy law included: new procedural rules designed to ensure that bankruptcy proceedings would be transparent; provisions that allowed for the appointment of receivers and administrators from the private sector to administer the estates of debtors; greater protection of debtors' assets, including protection against insider and fraudulent transactions; and limitations on the ability of secured creditors to foreclose on collateral during the proceedings, thus making reorganizations more likely. The new laws provided important incentives for both creditors and debtors to negotiate out-of-court, as well as providing a useful means by which debtors could bind dissenting creditors to a restructuring plan that received support from the requisite majority of creditors (Iskander and others 1999).

More recently, in 2009, Germany revisited its long-standing rule requiring company management to file for bankruptcy in certain situations, or face imprisonment. While this rule was originally instituted to ensure a level of debtor discipline and creditor confidence, the financial crisis prompted a fear amongst some policymakers that declining asset values would create widespread de facto balance-sheet insolvencies and prompt managers to put otherwise viable companies into insolvency proceedings. As a result, the filing requirement was amended to be less stringent. The current law allows companies to continue to operate in the case of over-indebtedness (Meier, Michael, and Schauenburg 2010). In addition, prior to the Asian financial crisis, Thai judicial procedures were fraught with large transactions costs: bankruptcy cases dragged on for more than two years on average, and there was no specialized court to implement expedited procedures. The law lacked financing and an automatic stay provisions for debtors in possession to protect assets. The law also did not explain how creditors and managers should prepare and implement a restructuring plan. A study of Thai firms (with the important caveat that the author could not disentangle the impact of the economic recovery) shows that, as a result of bankruptcy reforms in Thailand, both creditors and debtors experienced immediate financial gains from the new bankruptcy procedures (Foley 2001).

Examples from an earlier crisis in Latin America show that, by restructuring viable businesses and quickly liquidating nonviable ones, well-functioning bankruptcy regimes can reallocate and remobilize resources, thus speeding up the recovery from the crisis. For example, a comparative study of Chile and Mexico suggests that the decade-long divergent growth paths of the two countries since the financial crisis in the early 1980s are predominantly driven by differences in total factor productivity growth rates and laws that facilitate the entry and exit of firms (Bergoeing and others 2001). For instance, an explanation for Chile's relatively quick recovery from a deep recession in the early 1980s is that the country

reformed its bankruptcy law to allow firms to fail, while Mexico was unwilling to let inefficient firms go bankrupt. Similarly, during the financial crisis that spread across Latin America, Colombia introduced bankruptcy reform in 1999 to improve the efficiency of the bankruptcy process by streamlining reorganization proceedings. This reform improved the efficiency of the resolution of distress, leading to a significant improvement in the selection of viable firms into reorganization and a significant decrease in the duration of reorganization (Gine and Love 2008).

During financial crises, countries have also introduced new mechanisms to reduce the costs of reorganization by making it possible for an out-of-court system to circumvent the formal judicial process and its attendant costs. For example, the Mexican and East Asian crises spurred the introduction of arbitration rules—the "London rules" or "prepackaged" bankruptcies—which encouraged all creditors to sign an out-of-court agreement reached among the majority of creditors prior to the bankruptcy filing, which allows distressed firms to avoid lengthy and costly court procedures. This instrument is being used again during the current crisis; for instance, Italy now allows a distressed company to seek an agreement with creditors before filing for bankruptcy, which permits it to continue operating, and provides the possibility of paying secured creditors less than the full amount of debt. Prior to the reform, insolvency procedures were predominantly aimed at liquidating insolvent enterprises, while after the reform the law focused on efficient prebankruptcy procedures and reorganization (Novarese 2009).

Similarly the reform of the bankruptcy law in France improved the insolvency process by encouraging pre-insolvency workouts and by ending the requirement that a public auctioneer had to estimate the value of a firm's assets (World Bank 2009). While it remains too soon to judge how successful the new legislation will ultimately prove to be, early indications are that it has been successful in reducing the number of liquidations (Lucheux and Pusch 2009).

Miller and Stiglitz (1999) and Stiglitz (2002) proposed the use of "super-bankruptcy" to enhance recovery and provide protection against large macroeconomic shocks by keeping existing management in place and forcing debt-to-equity conversions. The super-bankruptcy mechanism aims to prevent liquidations that occur only as the result of a system-wide crisis by not punishing existing managers who become "victims" of external macroeconomic shocks. For instance, studies of distressed firms in the United States (Gilson 1989; Gilson and Vetsuypens 1994) find that after a bankruptcy filing managers receive significantly lower salaries and bonuses (on average, managers receive only 35 percent of their previous gross income), and more than half of the sampled managers are fired. The downside of such a policy is the moral hazard of protecting firm managers and owners (who caused the problems in the first place) and the incentive it gives to creditors to charge an interest rate premium in normal times because

their loan is at risk during business cycle downturns (Demirgüç-Kunt 2009). Evidence from East Asia suggests that adopting a temporary super-bankruptcy is unnecessary—corporations and banks avoided restructuring outstanding debt, in the hope that an economic recovery would preclude the need for write-offs (for banks) or the surrendering of equity control (for large shareholders) (Claessens, Djankov, and Xu 2000; Claessens, Djankov, and Klapper 2003).

During normal times, proposed changes to bankruptcy laws might face opposition from judges, administrators, and lawyers resistant to significant institutional changes (Djankov 2009). However, during financial crises, policymakers might be forced to addresses weaknesses in their insolvency codes in response to an increase in loan defaults and business closures. As the main goals of insolvency reforms enacted in times of crisis are to improve economic efficiencies and strengthen market resilience, the most popular trends among current reformers include the following.

Establishing Reorganization Procedures or Prepackaged Arrangements to Enable Viable Firms to Continue as Going Concerns

This occurs in, for example, Italy, Kuwait, the Czech Republic, Poland, Estonia, Mauritius, Uruguay, Rwanda, Sierra Leone, the Philippines, and France. For instance, Poland amended its bankruptcy law to introduce a "prepackaged" reorganization, which permits filings by either the insolvent firm's board of directors or by the creditors (World Bank 2010). Prepackaged reorganizations were also instituted by the new Insolvency Act of the Czech Republic and by the Estonian Restructuring Act. The first introduced reorganization as the preferred method for dealing with insolvency and established an electronic insolvency register (Osicka, Kucerova, and Mestanek 2008). The Estonian reform, which was modeled on the U.S. Chapter 11 approach, as well as on the German Insolvenzordnung and the Finnish Saneerauslaki, is designed to help financially troubled firms avoid liquidation and to optimize the possibility of retaining their reputation and the trust of their creditors. Creditors have found the new Act attractive because it offers them a clear nonbankruptcy means of maximizing the amount they are able to collect from a debtor, and it encourages them to purchase debt or equity in financially distressed firms (United States Department of Commerce 2010).

A slightly different approach was chosen by Uruguay, where the new law consolidated all the different procedures existing prior to the enactment of the new law in just one unique procedure called "Concurso." The new law aims to encourage companies to disclose financial difficulties in a prompt manner in order to facilitate direct agreement between debtors and creditors and to preserve viable firms (Garcia 2008). In comparison, Bolivia suspended accepting applications for voluntary restructuring. While this reform was aimed at preventing viable businesses

from exiting the market, the result was that many distressed companies that otherwise might have been able to recover were forced into a long liquidation process.

Introduction of Shorter Time Limits on Bankruptcy Procedures

This occurs in Italy, Lithuania, and Tajikistan. For instance, the Republic of Tajikistan introduced a new bankruptcy law, which streamlined the bankruptcy process and reduced the time required for closing a business from three to two years. The reform is expected to decrease the cost of bankruptcy from 9 to 2 percent of total asset values and increase the ratio of funds recovered for investors from 25.4 to 35.0 cents per U.S. dollar (USAID 2009). In Lithuania, reforms to commercial bankruptcy laws reduced the three-month wait-period for creditors to initiate bankruptcy proceedings to a 30-day grace period. During the first half-year of 2009, bankruptcy procedures were initiated for 936 enterprises, which is 55 percent more than at the same time in 2008 (prior to the financial crisis). On the other hand, at the beginning of the crisis, Italy enabled debtors to pursue immediate asset disposal plans (EStandards Forum 2010). Thus the time necessary for creditors to recover assets was significantly shortened and business restructurings were simplified.

Introducing Professional Requirements for Bankruptcy Administrators and Limiting the Payments they Are Permitted to Receive

This occurs in Albania, Colombia, Malawi, Lithuania, and Russia. These administrators play essential roles in insolvency procedures by taking part in managing insolvent companies and selling the assets of nonviable ones. For example, Colombia, Russia, and Albania introduced licensing requirements for bankruptcy receivers and training courses to improve professional qualification standards, aiming to reduce corruption among the bankruptcy administrators and debtors. Trying to achieve the same goals, Lithuania now sets higher standards of responsibility for persons executing bankruptcy procedures in order to prevent directors or owners from unfairly selling or hiding assets of a bankrupted company (World Bank 2009).

Several countries have focused on reducing corruption among administrators by limiting the amount of payments they can receive from assisting with the recovery of assets. In Malawi, for instance, the new Companies Regulation that took effect in June 2009 has made the mechanism for payment of liquidators more transparent. The new regulation sets a cap of 5 percent of the value of the estate on the liquidator's fees. Before, liquidators had the discretion to set their own fees, usually at around 10 percent of the value of the estate. Pursuing the same means, Romania amended its insolvency law to require 1.5 percent of the

amount recovered from each insolvency procedure to go to a fund for reimbursing the expenses of insolvency administrators (World Bank 2010). The aim was to ensure that insolvency administrators are paid even when debtors have no assets. However, the reform put additional constraints on closing businesses.

The financial crisis has also forced many legislators to take a fresh look at their bankruptcy codes. For instance, in May, 2009, in Abu Dhabi, representatives from 11 Middle East and North Africa jurisdictions (Egypt, Jordan, Lebanon, Libya, Oman, Qatar, Saudi Arabia, Sudan, the United Arab Emirates, the West Bank, and Gaza) established a dialog to reinforce insolvency laws and signed a joint declaration on intended reform in the region. Countries agreed to set up public–private partnerships to unify their insolvency laws; the insolvency laws in operation within the Dubai International Financial Centre have been proposed as a basis for the unification (Saidi 2009).

Conclusion

The 2008 global economic downturn highlighted once again that the effectiveness of insolvency laws has a profound effect on corporate and financial relationships and transactions among entrepreneurs, and is a powerful indicator of the impact of the legal system on commercial activities. As governments and policymakers use the current recession as an opportunity to engage in meaningful reform of the bankruptcy process, it is critical to examine and draw lessons from previous experiences.

Strong insolvency laws should ensure a quick and low-cost resolution of financial distress by incentivizing the liquidation of unviable firms and the restructuring of firms that are viable but suffering from insufficient access to credit or temporary drops in demand. Well-functioning bankruptcy laws should ensure that the resolution of financial distress maximizes the total value received by creditors, shareholders, employees, and other stakeholders. Yet insolvency laws can only function in a supportive and efficient judicial environment; speedier court resolutions can also reduce uncertainty for entrepreneurs, creditors and management, and improve assets value and transparency.

Following the 2008 financial crisis, the threat of widespread insolvencies in the financial and corporate sectors forced governments to start major reforms to improve their insolvency laws. Three popular trends emerged among insolvency reformers: (i) the establishment of reorganization procedures or prepackaged arrangements to enable viable firms to continue as going concerns; (ii) the introduction of shorter time limits on bankruptcy procedures; and (iii) the introduction of professional requirements for bankruptcy administrators. During a crisis, ensuring that viable companies can continue to operate as going concerns, and

preserving jobs, become especially important. Many new laws also address the growing complexity in insolvency caused by the rapid increase in credit and leverage in firms around the world and the greater diversity of creditors and shareholders. However, much more work must be carried out in order for many emerging markets to provide a quick, transparent, and efficient process to resolve financial distress.

Notes

References

The word *processed* describes informally reproduced works that may not be commonly available through libraries.

Acharya, Viral, Sundaram Rangarajan, and John Kose. 2008. "Cross Country Variations in Capital Structure: The Role of Bankruptcy Codes." AFA 2005 Philadelphia Meetings, Tuck Contemporary Corporate Finance Issues III Conference Paper.

American Bankruptcy Institute. 2010. "Annual U.S. Bankruptcy Filings."

Armour, John, and Douglas Cumming. 2005. "Bankruptcy Law and Entrepreneurship." ESRC Centre for Business Research Working Paper 300.

———. 2007. "Bankruptcy Law and Entrepreneurship." ECGI Law Working Paper 105/2008; University of Cambridge Centre for Business Research Working Paper 300.

Armour, John, and Simon Deakin. 2001. "Norms in Private Insolvency: The 'London Approach' to the Resolution of Financial Distress." *Journal of Corporate Law Studies* 1:21–51.

Audretsch, David. 1991. "New Firm Survival and the Technological Regime." *Review of Economics and Statistics* 73:520–6.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2009. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." National Bureau of Economic Research Working Paper 15490.

Bebchuk, Lucian. 1988. "A New Approach to Corporate Reorganizations." *Harvard Law Review* 101: 775–804.

Bergoeing, Raphael, Norman Loayza, and Facundo Piguillem. 2010. "Why Are Developing Countries So Slow in Adopting New Technologies? The Aggregate and Complementary Impact of Micro Distortions." National Bureau of Economic Research Working Paper 5393.

Bergoeing, Raphael, Patrick Kehoe, Timothy Kehoe, and Raimundo Soto. 2001. "A Decade Lost and Found: Mexico and Chile in the 1980s." Federal Reserve Bank of Minneapolis Staff Report 292. *Review of Economic Dynamics* 5(1): 166–205.

Berk, Jonathan, Richard Stanton, and Josef Zechner. 2010. "Human Capital, Bankruptcy and Capital Structure." *The Journal of Finance* 65:891–926.

Berkovitch, Elazar, Ronan Israel, and Jaime Zender. 1997. "Optimal Bankruptcy Law and Firm-Specific Investments." European Economic Review *41(3–5)*: 487–97.

Berkowitz, Jeremy, and Michele White. 2004. "Bankruptcy and Small Firms' Access to Credit." *Rand Journal of Economics* 35:69–84.

Bignon, Vincent, and Jerome Sgard. 2007. "The Two Uses of Bankruptcy Law in 19th Century France: Dealing With the Poor and Restructuring Capital." GEHN conference on Law and Economic Development, Utrecht, September 20–22, 2007.

Caprio, Gerard, Asli Demirgüç-Kunt, and Edward Kane. 2008. "The 2007 Meltdown In Structured Securitization: Searching for Lessons, Not Scapegoats." World Bank Policy Research Working Paper 4756.

Carruthers, Bruce, and Terrence Halliday. 2007. "Institutionalizing Creative Destruction: Predictable and Transparent Bankruptcy Law in the Wake of the East Asian Financial Crisis." In Meredith Woo-Cummings, ed., *Neoliberalism and Institutional Reform in East Asia: A Comparative Study.* Ithaca: Cornell University Press, 238–72.

Claessens, Stijn, and Leora Klapper. 2005. "Bankruptcy Around the World: Explanations of its Relative Use." *American Law and Economic Review* 7:253–83.

Claessens, Stijn, Simeon Djankov, and Leora Klapper. 2003. "Resolution of Corporate Distress in East Asia." *World Bank Journal of Empirical Finance* 10:199–216.

Claessens, Stijn, Simeon Djankov, and Colin Xu. 2000. "Corporate Performance in the East Asian Financial Crisis." *World Bank Research Observer* 15(1): 23–46.

Correa, Paolo, and Mariana Iootty. 2010. "The Impact of the Global Economic Crisis on the Corporate Sector in Europe and Central Asia: Evidence from a Firm-Level Survey." Enterprise Surveys Note Series, 9. Financial Crisis. EU 10 Regular Economic Report. The World Bank, Washington D.C.

Davydenko, Sergey, and Julian Franks. 2008. "Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany and the UK." *Journal of Finance* 63(2): 565–608.

Demirgüç-Kunt, Asli. 2009. "Dealing with Financial Distress in Systemic Crises." Viewpoint Note, World Bank, Washington DC.

Demirgüç-Kunt, Asli, and Ross Levine. 2008. "Finance, Financial Sector Policies, and Long-run Growth." World Bank Policy Research Working Paper 4469.

Demirgüç-Kunt, Asli, and Luis Servén. 2009. "Are all the Sacred Cows Dead? Implications of the Financial Crisis for Macro and Financial Policies." World Bank Policy Research Working Paper Series, 4807.

Dewaelheyns, Nico, and Cynthia Van Hulle. 2006. "Corporate Failure Prediction Modeling: Distorted by Business Groups' Internal Capital Markets?" Journal of Business Finance & Accounting *33(5–6)*: 909–31.

Martino Di, Paolo. 2002. "Approaching Disaster: A Comparison between Personal Bankruptcy Legislation in Italy and England (1880–1930)." *Business History* 47(1): 23–43.

Djankov, Simeon. 2009. "Bankruptcy Regimes during Financial Distress." Processed. The World Bank, Washington D.C.

Djankov, Simeon, Oliver Hart, Caralee McLiesh, and Andrei Shleifer. 2008. "Debt Enforcement Around the World." *Journal of Political Economy* 116(6): 1105–49.

Eslava, Maricela, John Haltiwanger, Adrian Kugler, and Maurice Kugler. 2004. "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia." *Journal of Development Economics* 75(2): 333–71.

Forum, EStandards. 2010. "Financial Standards Foundation, Insolvency Framework, Italy." Country framework.

Foley, Fritz. 2001. "Going Bust in Bangkok: Lessons from Bankruptcy Law Reform in Thailand." Harvard Business School Mimeograph.

Foster, Lucia, John Haltiwanger, and C.J. Krizian. 1998. "Aggregate Employment Dynamics: Building from Microeconomic Evidence." *American Economic Review* 87(1): 115–37.

Funchal, Bruno. 2008. "The Effects of the 2005 Bankruptcy Reform in Brazil." *Economics Letters* 101(1): 84–6.

Garcia, Ricardo. 2008. "Ley de Concursos y Reorganización [Bankruptcy and Reorganization Law]." El País *(newspaper online)*.

Gilson, Stuart. 1989. "Management Turnover and Financial Distress." *Journal of Financial Economics* 25:241–62.

Gilson, Stuart, and Michel Vetsuypens. 1994. "CEO Compensation in Financially Distressed Firms: An Empirical Analysis." *Journal of Finance* 48(2): 425–58.

Gine, Xavier, and Inessa Love. 2008. "Do Reorganization Costs Matter for Efficiency? Evidence from a Bankruptcy Reform in Colombia." World Bank Policy Research Working Paper 3970.

Hart, Oliver. 2000. "Different Approaches to Bankruptcy." Harvard Institute of Economic Research Working Papers 1903.

Iskander, Magdi, Gerald Meyerman, Dale Gray, and Sean Hagan. 1999. "Corporate Restructuring and Governance in East Asia." *Finance & Development* 36:42–5.

Jackson, Thomas. 1982. "Bankruptcy, Non-bankruptcy Entitlements, and the Creditors' Bargain." *Yale Law Journal* 91:857–907.

Klapper, Leora, Luc Laeven, and Raghuram Rajan. 2006. "Barriers to Entrepreneurship." *Journal of Financial Economics* 82:3.

Lee, Seung-Hyun, Mike Peng, and Jay Barney. 2007. "Bankruptcy Law and Entrepreneurship Development: A Real Options Perspective." *Academy of Management Review* 32(1): 257–72.

Lim, Youngjae, and Chin Hee Hahn. 2003. "Bankruptcy Policy Reform and Total Factor Productivity Dynamics in Korea: Evidence From Microdata." In T. Ito, and A.K. Rose, eds., *Growth and Productivity in Asia*. National Bureau of Economic Research Working Paper 9810, East Asia Seminar on Economics, vol. 13. London: University of Chicago Press, 297–322.

Lucheux, Jean-Michel, and Olivier Pusch. 2009. "An Overview of French Insolvency Law." *International Financial Law Review*, online edition.

Meier, Werner, Kern Michael, and Christoph Schauenburg. 2010. "Proposed Insolvency Reform to Boost Restructurings in Germany." *Lexology*, online edition.

Miller, Marcus, and Joseph Stiglitz. 1999. "Bankruptcy Protection Against Macroeconomics Shocks: The Case for a 'Super Chapter 11'." CSGR Hot Topics: Research on Current Issues 08, Centre for the Study of Globalisation and Regionalisation (CSGR), University of Warwick.

Ministry of Justice of the United Kingdom. 2010. "Quarterly National Statistics."

Modigliani, Franco, and Enrico Perotti. 2000. "Security Markets versus Bank Finance: Legal Enforcement and Investors' Protection." *International Review of Finance* 1(2): 81–96.

Nickell, Stephen. 1996. "Competition and Corporate Performance." *Journal of Political Economy* 104: 724–46.

Novarese, Aldo. 2009. "Bankruptcy Law and Reforms." In H. Gibbon, and Q. Carruthers, eds., "Corporate Restructuring: The Breaking Wave." Thomson Reuters IFR, *Market Intelligence* 133–6.

OECD (Organisation for Economic Co-operation and Development). 2009. "The Impact of the Global Crisis on SME and Entrepreneurship Financing and Policy Responses." Osicka, Tomas, Ida Kucerova, and Petr Mestanek. 2008. "The Modernisation of Czech Insolvency Law." *Linklaters 2*, online newsletter.

Porter, Michael. 1990. *The Competitive Advantage of Nations*. New York: Free Press.

Saidi, Nasser. 2009. "Insolvency and Creditor Rights Systems in MENA." Presentation at the Hawkamah Symposium on Insolvency Laws and Creditor Rights Systems. Dubai International Financial Center.

Scott, W. Richard. 1995. *Institutions and Organizations*. Thousand Oaks, CA: Sage.

Statistisches Jahrbuch Für die Bundesrepublik Deutschland mit "Internationalen Übersichten." 2009. Statistisches Bundesamt Deutschland (Federal Statistical Office), Wiesbaden, Germany.

Stiglitz, Joseph. 2002. "Globalization and its Discontents." New York: W.W. Norton & Company.

Sullivan, Teresa, Elizabeth Warren, and Jay Lawrence Westbrook. 1989. "As We Forgive Our Debtors." New York: Oxford University Press.

Databank, Teikoku. 2010. Bankruptcy Report for 2009, online edition.

UNCITRAL (United Nations Commission on International Trade Law). 2005. "Legislative Guide on Insolvency Law." New York: United Nations Publication no. E.05.V.10

United States Department of Commerce. 2010. "Doing Business in Estonia: A Country Commercial Guide for U.S. Companies." Country Report.

USAID. 2009. "Significant Business Reforms in Tajikistan signed into Law." Press Release no. 090608.

Uttamchandani, M. 2008. "From Crisis to Crisis." In J. Sarra, ed., *Annual Review of Insolvency Law*. Toronto: Thomson-Carswell.

Uttamchandani, Mahesh, and Antonia Menezes. 2010. "Freedom to Fail: Why Small Business Insolvency Regimes are Critical for Emerging Markets." *International Corporate Rescue*, 7(4): 262–8.

Visaria, Sujata. 2009. "Legal Reform and Loan Repayment: The Microeconomic Impact of Debt Recovery Tribunals in India." *American Economic Journal: Applied Economics* 1(3): 59–81.

Wang, Jialan. 2009. "The Role of Human Capital in Corporate Bankruptcy." Massachusetts Institute of Technology, Job Market Paper.

Warren, Elizabeth, and Jay Lawrence Westbrook. 1999. "Financial Characteristics of Business in Bankruptcy." *American Bankruptcy Law Journal* 73:499–589.

World Bank. 2005. "World Bank Principles for Effective Insolvency and Creditor Rights Systems." Processed.

———. 2009. *Doing Business Report, 2009*. Washington, DC: World Bank.

———. 2010. *Doing Business Report, 2010*. Washington, DC: World Bank.

School Feeding Programs and Development: Are We Framing the Question Correctly?

Harold Alderman · Donald Bundy

School feeding programs are politically popular interventions. They are, nevertheless, difficult to assess in terms of effectiveness since their impact is partially on education and partially on school health. They are, additionally, a means to augment consumption by vulnerable populations. The authors look at recent evidence from in-depth studies and argue that while school feeding programs can influence the education of school children and, to a lesser degree, augment nutrition for families of beneficiaries, they are best viewed as transfer programs that can provide a social safety net and help promote human capital investments. JEL codes: H, I, O

Nearly every country in the world today, whether high or low income, seeks to feed at least some of its school children through government sponsored programs. Moreover, when the financial crisis emerged in 2008, the World Bank crisis response mechanisms experienced unprecedented demand to strengthen support for school feeding programs. Yet despite this popularity there remain questions about the evidence of its effectiveness, and there is a continuing struggle to identify what makes for a successful program. For example in 2002 the United States General Accounting Office (USGAO) published a report that claimed "school feeding programs may not be cost effective when compared with alternative interventions such as providing quality teaching and offering nutritional and health packages directed at pregnant women and at mothers with their preschool children" (USGAO 2002, p. 3) and, at the same time, laid out a plan for a pilot to reassess school feeding programs. With a similar motive, in 2009 the World Bank and the World Food Programme (WFP) conducted a joint analysis with the title

"Rethinking School Feeding," explicitly acknowledging the need to clarify the underlying issues (Bundy and others 2009).

A key question relates to the specific benefits of school feeding. It is claimed, for example, that school feeding programs which provide meals at school (SFPs) or related take home rations (THRs) can improve enrollment and attendance, can address chronic hunger or micronutrient deficiencies and, by improving health or by increasing a child's focus in the classroom, can enhance learning. Given the range of countries that employ these two categories of programs—collectively called food for education (FFE)—in one context or another, the results of studies of FFE programs are quite heterogeneous apart from any differences in research methodology (Adelman, Gilligan, and Lehrer 2008). Additionally the conclusions drawn from such studies depend, in part, on how the questions are framed.

We review some recent evidence on school feeding and make the case that the strongest direct consequence of school feeding is best viewed as a form of an income transfer to assist low income households, although there is also a case to be made for a complementary role in education. As such, a primary role is to reduce current poverty with the additional benefit of promoting the accumulation of human capital by jointly influencing education and health. That is, FFE may address both equity and economic efficiency (Das, Do, and Özler 2005).

Figure 1 serves as a starting point for this discussion. The country pattern of a declining ratio of school feeding to education expenditures is analogous to Engle's law. Food budgets (costs of SFP) increase somewhat over GNP range but other schooling expenditures increase more rapidly (figure 2). Over much of the range for middle-income and rich countries the ratio is surprisingly constant at 10 to 20 percent, but for a few countries, mostly low-income African nations, SFP cost per beneficiary is as much as is spent on the average student in basic education or nearly so.

Is the comparison to education expenditures fair? At one level it is useful in providing a comparator with another important intervention for the same age group, but the real question is whether we should view FFE as a cost to education or as a cost to some larger development goal. While it is conceivable that there is some notional tradeoff between school feeding budgets and the budget that is made available for other educational programs—or other investments in nutrition—there is little empirical evidence that tests this conjecture. Conceivably expenditures for FFE crowd out other school expenditures—for example when they are funded from a fixed Education for All Fast Track Initiative allocation. However, in the absence of research on the budgeting process they may also be considered as the core of a country's food security budget, as in the case of the 2001 order of the Indian Supreme Court, mandating midday meals as part of fulfilling the constitutional right to food, or as a component of the Zero Hunger program of Brazil. Indeed, the current political trend is clearly to view FFE as a social intervention that transcends the education goals.

Figure 1. Ratio of per Child Cost of School Feeding in Relation to per Child Cost of Basic Education, Plotted against GDP per capita
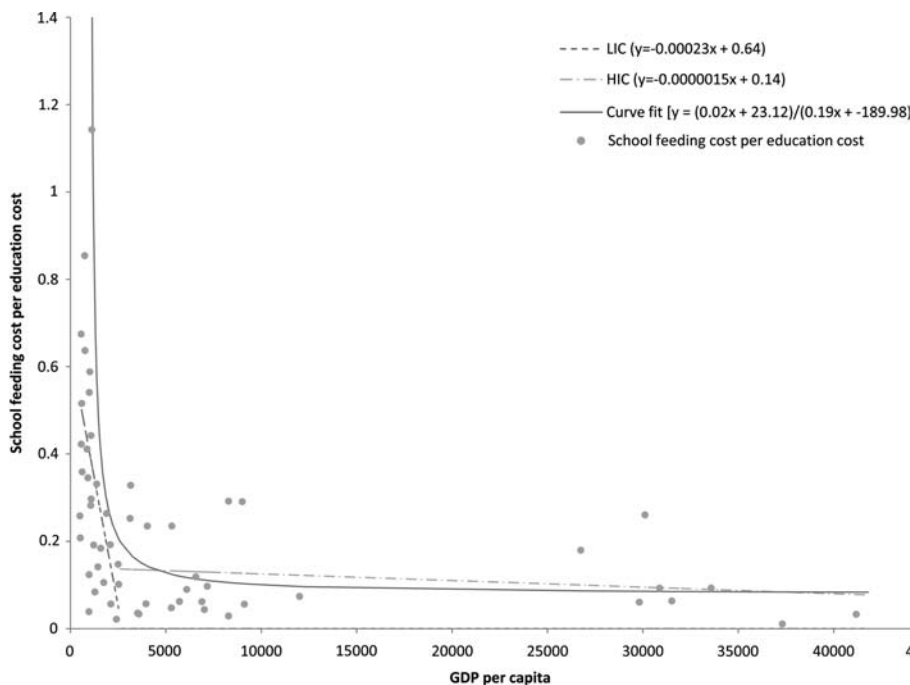
Notes: HIC is high income country, LIC is low income country.
Source: Bundy and others (2009).

If SFPs are social protection expenditures then should they not be compared to levels of other safety nets? On this criterion SFPs are similar to annual transfers per beneficiary in many conditional cash transfer (CCT) programs. Globally, SFPs cost $40–50 a year per beneficiary (and may be several times this per family, depending on the number of children benefiting). This is roughly half of the average magnitude of transfers per household in CCTs.[1] This comparison is particularly appropriate to the degree that FFE can be viewed as conditional-in-kind transfers. However, it is not the objective of this review to compare the two programs—few, if any, direct evaluations have been undertaken—but rather to look at FFE both from the perspective of the efficiency impact on human capital investments and from its role as a transfer program.

To do this we look at the monitored effect of meals compared to alternatives including THRs and snacks. Ideally one would also want to know the costs per outcome. This is hindered both by the scarcity of detailed administrative costs and by the relative scarcity of studies comparing modality of delivery in the same

The World Bank Research Observer, vol. 27, no. 2 (August 2012)

Figure 2. Changes in the Costs per Child of School Feeding and Primary Education with Economic Growth, per capita GDP for 58 countries
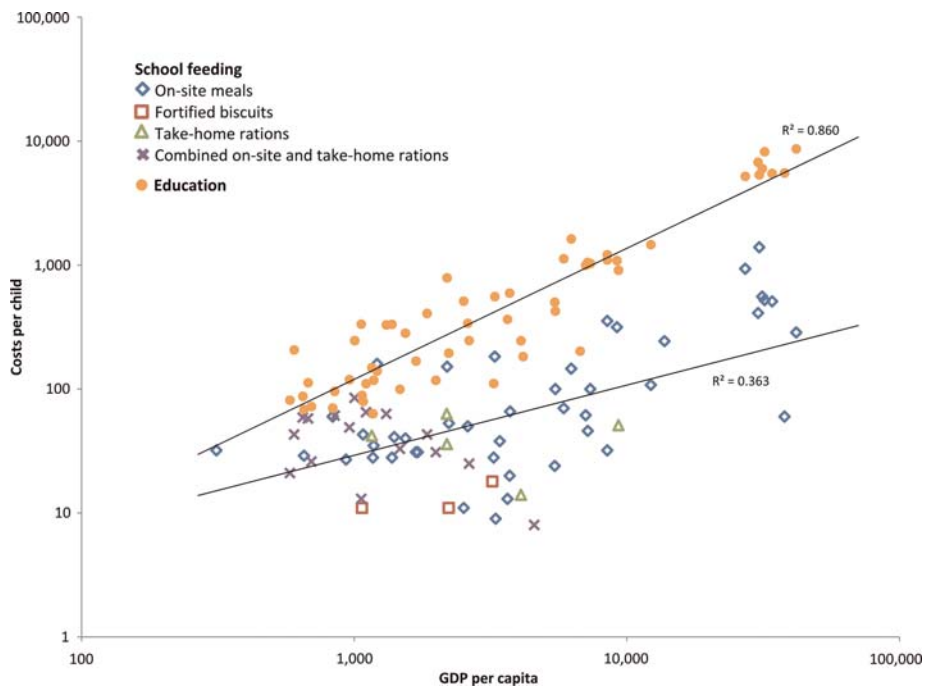
Source: Bundy and others (2009).

setting and time period. Thus, while we consider the general literature on FFE, we pay particular attention to a set of three studies undertaken by the World Bank in conjunction with the WFP. These studies used a randomized longitudinal experimental design to compare SFPs with THRs, and to compare both with a control group. The common study design used baseline and follow-up surveys with a household sample that allowed for an assessment of the ability of a program to attract new students as well as to facilitate the measurement of the spillover from the program to other household members. These three projects on which we particularly focus on are:

- A study in Uganda undertaken between 2005 and 2007 in internally displaced people's (IDP) camps in the Pader and Lira districts of Northern Uganda. While the IDP setting is somewhat unprecedented for studies, it does not necessarily rule out external validity since over half of the WFPs are in emergency situations.

- A parallel study in Burkina Faso that was conducted in four provinces in the Sahel region (Gorom, Oudalan, Soum, and Yahga) with the program delivering food in the 2006/07 school year.
- An assessment of school feeding extended to two northern provinces of Lao PDR in 2006–08.

School Feeding as a Nutrition Program

The direct impact of FFE programs on nutrition has often been measured in terms of the net increase of food consumed by the student over a 24-hour period. For this increase one needs to take into account not only the content and frequency of school meals[2] but also any reallocation of resources within the household. In the case of meals consumed at school, this sharing would come about from reallocation of food provided at home during other meals. This could partially offset the increment in school and, thus, achieve an indirect sharing of the meal or snack. This is often referred to as leakage, although such a phrase is misleading as it differs markedly from a more common concept of leakage—that is, it differs from mistargeting of transfers intended for the poor to wealthy households or from private diversion of public resources.

Using a random assignment of the dates of a 24-hour food recall survey, Jacoby (2002) ascertained that school snacks in the Philippines were completely additional resources to the students in the program. That is each additional calorie provided in school led to an identical increase to the total calories consumed by the student during the day. This is deemed a flypaper effect, as the food resources stick with the school-aged child. However, unless the snack was unknown to the rest of the household, the full capture by the student is not compatible with most household allocation models (Haddad, Hoddinott, and Alderman 1997). Even bargaining models are unlikely to produce a polar case with no sharing of resources with other household members.

While the absence of any sharing is a puzzle, Jacoby's empirical strategy is, nevertheless, solid. Moreover subsequent studies have used a similar methodology to replicate and expand upon Jacoby's result. For example Afridi (2010) looked at school meals in India. While the point estimates for the unit increase of total nutrient intake for each of five nutrients provided in this school meal program that was studied are less than one, these were often not significantly different from one. A coefficient of one implies that one calorie or other nutrient consumed from the school meal leads to a one calorie increase in total consumption for the day. Thus this study is consistent with Jacoby's results. In addition Ahmed (2004) used

an individual fixed-effect variant of Jacoby's approach in Bangladesh and found, again, virtually a one-to-one increase in total calorie intake from a snack provided in school. Islam and Hoddinott (2009) find some reallocation of food to other family members—and also note that reallocation from each child's school meals may be limited by the fact that in many families more than one child is a program beneficiary—but they also find that diet quality improves. This is indicated by the fact that half of the calories provided were reallocated within the household, while only 20 percent of the protein was reduced by household sharing.

But, in fact, from the standpoint of nutrition, the amount of calories that is additional in the diet of the student is not the core issue. Rather the main limitation of school feeding programs—and studies of school feeding—is that they generally do not focus on the most vulnerable period for malnutrition, which is the period spanning development *in utero* through to two years of age (Shrimpton and others 2001).[3] A few recent studies have turned the flypaper studies on their head and looked at the impact of school feeding on the younger, more vulnerable, age group by including siblings of students in impact evaluations using randomized design.

For example, in Burkina Faso, weight for age increased by 0.38 standard deviations for children aged 12–60 months whose sisters were eligible for a THR compared to a control group (Kazianga, de Walque, and Alderman 2009). Comparable children in the treatment villages who did not have a school-aged sister and thus were not eligible for the program did not show this improvement, implying that local area affects are unlikely to account for the result. This increase was greater than could be expected from the implicit income transfer. This may reflect what is referred to as a labeling effect by which a program encourages a reallocation of household resources (Kooreman 2000). Such an increase of allocation toward food and nutrition beyond the preprogram marginal budgets has been observed for food stamps in the United States (Breunig and Dasgupta 2005) and for cash transfers in Ecuador (Paxson and Schady 2008).

In Uganda, younger siblings of beneficiaries of a SFP had a significant improvement in height for age of 0.36 standard deviations. In contrast to the Burkina Faso results a similar increase was not observed for children in families that received THRs. Also the Uganda investigation found that both THRs and SFPs contributed to a significant relative improvement in anemia prevalence of adolescent girls, an age at which anemia rates tend to increase, an outcome that was not studied in the Burkina Faso study. The mothers of young girls in the Uganda THR programs also had lower anemia rates than the control group, although the SFP did not show a similar benefit.

Since SFPs are widespread even in middle- and upper-income countries, evaluations of their nutritional impact also need to consider their potential contribution to obesity. While countries such as Brazil and Chile have redesigned

their school meal programs to address this risk (Doak 2002), others have yet to consider the problem of obesity.[4] Often the most successful programs to address the risk of obesity combine changes in the composition of meals provided with nutrition education (Foster and others 2008).

Using school meal programs as a vehicle for education is not confined to the prevention of obesity and related chronic illness. Such programs can be a means to promote basic health services such as hand washing or deworming. While the biannual schedule advised for deworming does not coincide with the delivery of either school meals or most THRs, it is now very common to include deworming in the planning for FFE (Del Rosso 1999; Bundy and others 2006).

School meal programs can also be a vehicle for improved micronutrient status if the meals or rations are fortified or if they contribute to an increase of diet diversity. While studies often—but not universally—find benefits from the inclusion of meat in school meal programs (Whaley and others 2003), such meals are often impractical or too expensive for low income settings. In contrast, fortification generally adds very little to the costs of FFE. For example, biscuits fortified with iron and iodine were found to reduce absenteeism as well as to improve some dimensions of cognitive function relative to a similar snack without fortification (van Stuijvenberg and others 1999). As the control group also received a snack, the impact of the fortification was additional to the unmeasured impact of the provision of food at the start of the school day.

Nevertheless the logistics of fortification may be influenced by local procurement strategies. Although some foods such as wheat or maize flour can be fortified in decentralized milling, other commodities are harder to fortify. This is especially the case when multiple fortification is recommended. As a general rule, the more processed the items in a FFE program the greater the share of costs for transport and packaging. Moreover, fortification is less likely when FFE is locally procured. Currently there are few programs where local procurement is the sole source of food, so there remain opportunities for centralized fortification. As decentralized procurement increases, there may be an increased role for school fortification using prepackaged mixes. This remains an area for research.

School Feeding as an Education Program

Numerous studies show that in-school feeding has a positive impact on school enrollment or participation in areas where initial indicators of school participation are low (Jukes, Drake, and Bundy 2007; Kristjansson and others 2007; Adelman, Gilligan, and Lehrer 2008). In many cases the impact may appear modest because initial enrollment rates are high and thus cannot be substantially increased. However, impacts may also be low because the time frame of studies—

particularly randomized studies that require a control group to be phased in at a later date—often do not have adequate time to show the cumulative impact of a program (Behrman and King 2009). For example, while *overall* enrollment in the Uganda study did not increase significantly in an 18-month period, an SFP led to a significant 9 percent increase in the share of children aged 6–13 who *started* school compared to the control group (Alderman, Gilligan, and Lehrer 2010). THRs also contributed to an increase that, while not significantly different from zero, was also not significantly less than the increase attributed to SFPs. In both modalities of delivery of FFE children entered at a younger age than children in the control communities.

Results from Burkina Faso are similar: both school meals and take home rations increased new enrollment of girls by about 5 to 6 percent. Even fortified biscuits provided as snacks may impact on enrollment; Ahmed (2004) reports a 14 percent difference in enrollment in Bangladesh using a matched (non-experimental) cross-sectional analysis of communities with and without such a program.

The gender specific impacts reported from Burkina Faso are in keeping with a common expectation that FFE will have greater impacts on girls than boys (Dréze and Kingdon 2001; Gelli, Meir, and Espejo 2007). Indeed THRs are often targeted only to girls, as was the case in Burkina Faso. However, not all studies of enrollment have a difference by gender; the enrollment impacts in Uganda were gender neutral. This may reflect the fact that, unlike Burkina Faso, there was no gender difference in primary enrollment rates at baseline.

Studies of FFE regularly report increased attendance, often using school based samples and thus these studies generally present results conditional on enrollment. Most studies show a positive impact although the results are often nuanced. For example in Uganda there was no effect on self-reported attendance. However, there were higher rates of attendance based on results of four randomly timed spot visits for both SFPs and THRs. The increase in morning attendance compared to controls was around 9 percent in both programs, although the increase was mainly for boys in THR and for girls when the intervention was SFP. The impact on afternoon attendance was somewhat larger than it was on morning attendance but there was no difference by gender or by program type in the afternoon.

The Burkina Faso study also indicated heterogeneity on attendance with respect to household size. Attendance, recorded close to the planting season, increased in both THR and SFP when the household had spare labor (three or more children in addition to the student) but decreased when there was no other child or only one sibling. This decrease may be due to the program attracting children with higher opportunity costs into the schools.

Vermeersch and Kremer (2005) also indicate a significant increase of attendance when school meals were offered to a randomized sample of children in

Western Kenya. The 30 percent increase is relatively large, but this may reflect the fact that their sample was of preschool children in which initial school participation was much lower than it is in basic education; only a third of their sample participated in preschool at baseline. As preschools generally have lower enrollment then primary schools—and where enrollments are more skewed to relatively well off children—this example may point to an area where SFPs may be particularly efficacious.

Vermeersch and Kremer also found that the school meal program led to an increase in scores on written and oral tests of performance, relative to the school curriculum, after two years participation in school. While the school meal program improved performance this was only noted in schools where the teachers had greater than average experience. The absence of a more general improvement was attributed, in part, to an increase in class size and in the reduced time for teaching necessitated by food preparation.

Improved performance as measured by tests of achievement is often reported for FFE, although there is a fair amount of variance as to which ages and which skills are most affected (Jukes, Drake, and Bundy 2007; Adelman, Gilligan, and Lehrer 2008). For example, in the recent study on Uganda, both SFP and THR had significant impacts on math test scores of children aged 11–14, but there was no impact on the test of literacy and only THR had a significant impact on Primary Leaving Exam scores.

Improvements in test scores may either reflect total time in the classroom, the possibility that FFE increases the amount of learning per day of schooling, or both. A few studies have attempted to investigate this second avenue of increased receptivity to instructions by looking at the tie between hunger and classroom performance using an experimental design. Available results, however, are not conclusive regarding long-term consequences, perhaps, in part, because controlled studies are hampered by difficulties in running experiments for an appreciable duration as well as the difficulty of encouraging parents to conform to the protocols of research design and the inability to use a placebo. Moreover, as shown in Grantham-McGregor, Chang, and Walker (1998), while feeding children may improve attention, its impact on learning depends on the classroom organization. The impact also depends on the timing; school lunches may have a very different impact on classroom performance.

Additional evidence on the impact of FFE may come from comparisons of measures of cognitive ability such as scores on Raven matrices, forward digit span (this is a test of working memory that asks a child to repeat strings of numbers of different lengths), or backwards digit span (which also assesses executive function since this involves manipulating information). While results on such tests from Kenya (Whaley and others 2003) as well as Uganda contribute to the evidence base that FFE can influence cognitive ability, this pathway to improved outcomes

may be less direct than that mediated by attention or attendance since it depends on the quality of education that is available. This is commonly observed with other school health interventions as well. For example malaria reduction in school age children in Kenya resulted in a decline in the prevalence of anemia and a concomitant enhancement in performance on cognitive tests, but no measurable improvement in education outcomes due to the lack of quality education inputs (Clarke and others 2008). This finding helps emphasize that FFE programs can only be effective in education terms if combined with quality education programs.

Another perspective of the impact of FFE on learning is provided by Ahmed and Arends-Kuenning (2006). They find a decrease of scores on the government test administered in the fourth grade in a THR program in Bangladesh. They attribute this to peer effects; not only did the targeted program bring in new students with lower than average scores, the scores for nontargeted students declined. However, the study ruled out the possibility that this was due to more crowded classrooms.

School Feeding as a Safety Net

If FFE is viewed as a transfer program, one criterion for assessing effectiveness is targeting efficiency. In general, SFPs are not targeted within schools—although some programs have sliding scales of payments for meals. Thus targeting will mainly reflect the choice of schools to be included, often on a geographic basis. Lindert, Skoufias, and Shapiro (2010) indicate that FFE programs in Latin America are generally progressively targeted. However, they noted that in Guatemala the poorest quintile received less assistance than the middle class, perhaps reflecting exclusion of schools in more remote areas.

THRs often have an additional layer of targeting in that the individuals within a school may not all be eligible. As with much of the targeting literature, results are mixed. One of the more detailed studies of targeting of FFE showed pro-poor targeting within schools but little evidence that the geographic targeting was pro-poor or designed to increase allocations to those schools where targeting was more effectively carried out (Galasso and Ravallion 2005).

THRs are often targeted by gender, reflecting both the evidence that girl's schooling frequently lags behind boys schooling and the expectation that girls schooling is more responsive to supply-side interventions (de Janvry and Sadoulet 2006). While gender-based targeting is administratively simple to implement, over recent years the number of settings where gender discrimination occurs in basic schooling has been substantially reduced (Grant and Behrman 2010). Thus in many communities gender-based targeting may be less effective at reducing unequal school participation than income or asset targeting.

Exclusion of poorer schools, however, is not always a case of these schools being excluded from program eligibility; in Laos the probability that a school would take up FFE assistance that was offered was negatively associated with the education of the community or with current enrollment rates, as well as the altitude of the community (Buttenheim, Freidman, and Alderman 2011). The percent of villages that had schools and were offered FFE and took up the offer ranged from 58 to 75 percent in the three districts that were included in the program. Even when the school participated in the program, meals were not regularly provided; the two districts that had SFPs reported that meals were provided between 47 and 58 percent of the days when the school was in session. This latter issue of irregular supply of meals is one that has challenged school feeding in remote areas for years (Levinger 1986). Irregular supply not only dilutes the impact but may have a negative impact to the degree that the unrealized expectation of a school meal crowds out meals or snacks that a parent might have otherwise provided.

In Laos, the cost of transport and storage was often cited by schools as a reason for not taking up the program. Elsewhere it may be the preparation of meals that influences the cost and accounts for irregularity of delivery. Data on costs are, however, often not reported and, in any case, estimates of costs are heterogeneous due to both differences in accounting as well as differences in programs. Galloway and others (2009) report the costs for four programs in Africa as ranging between $28 and $63 per child per year with nonfood costs ranging between 26 and 49 percent of the total.[5] Comparing across modalities is similarly subject to the difficulty in standardizing programs. Gelli and others (2011) come up with an estimate of $48 on average for FFE costs (exclusive of in-school costs) using data from 72 WFP projects with snacks costing only half of meal programs. Thus biscuits were found to be more cost effective for distribution of micronutrients, although SFPs were on average more cost effective in terms of calories provided than biscuits. Likely this would also be the case in terms of implicit transfers, although the calculation comparing biscuits with school meals was not provided.

THRs cost more than twice the average cost of meal programs. However, THRs in the review by Gelli, Al-Shaiba, and Espejo (2009) also provided twice as much food as SFPs, so the transfer benefits were correspondingly higher. The most expensive THR in this review still devoted less than 20 percent of all costs to indirect costs including transport.[6] If one considers the cost of calories provided to the recipient family, THRs are generally more effective than SFPs; only under the criteria of calories provided to school children alone (that is, considering all other transfers to be outside the benefits of the program) do SFPs appear to be more cost effective as a transfer than THRs.

The impact of FFE on a household budget is not identical with the unit cost of the food. That is, food that cost the program a dollar might be valued differently by the household. In more remote areas a FFE program may be able to bring in food at a cost lower than the household would otherwise pay (at some disadvantage to local producers). More commonly the local cost of comparable foods to the beneficiaries will be less than the cost to the program, leading to a transfer value somewhat less than the budgetary outlay. This, of course, is not an issue for cash transfers.

Given the heterogeneity of costs for FFE as well as the range of objectives, only a rough comparison can be offered with the costs of CCTs. As indicated in Caldes, Coady, and Maluccio (2006) CCTs may devote up to 60 percent of costs to identifying beneficiaries in initial years, although this upfront cost is not repeated annually. In contrast, SFPs incur only minimal costs for geographic targeting. The THR programs that use poverty targeting, however, would have associated costs for this screening. It can be assumed that these costs would differ little from a CCT covering the same community. SFPs also do not incur costs for monitoring conditions; the meal is delivered if and only if the child is present. Again since THRs are generally based on attendance there might be costs for verifying compliance. However, as most programs are administered at the school level, the data collection and transmission costs are not generally extensive. Thus the main difference in costs of cash and food programs are, as expected, the difference in the physical transport and handling of commodities.

One study (in Bangladesh) compared school meals to cash support with enrollment, as well as food budgets as a tracked outcome, finding that the former had a larger impact on enrollment. However, the increase in enrollment attributed to school meals relative to cash (36 percent) was virtually the same as the difference in the size of the transfer (41 percent). The main difference in outcomes of the two modes of delivery was that only the food transfer increased household food consumption. The majority of households—80 percent—indicated that they preferred cash to food for the oft recognized flexibility that cash provided. That study, however, did not use an experimental design and, indeed, did not compare programs undertaken in the same year. Thus there is remaining scope to improve programmatic knowledge relevant both to school programs as well as to the broader knowledge of cash programs.

Another criterion to assess FFE as a safety net is its ability to respond to crises. These programs have been relatively easy to scale up in emergencies. For example they were widely used in Africa in the wake of the 2007–08 food price spike; Burundi, the Central African Republic, Ghana, Liberia, and Togo all established or expanded their SFPs (Wodon and Zaman 2010). While Africa relied more heavily on in-kind transfers (as opposed to cash) in response to the food price spike than other regions, the expansion of school feeding during this global crisis was not confined to that region. In one notable example, the Philippines employed

expanded school feeding as part of a multipronged program to protect its poor from a precipitous rise in the price of rice (World Bank 2010, box 2.4). Thus, despite concerns over capacity mentioned above, FFE has proven flexible in response to crises.

A key change in the context of FFE programs over the last four to five years has been the move away from food aid. This reflects many interacting factors in the global economy, including rising commodity prices, increased demand for agricultural products for nontraditional purposes (such as fuel and alcohol production), and trends in agricultural subsidies. Whatever the reasons, today there is a tendency to favor the local purchase of food for FFE programs. This has increased focus on procurement and quality. In particular, there is a movement towards so-called home grown school feeding in general, with the emphasis on food procured in the communities around the school, thus enhancing both the rural economy and food quality.[7] Where local prices are below import parity prices (or where FFE assistance has requirements that put the cost of food above import parity prices) such programs can reduce the cost of school feeding. Their impact on farmers' incomes or on the prices that local food purchasers face depends on market integration and, thus, will vary according to local conditions. FFE programs in Osun State in Nigeria and in Côte d'Ivoire have, however, demonstrated the sustainability of such programs. Further research is required to confirm their apparently major contribution to local economies.

Conclusion

Do the results reviewed here imply that FFE is among the best investments in nutrition? Despite new evidence indicating favorable externalities to siblings of students, and the clear benefit in addressing hunger in schoolchildren, the fair answer to this question is no. While FFE can provide iron and other key micronutrients, these programs are not designed to address the most critical nutritional constraints in low income settings, simply because they are not targeted at the most vulnerable period in child development, which is between conception and two years of age.

Do the results imply that FFE is the best way to use funds for education? Again, the quick answer is likely no. However, in this case, the answer is more nuanced. FFE is not a substitute for a well-organized education system and teacher performance. However, there is extensive evidence that FFE can complement a good education program. So although FFE may not be the best education response it may be an important element in achieving an effective education system. In Addis Ababa, in February, 2010, the 9th Annual Meeting of the High Level Group on Education for All recognized this contribution in including school feeding in their

call for "Education for All Partners to intensify efforts to support initiatives targeted at the most marginalized, such as cash transfers, school health and school feeding, scholarships and gender-specific interventions" (Bundy 2011). Most clearly this comes from demand-side encouragement of schooling in settings where universal basic schooling is not yet achieved and, perhaps, where preschool programs reach low income households. FFE may also have a particular role in programs that are attempting to expand schooling to cover a longer day. These programs may enhance learning per time invested in school but, as mentioned, such a desired impact is not inevitable.

Do the results imply that FFE is a plausible candidate for a social protection investment on a par with CCTs? Here the fair answer appears to be: quite likely. FFE can increase human capital investments while also providing support to poor households. Thus they serve as a support to current poverty reduction while making the need for future transfers and assistance less likely. The dual objectives of raising current consumption while promoting investments, however, make it difficult to compare outcomes of either CCTs or FFE with direct investments. The value of transfers does not easily aggregate with outputs in a benefit cost assessment. For one thing such a summation requires a quantification of the weight society puts on consumption of the poor relative to that of the average citizen. Absent this calculation, a direct comparison of demand-side interventions for education or direct investments in health with a FFE transfer does not put both categories of expenditures on the same metric. A benefit–cost analysis or a cost effectiveness comparison within a sector generally assumes away the value of the transfer or ignores the benefits outside the sector being considered. However, if the question is phrased as "Can FFE give a government additional value over other forms of transfers?" the answer is clearer: the investment component of FFE has a positive value that can be quantified and which adds to the social value of the transfer to low income households.

Targeting of programs, then, has to balance the dual objectives of equity and efficiency. The former case suggests efforts to include poor households whether or not there is a risk of nonattendance in school while, in the latter case, the prioritization is for the relatively smaller cohort of children who do not participate in education opportunities, including preschool programs where they are available. Improved targeting, however, may find a convergence of equity and efficiency; to the degree that there is heterogeneity of impacts it is likely to show greater improvement in health and schooling among the poorest (Bundy 2011).

There is yet no clear dominance of types of programs in regards to these impacts. For example, while the automatic link of SFPs to attendance might lead one to expect a larger impact of meals compared to THRs, this has not been found in the few direct comparisons of these two modalities. Similarly, as with CCT programs, it is not clear that an increase in the value of a transfer leads to a

proportional increase in the impact on students; a few studies of the impact of school snacks show substantial impact on enrollment comparable to similar studies (undertaken in other settings) of meals.

Ultimately, then, the relative priority of FFE programs hinges on the costs of delivery and on sustainability. THRs, with their potential for targeting, may be a promising part of such a package. Other program modifications to reduce costs, such as local sourcing of inputs and the use of vouchers in lieu of the direct provision of meals, may further the objectives of FFE at lower costs, but at this time innovations are supported more by qualitative reviews than by empirical studies. Still, given the political energy behind FFE, there is likely to be substantial value in understanding where best to place FFE in the range of instruments to reduce the intergenerational transmission of poverty.

Notes

1. Estimated from Fiszbein and Schady (2009), table 2.

2. While some THRs may be delivered throughout the year, SFPs are rarely available when the school is not in session. As such the contribution of SFPs to the diet averaged over a year is often between a half and two-thirds of the daily contribution when the school is open. It is often far less since many SFPs are plagued by irregular availability even on days when schools are in session. Absenteeism—for example during peak agricultural seasons—further reduces the contribution of SFPs to food consumption.

3. Until recently very few studies considered the indirect contribution of FFE to nutrition of young children. For example a recent comprehensive meta-analysis of medical and nutritional literature covering various dimensions of school feeding (Kristjansson and others 2007) does not address the impact on siblings, although it does find an impact on the weights of direct beneficiaries.

4. Chile provides more calories to schools with greater poverty incidence. While regression discontinuity analysis does not show that this has an impact on school performance among the poorest students (McEwan 2010)—few of whom are malnourished by international standards— there is yet no analysis of the impact on obesity.

5. This range partially reflects accounting procedures. Also Lesotho purchases food locally and thus has the highest food costs but no external transport and handling.

6. Excluding this program in the average costs also brought the estimated average down by more than a third.

7. 'Home grown' refers to local procurement. It is not linked to school gardens which are virtually never of adequate scale to address the requirements of SFPs and are detrimental to the objectives of education in general (Bundy and others 2009).

References

The word *processed* describes informally reproduced works that may not be commonly available through libraries.

Adelman, Sarah, Daniel O. Gilligan, and Kim Lehrer. 2008. How Effective are Food for Education Programs? A Critical Assessment of the Evidence from Developing Countries. *IFPRI Food Policy Review 9*. Washington, DC: International Food Policy Research Institute.

Afridi, Farzana. 2010. "Child Welfare Programs and Child Nutrition: Evidence from a Mandated School Meal Program in India." *Journal of Development Economics* 92:152–65.

Ahmed, Akhter. 2004. "Impact of Feeding Children in School: Evidence from Bangladesh." Processed. Washington, DC: International Food Policy Research Institute.

Alderman, Harold, Daniel Gilligan, and Kim Lehrer. 2010. "The Impact of Food for Education Programs on School Participation in Northern Uganda." International Food Policy Research Institute, Washington DC. Processed.

Behrman, Jere, and Elizabeth King. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *World Bank Research Observer* 24:55–82.

Breunig, Robert, and Indraneel Dasgupta. 2005. "Do Intra-household Effects Generate the Food Stamp Cash-Out Puzzle?" *American Journal of Agricultural Economics* 87(3): 552–68.

Bundy, Donald. 2011. *Rethinking School Health: A Key Component of Education for All.* Directions in Development. Washington, DC: World Bank.

Bundy, Donald, S. Shaeffer, M. Jukes, K. Beegle, A. Gillespie, L. Drake, F.L. Seung-hee, A-M. Hoffman, J. Jones, A. Mitchell, C. Wright, D. Camara, C. Golmar, L. Savioli, T. Takeuchi, and M. Sembene. 2006. "School Based Health and Nutrition Programs." In D. Jamison, J.G. Breman, A.R. Measham, G. Alleyne, M. Claeson, D. Evans, P. Jha, A. Mills, and P Musgrove., eds., *Disease Control Priorities in Developing Countries*. 2nd edn. New York: World Bank and Oxford University Press: 1091–108.

Bundy, Donald, Carmen Burbano, Margaret Grosh, Aulo Gelli, Matthew Jukes, and Lesley Drake. 2009. *Rethinking School Feeding: Social Safety Nets, Child Development, and the Education Sector.* Joint publication of the World Food Programme and the World Bank. Directions in Development. Washington, DC: World Bank.

Buttenheim, Alison, Jed Freidman, and Harold Alderman. 2011. "Impact Evaluation of School Feeding Programs in Lao PDR." World Bank Policy Research Working Paper 5518.

Caldes, Natalia, David Coady, and John Maluccio. 2006. "The Cost of Poverty Alleviation Transfer Programs: A Comparative Analysis of Three Programs in Latin America." *World Development* 34(5): 818–37.

Clarke, Siân, Matthew Jukes, J. Kiambo Njagi, Lincoln Khasakhala, Bonnie Cundill, Julius Otido, Christopher Crudder, Benson Estambale, and Simon Brooker. 2008. "Health and Education in Schoolchildren: A Cluster-randomised, Double-blind, Placebo-controlled Trial." *Lancet* 372: 127–38.

Das, J., Q. Do, and B. Özler. 2005. "Reassessing Conditional Cash Transfer Programs." *World Bank Research Observer* 20(1): 57–80.

Del Rosso, J.M. 1999. "School Feeding Programs: Improving Effectiveness and Increasing the Benefit to Education. A Guide for Program Managers." Partnership for Child Development, Oxford, UK.

Doak, Colleen. 2002. "Large-scale Interventions and Programmes Addressing Nutrition Related Chronic Diseases and Obesity: Examples from 14 Countries." *Public Health Nutrition* 5(1a): 275–7.

Dréze, J., and G. Kingdon. 2001. "School Participation in Rural India." *Review of Development Economics* 5(1): 1–24.

Fiszbein, A., and N. Schady. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty.* Policy Research Report. Washington, DC: World Bank.

Foster, G.D., S. Sherman, K.E. Borradaile, K.M. Grundy, S.S. Vander Veur, J. Nachmani, A. Karpyn, S. Kumanyika, and J. Shults. 2008. "A Policy-based School Intervention to Prevent Overweight and Obesity." *Pediatrics* 121(4): 794–802.

Galasso, Emanuela, and Martin Ravallion. 2005. "Decentralized Targeting of an Antipoverty Program." *Journal of Public Economics* 89(4): 705–27.

Galloway, Rae, Elizabeth Kristjansson, Aulo Gelli, Ute Meir, Francisco Espejo, and Donald Bundy. 2009. "School Feeding: Outcomes and Costs." *Food and Nutrition Bulletin* 30(2): 171–82.

Gelli, Aulo, U. Meir, and F. Espejo. 2007. "Does Provision of Food in School Increase Girls' Enrollment? Evidence from Schools in Sub-Saharan Africa." *Food and Nutrition Bulletin* 28(2): 149–55.

Gelli, Aulo, Najeeb Al-Shaiba, and Fransico Espejo. 2009. "The Costs and Cost-Efficiency of Providing Food Through Schools in Areas of High Food Insecurity." *Food and Nutrition Bulletin* 30(1): 68–76.

Gelli, Aulo, Andrea Cavallero, Licia Minervini, Mariana Mirabile, Luca Molinas, and Marc Regnault de la Mothe. 2011. "New Benchmarks for Costs and Cost-efficiency for Food Provision in Schools in Food Insecure Areas." Food and Nutrition Bulletin, forthcoming.

Grant, Monica, and Jere Behrman. 2010. "Gender Gaps in Educational Attainment in Less Developed Countries." *Population and Development Review* 36(1): 71–89.

Grantham-McGregor, S.M., S. Chang, and S.P. Walker. 1998. "Evaluation of School Feeding Programs: Some Jamaican Examples." *American Journal of Clinical Nutrition* 67(4): 785S–789S.

Haddad Lawrence, John Hoddinott, and Harold Alderman, eds. 1997. *Intrahousehold Resource Allocation in Developing Countries: Methods, Models, and Policy.* Baltimore: Johns Hopkins University Press.

Islam, Mahnaz, and John Hoddinott. 2009. "Evidence of Intrahousehold Flypaper Effects from a Nutrition Intervention in Rural Guatemala." *Economic Development and Cultural Change* 57(2): 215–38.

Jacoby, Hanan G. 2002. "Is there an Intrahousehold 'Flypaper Effect'? Evidence from a School Feeding Programme." *Economic Journal* 112(476): 196–221.

de JanvryAlainSadouletElisabeth. 2006. "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality." *World Bank Economic Review* 20(1): 1–29.

Jukes, M.C.H., L.J. Drake, and D.A.P. Bundy. 2007. *School Health, Nutrition, and Education for All: Leveling the Playing Field.* Wallingford, Oxfordshire, UK: CAB International.

Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2009. "Educational and Health Impact of Two School Feeding Schemes: Evidence from a Randomized Trial in Rural Burkina Faso." World Bank Policy Research Working Paper 4976.

Kooreman, Peter. 2000. "The Labeling Effect of a Child Benefit System." *American Economic Review* 90(3): 571–83.

Kristjansson, E.A., V. Robinson, M. Petticrew, B. MacDonald, J. Krasevec, L. Janzen, T. Greenhalgh, G. Wells, J. MacGowan, A. Farmer, B.J. Shea, A. Mayhew, and P. Tugwell. 2007. "School Feeding for Improving the Physical and Psychosocial Health of Disadvantaged Elementary School Children." *Cochrane Database of Systematic Reviews* Issue 1, Art. CD004676. (www.cochrane.org/reviews/en/ab004676.html).

Levinger, Beryl. 1986. "School Feeding Programs in Developing Countries: An Analysis of Actual and Potential Impact." Washington, DC: USAID Evaluation Special Study 30. (http://pdf.usaid.gov/pdf_docs/PNAAL060.pdf).

Lindert, Kathy, Emmanuel Skoufias, and Joseph Shapiro. 2010. "Redistributing Income to the Poor *and* the Rich: Public Transfers in Latin America and the Caribbean." *World Development* 38(6): 895–907.

McEwan, Patrick. 2010. "The Impact of School Meals on Education Outcomes: Discontinuity Evidence from Chile." Processed. Department of Economics, Wellesley College, MA.

Paxson, Christina, and Norbert Schady. 2008. "Does Money Matter? The Effects of Cash Transfers on Child Health and Cognitive Development in Rural Ecuador." Processed. World Bank, Washington DC.

Shrimpton, R., C. Victora, M. de Onis, R. Costa Lima, M. Blössner, and G. Clugston. 2001. "Worldwide Timing of Growth Faltering: Implications for Nutritional Interventions." *Pediatrics* 107:75–81.

van Stuijvenberg, M.E., J.D. Kvalsvig, M. Faber, M. Kruger, D.G. Kenoyer, and A.J. Benadé. 1999. "Effect of Iron-, Iodine-, and β-carotene-fortified Biscuits on the Micronutrient Status of Primary School Children: A Randomized Controlled Trial." *American Journal of Clinical Nutrition* 69: 497–503.

USGAO (United States General Accounting Office). 2002. *Global Food for Education Initiative Faces Challenges for Successful Implementation*. Report to Congressional Requesters, February, GAO-02-328, Washington DC.

Vermeersch, C., and M. Kremer. 2005. "School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation." World Bank Policy Research Working Paper 3523.

Whaley, Shannon, Marian Sigman, Charlotte Neumann, Nimrod Guthrie, Robert E. Weiss, Susan AlberSuzanne P. Murphy. 2003. "The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children." *Journal of Nutrition* 133(11): 3965S–3971S.

Wodon, Quentin, and Hassan Zaman. 2010. "Higher Food Prices in Sub-Saharan Africa: Poverty Impact and Policy responses." *World Bank Research Observer* 25(1): 157–76.

World Bank. 2010. *Global Monitoring Report 2010. After the Crisis*. Washington, DC: World Bank.

International Grain Reserves And Other Instruments to Address Volatility in Grain Markets

Brian D. Wright

In the long view, recent volatility of prices of the major grains is not anomalous. Wheat, rice, and maize are highly substitutable in the global market for calories, and when aggregate stocks decline to minimal feasible levels, prices become highly sensitive to small shocks, consistent with the economics of storage behavior. In this decade, stocks declined due to high global income growth and biofuels mandates, making markets unusually sensitive to subsequent unanticipated shocks, including biofuels demand boosts in reaction to high petroleum prices, the Australian drought, and other regional grain production problems. To protect their own vulnerable and politically influential consumers, key exporters restricted supplies in 2007, exacerbating the price rise. Understandably, vulnerable importers are now building strategic reserves. To reduce costs and disincentive effects, reserves should have quantitative goals related to targeted distribution to the most vulnerable in severe emergencies. For countries with significant animal feeding or biofuels industries, options contracts to protect the consumption of the most vulnerable from harvest shocks are likely to be more cost-effective than emergency reserves.

1. Introduction: The Food Price Crisis of 2007/08 and the Re-emergence of Concerns over Commodity Price Volatility

The increases during 2007/08 in the prices of many consumption commodities, including the major grains, came as a shock to consumers and governments. Millions of the world's poor were likely forced to reduce their daily calorie intake, and urban consumers participated in protests, often violent, that placed serious pressure on governments in developing countries.[1]

In response, many nations adopted short run policies to reduce the effects of rising world prices on domestic consumers. Though perhaps rational for each country acting individually, these policies exacerbated international price volatility, and often penalized the domestic farmers and traders whose supplies to the market prevented more serious shortages. To make matters worse, importers' concerns about food market access were heightened by news that key rice exporters were discussing the possibility of an export cartel.[2]

Grain prices have receded significantly from their 2008 highs. But food prices remain volatile.[3] The policy focus has switched from short-term tactics for crisis management to strategies to manage price volatility and assure that consumers worldwide not be denied access to the grain they need. Global grain reserves have figured prominently in international discussions (United Nations, Food and Agricultural Organization, 2009). Proposals have been made for special emergency reserves, international reserves, and "virtual reserves" controlled via commodity futures and options trading. Some observers have also recommended regulation of commodity futures trading by noncommercial investors. Others have pressed for reductions in subsidies or mandates for biofuel production, on the grounds that such policies threaten the stability of food markets.

This paper addresses the role of grain reserves and related policies in managing grain market volatility. It is obviously important to begin with questions about the nature of the problem and its underlying causes. Are we witnessing the beginning of a new regime characterized by more volatile, if not higher, commodity prices? Is the recent turmoil in prices an aberration, involving irrational bubbles, unconnected to market fundamentals? Does it reflect purposeful manipulation by global monopolies? What have been the roles of futures and options markets, noncommercial speculators, and global international financial flows in all this?

Or is global warming already changing the volatility of crop yield disturbances, or is the world finally facing a global land or water constraint? Have fertilizer and oil prices been major causes of market gyrations? How significant is the role of expansion of biofuel supply in destabilizing grain markets?

Although many of these questions cannot be answered definitively, information is available to shed considerable light on appropriate policy responses. The purpose of this paper is, given the evidence at hand, to address the merits of the types of proposals formulated in response to the sharp price spikes experienced recently, and to focus on increasing the food security of vulnerable consumers.

Fortunately, the topic is not new. Nor are the proposed policy responses; most have precursors in programs advocated or adopted after previous periods of market instability. We can draw upon experience with previous policies, and on models that show why prices in food markets can jump so abruptly, to assess the merits of recent policy proposals.

Figure 1. UN FAO Food Price Index (Jan. 1990–Jun. 2010) (2002/04 = 100)
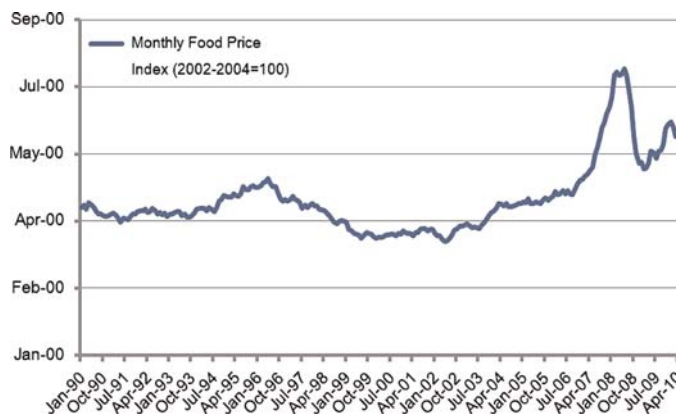
(*Source:* FAO).

2. Price Volatility: Recent Evidence

First consider the evidence about the recent behavior of aggregate food price, which was less variable than the prices of many of its components, including food grains in particular.[4] As demonstrated in Figure 1, in 2005 the United Nations FAO food price index showed evidence of a modestly rising trend that had moved the index less than 20% higher than the 1998–2000 average. In 2006 prices started to accelerate, and by October were on a sharp uptrend that continued until summer 2008, when the index exceeded twice its 2005 level.

By late summer, prices had fallen from their peaks. At year's end the index had reverted to the range observed in early 2007, still much higher than in its level at the turn of the century.

Figures 2 and 3 focus on the prices of wheat and maize. Their prices followed downward trends for decades, reflecting the fact that yields have generally out-paced demand growth, contrary to Malthusian predictions of the 1960's. Along their downward paths, prices generally fluctuate moderately within a fairly well defined range. However, episodes of steeply rising prices, followed by precipitous falls, are prominent features of the data. The price series are asymmetric; there are no equally prominent troughs in the price series to match these spikes. When price is relatively low, the probability of a sudden fall becomes negligible.

Figure 4 confirms that these features are characteristic of commodities more generally. It is interesting that the recent episode of price spikes in so many agri-cultural commodities, including minerals and petroleum, comes just over 30 years after the multi-commodity price turmoil of the mid-1970s. Note also that,

Figure 2. Price of Wheat (1950–2010) in Dollars per Bushel Deflated by U.S. CPI (1982–1984 = 1)
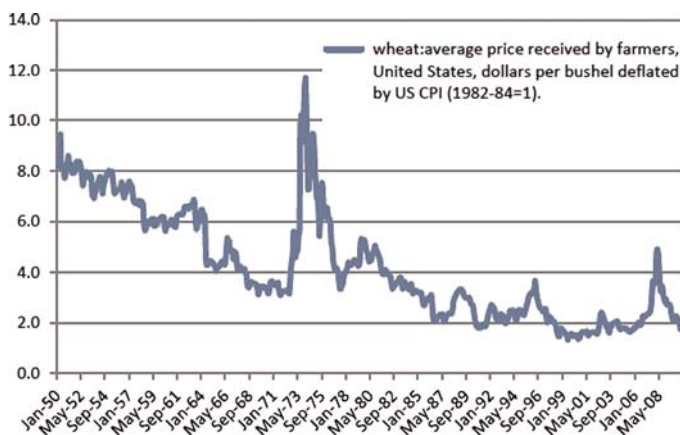
(*Source:* USDA).

Figure 3. Corn, Average Price Received by Farmers in Dollars per Bushel Deflated by U.S. CPI ((1982–1984 = 1)
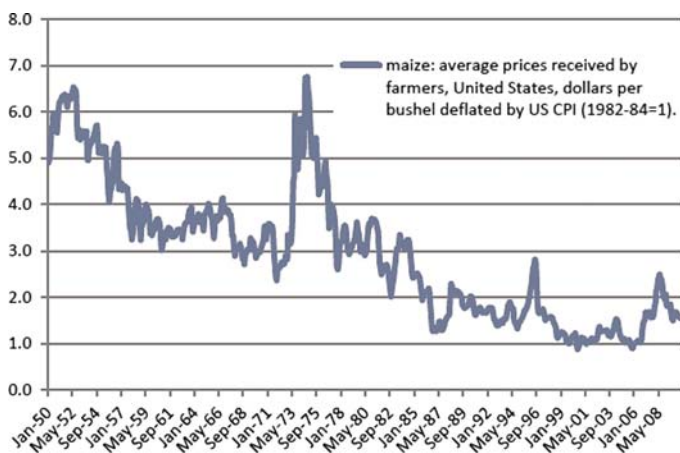
(*Source:* USDA).

relative to other grain price peaks in the figure, those of the last few years, adjusted for inflation, are not particularly high.

The overall downward trend in prices can be attributed principally to the remarkable success of plant breeders and farmers in continually developing and

Figure 4. Long Run Movements of Prices
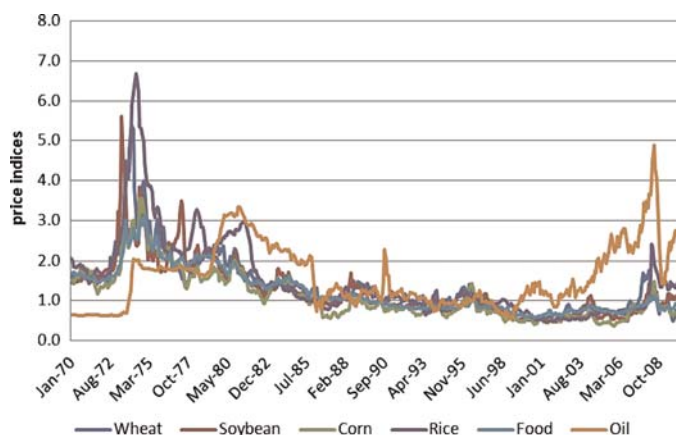
Normalized commodity price indexes deflated by the U.S. CPI.

Figure 5. Global Consumption of Grains
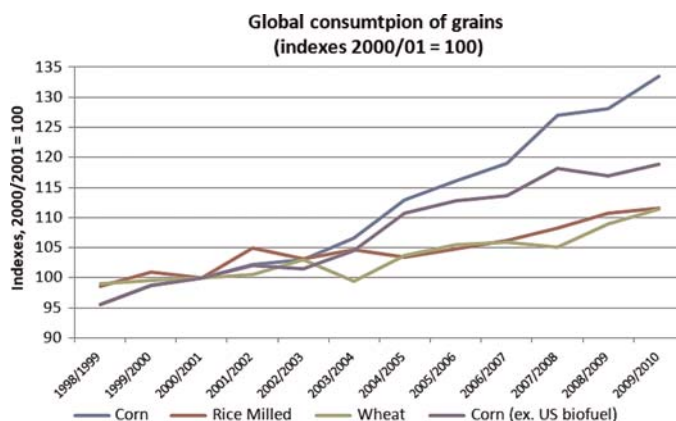
(*Source:* USDA Foreign Agricultural Service–Production Supply and Distribution Online).

adopting new crop varieties offering higher yields, and to the development of cheap and plentiful supplies of such inputs complementary to the new biotechnology. Figure 5 shows the increases in world consumption of the major grains that have occurred even as the scope for expanding the area of cultivated land has diminished or disappeared in most countries. Note also the recent large surge in diversion of maize to biofuel uses.

Figure 6. World Rice Production 1961–2010 (3rd Order Polynomial Trend)
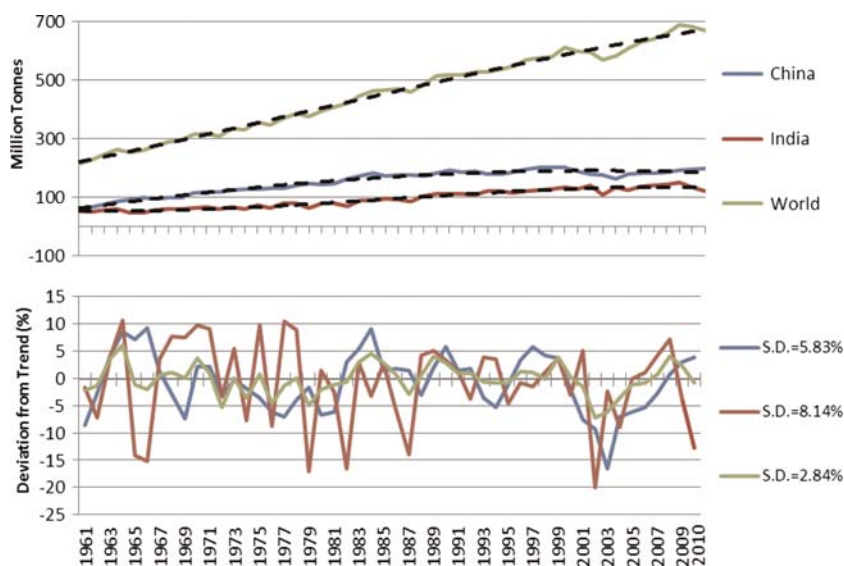
These aggregate figures mask great regional variation in prices and consumption. But globalization of markets and reduction in shipping costs offer great opportunities for smoothing local fluctuations. Figure 6 shows rice production for China and India, both major producers and consumers, and for the world as a whole. The bottom panel shows deviations from trends. Both China and India cover so many production environments that each can, to some extent, smooth out internal regional supply and demand variations via internal trade and public reallocations. Nevertheless, pooling the entire world's output variation and sharing it proportionately would reduce the variation of China's and India's shares by about 40% and 60%, respectively. For many smaller countries the effects would be far greater. These figures for wheat and maize show that the international pooling of production risks could similarly smooth national supplies. Currently, global cereal trade achieves only a fraction of these potential pooling benefits.

The trend increase in demand for grain for direct human consumption has recently been driven mainly by the increase in the global population, and the rate of increase appears to have been slowing down in recent decades. Only in poorer countries is increase in income an important driver of grain consumption per capita, which is naturally limited by the capacity of the human stomach. For grains used for animal feed, the trend increase in consumption has been greater,

because human consumption of animal products continues to rise with income long after minimum calorie requirements have been satisfied. Use of maize as an animal feed boosts maize demand far beyond what would be expected from its use as a staple food in many countries. Animal feed accounts for a smaller but still significant share of wheat production, notably in Europe. Rice is used predominantly as a food.

There is substantial agreement about the drivers of these longer run trends in grain consumption and prices. By contrast, there is a wide diversity of opinion regarding the causes of recent grain price volatility.

3. What Caused Recent Grain Price Fluctuations?

In 2008, when the rise in food prices had caught the attention of the worldwide press, observers quickly lined up a confusing array of suspects as the cause. Economists stepped in to assist in apportioning blame.

The roles played by several of these suspects are no longer controversial. These include, first, recent rapid increases in income in many countries, especially China and India, and recent neglect of crops research on a global basis. Excellent discussions of these factors are available elsewhere.[5] I do not address them here beyond noting that they could hardly have been surprises in 2007/08, except to the extent that continuation of already established trends was unexpected. Factors such as the unprecedented extension of the severe Australian drought and exchange rate movements were much less predictable. However, as noted elsewhere, their influence was insufficient to explain price spikes of all major grains of the magnitudes seen recently. Three other market disturbances that could not have been well predicted before 2007 were global in influence, and deserve particular attention. They are the changes in biofuel policies and biofuel demand, and spikes in the prices of fertilizers and fuel, which relate directly to recent price spikes in the petroleum market.

Biofuel Demand

In addition to income and population increases in the emerging economies, another currently popular suspect for aggravating recent price increases is the conversion of oilseeds into biodiesel in Europe, the United States, and elsewhere and of maize into ethanol in the United States.[6] In the United States in particular, the diversion of corn and soybeans to biofuel was increased substantially by the Energy Independence and Security Act of 2007. Biofuel use now approaches 30% for corn and 20% for soy, and will continue to increase under current policies which use subsidies and mandates, and protect the domestic biofuel industry

from competition from more efficient Brazilian sugar-based ethanol production that would place less stress on short-run food supplies.

To put the magnitude of these reductions into perspective, a drought or pest infestation that reduced United States maize output by 30% in a given year would be viewed as a production catastrophe. The southern corn leaf blight infestation of 1970, which cut U.S. corn supply by only half that percentage, was viewed at the time as a very serious shock. It directed new attention to the security of the U.S. food supply in general, and in particular the danger of genetic uniformity of a staple crop. The result was a major effort to ensure the conservation of plant varieties for agriculture and diversification of genetic resources available to plant breeders. Furthermore, relative to equivalent yield drops due to transitory disease outbreaks and weather-related shocks, the mandates for diversion of United States maize for biofuel, being quasi-permanent, and indeed slated to increase, have had much more serious implications for supplies of maize for feed and food.

On the other hand, diversion of grains and oilseeds to biofuel was not a complete surprise by 2006. To the extent that existing government mandates for ethanol use were perceived as solid policy commitments, strong demand for biofuel was clearly foreseeable before prices took off. Similarly, increased demand for oilseeds for biofuel use in Europe was no short-run surprise. In both cases, however, unexpected oil price jumps must have encouraged upward revisions in expected growth of biofuel-related demand for grains and oilseeds, as did increases in biofuels mandates in the United States in 2007. As additions to biofuel feed-stock demands resulting from previous policies, the diversions were too great to be made up in the short run by increased yields. They must have had large effects on the decreases in grain stocks, and the steady increases in prices, in the years immediately preceding 2007/08. The result was that food markets became much more susceptible to further shocks.

To substitute for maize diverted to ethanol, and oilseeds diverted to biodiesel, wheat and other food grains were diverted to animal feed. Consumers in some developing countries increased their demand for rice to replace the wheat used for feed. Some rice land might have been diverted to production of corn or soy-beans, but this is unlikely to have had a strong impact on overall rice production; the best rice land tends to be ill-suited to corn or soy production in the temperate zones where much of the world's corn and soybeans are grown. However, on Asian croplands where two or three crops are grown in succession each year, wheat can be substituted for rice as a dry-season irrigated crop when its relative price increases. In India, diversion of sugar land to rice in 2008 reportedly induced a sugar supply crisis in 2009.

Biofuel demands and surges in meat demand caused by rising incomes also affected food grain markets less directly, by diverting inputs including land and

fertilizer from some food crops to others used as animal feed or feedstock for biofuels.

Prices of Fertilizers and Fuels

Worldwide adoption of modern high-yield plant varieties and a decline in the scope for expansion of cultivated area have increased the demand for fertilizers. Prices of some fertilizers rose faster than any agricultural commodity price in the last few years, reflecting short run supply constraints, energy costs, transport costs, and a 100% export tax announced by China for fertilizers.[7] Recently, maize farmers and ethanol producers in the United States have blamed fertilizer and oil prices for jumps in grain prices.

As Figure 7 shows, prices of major fertilizers other than DAP did not really form peaks until well into 2008, after many of that year's crops were in the ground. It appears that grain prices associated with previous harvests generally preceded fertilizer price movements, rather than vice versa. Although there have been reports that farmers are reducing fertilizer applications, worldwide fertilizer supply is not likely to have diminished. There may of course have been reallocations to biofuel production and high-value crops. Reductions in fertilizer use should show up as yield or acreage reductions, but yields in 2008 generally appear to have been good.

Figure 7. Fertilizer Price and Food Price Index
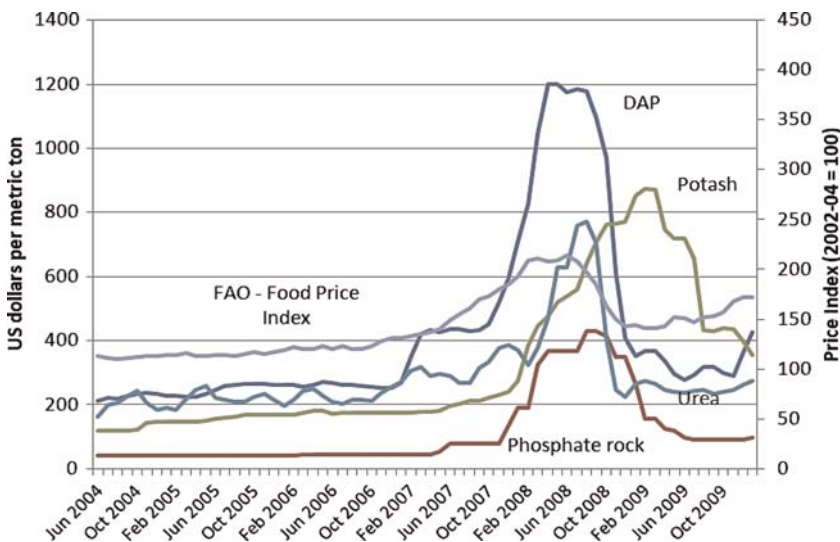
(*Source:* World Bank Prospects Group and FAO).

When prices are already high, subsidies have little effect on supply in the short run, but tend to divert global supplies from unsubsidized uses to less efficient subsidized uses, reducing overall production efficiency. Given a few years to invest in capacity, supplies can expand. But for fertilizers dependent on mineral deposits, increased demand might generate sustained higher prices and greater rents, without inducing much more production in the short run. Injudicious advice to further subsidize particular uses of such inelastically supplied fertilizers will, if heeded, certainly increase the profits of their producers, but is unlikely to increase the social value of agricultural production.

Crude oil, like fertilizer, is an important input—both directly and indirectly—into modern agriculture. Its price has been very high recently, but again there does not seem to have been a negative net effect on acreage or yield even in the countries that use petroleum intensively in production. Farm land prices in the United States rose dramatically as grain, fuel, and fertilizer prices were all rising, indicating that the net effect of all these changes on farmers' profits, and their incentives to produce more grain in the short run by any means possible, was positive and large.

The dominant effect of petroleum price jumps has been to increase demand, rather than to decrease supply. Petroleum prices shift the demand for the grain indirectly, by shifting biofuel demand. This is a new phenomenon. When ethanol production exceeds mandated levels, marginal fuel price changes increase total demand for grains even as they are raising input costs. High petroleum prices might also influence politicians to increase biofuel mandates.

From this line of reasoning, one might infer that income growth and biofuel demand should have had less influence on the volatility of rice prices relative to maize and wheat prices. However, in 2008 the price spike was actually highest for rice. Does this mean that biofuel demands had no significant role in the grain price spikes after all? To answer this question we must consider two additional, interelated factors: panic in the rice trade and inter-grain substitution by significant numbers of consumers.

Panic in Vulnerable Markets

On October 9, 2007, the Indian government, concerned about the effects of a poor domestic wheat harvest, announced a ban on exports of rice other than basmati. Large numbers of Indian consumers who eat both wheat and rice were able to substitute the rice intended for export for wheat, moderating the effects of the wheat harvest shortfall. But the ban[8] meant that the supply of exports on the world market fell, and the price of rice outside of India began to rise (Figure 8, after Mitchell (2008)). The subsequent chain of events in the rice market are discussed in colorful detail by Slayton (2009).

Figure 8. Thai Rice Price and The Indian Export Ban
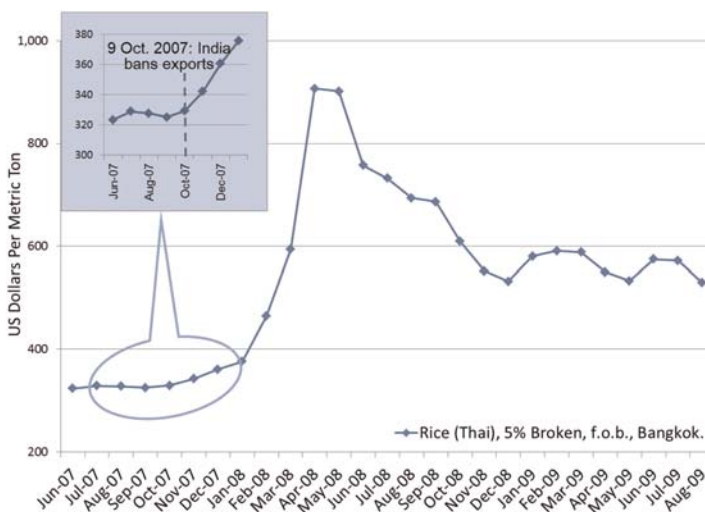
(*Source:* World Bank Development Prospects Group, and Mitchell (2008)).

5% broken = percentage of rice broken during transport; f.o.b. = "free on board"; Bangkok = where the rice is boarded.

As reports of production problems in other countries surfaced, governments of grain exporting countries were pressured by their own urban consumers to act to reduce grain prices. These pressures outweighed the interests of producers and traders in selling to the highest bidder. One by one, rice exporters imposed their own export restrictions, including, in March 2008, Vietnam, an important supplier.[9] It also became clear that China, apparently adequately supplied, would also act to insulate itself from market turmoil, rather than make its substantial grain stocks available to the international market as supplier of last resort. Key wheat suppliers also imposed export bans or taxes.

On the other side of the market, countries that relied on imports for an important share of their food became increasingly anxious to secure foreign supplies adequate for their needs so they could satisfy politically powerful urban consumers concerned about food security. Many also reduced their tariffs on imports. Reductions in import tariffs reduce domestic prices relative to world prices, but also contribute to those world prices.

One discouraging example of inadequate international cooperation on the part of a developed country importer was the failure to negotiate the timely sale, to desperate international importers, of Japanese stocks of rice, imported in reluctant compliance with World Trade Organization mandates, and never destined for domestic consumption.[10] The crisis in trade access and prices was resolved only

after it became clear, in the Northern summer, that the current harvest was good and that, overall, 2008 rice production would be close to its trend line.

Several reviews of the above influences on the grain price volatility of the past few years have allocated percentage shares of responsibility to each. This approach makes sense if the factors have a linear cumulative effect on food price volatility. But their effect is highly nonlinear. When supplies are already tight, a small reduction can cause an unusually large price increase. It makes no sense, then, to allocate percentages of responsibility for the crisis to different causes. But at the margin, alleviation of demand pressure from non-food uses has a disproportionately large effect when supplies are short. This fact is a key to understanding recent market events and constructing appropriate policy responses.

The economics of storage activity explains the relationship between grain prices and storage, and helps in the evaluation of other factors mentioned in discussions of recent grain price behavior, including distortion of futures markets by international financial flows, and an irrational or manipulative bubble in grain prices. These issues are best discussed after a review of some features of grain storage as an economic activity.

4. The Nature of Grain Storage

To interpret the behavior of grain market prices, and identify the causes of high volatility, it is crucial to understand the relation between prices and stocks. A glance at Figure 9 reveals that the wheat price spikes in the 1970s and in 2007/08 occurred when world stock-to-use ratios were low. For the market to function effectively, a virtually irreducible minimum amount of grain must be held in the system to transport, market, and process grains. Though stocks data are notoriously imprecise, minimum working stocks are apparently close to 20% of use.[11] Comparison of Figure 9 with Figure 2 reveals that stocks are very unresponsive to price at these minimum levels. Similarly, comparison of Figures 3 and 10 shows that spikes in corn price occurred when stock-to-use ratios were low.

A common feature of all such physical storage activity is that aggregate stocks are constrained to be non-negative. If current aggregate stocks (beyond essential working levels) are zero, it is impossible to "borrow from the future." Another important feature of these grains (and of most minerals) is that the marginal cost of storage per period, including physical protection, insurance, and spoilage, in practice is usually modest, and the assumption of constant unit costs is a generally reasonable approximation.[12] Increases in global grain stocks are not generally limited by storage capacity.[13]

The fact that their supply is usually seasonal is a distinctive feature of major storable agricultural commodities. For simplicity, the discussion here considers

Figure 9. World Wheat Stock-To-Use Ratios
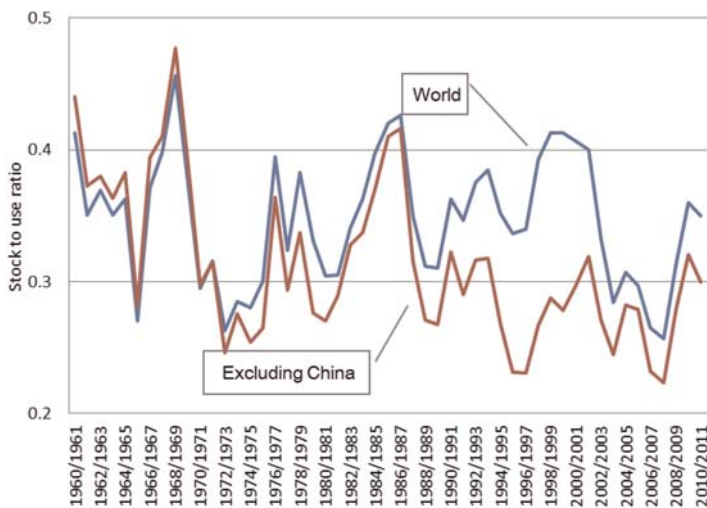
(*Source:* USDA Foreign Agricultural Service–Production Supply and Distribution Online).

annual variation and assumes a fixed interest rate. Like most studies of grain storage, the focus is on market aggregates, ignoring spatial variation and product heterogeneity, as well as national policy variation regarding trade barriers, subsidies, and taxes, all of which affect the relation between reported global prices and prices faced by consumers.[14] The observation that spikes occur only if stocks are near minimum levels reflects the constraint that intertemporal transfers via storage are unidirectional; negative storage is not feasible for the market as a whole. This reality makes modeling storage behavior interesting and challenging.

A profit is realized only if the value of the grain when released exceeds both the cost of storing it and the interest on capital.[15] Thus the value of storage today depends on its expected value tomorrow, and so on to infinity. It seems necessary to know the answer for tomorrow before solving for the problem today. Fortunately, this problem this problem can be solved by dynamic programming.[16] Here the focus is on the implications of that solution for arbitrage and grain price behavior.

5. The Economics of Competitive Storage Activity

Assume that there is one crop, sown annually. The harvest in year t, h_t, is random, due to weather and other unpredictable disturbances. The effects of storage on consumption and price of grains, illustrated in Figure 11, are the

Figure 10. World Corn Stock-to-Use Ratios
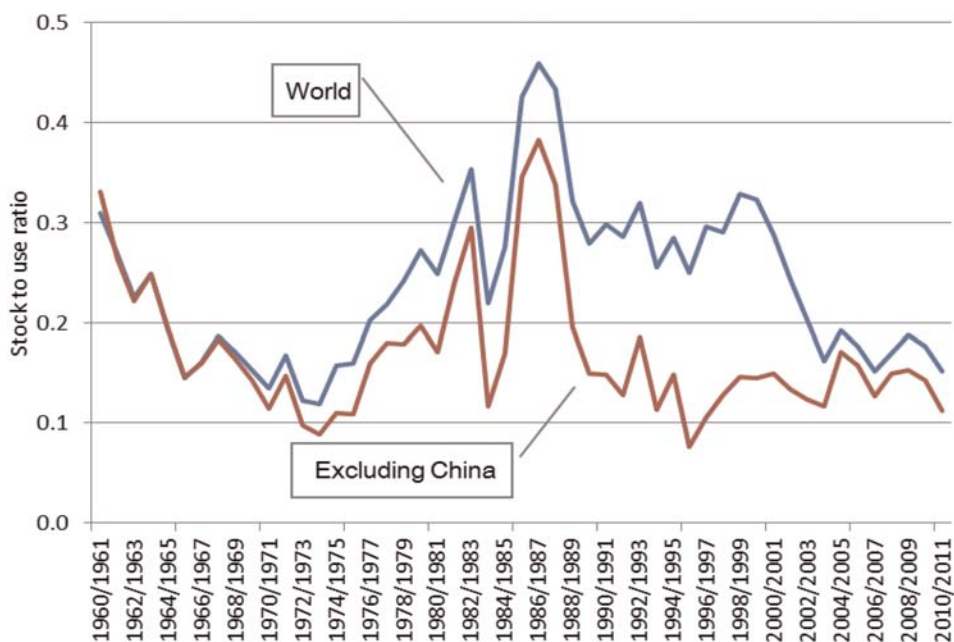
result of the horizontal addition of two demands. One, assumed to be linear in the figure, is the demand for consumption in the current period, c_t; the other is the demand for grain stocks in excess of essential working levels, x_t, carried forward for later consumption. To keep things simple, deterioration is ignored. In any period, regardless of the economic setting (monopoly, competition, state control of resource allocations) two accounting relations hold. The first defines available supply A_t is the sum of the harvest and (non-negative) stocks carried in from the previous year:

$$A_t \equiv h_t + x_{t-1}.$$

The second states that consumption is the difference between available supply and the stocks carried out:

$$c_t = A_t - x_t.$$

Assuming competitive storage, stocks x_t are positive (in excess of minimal working stock levels) only if the expected returns cover costs. (Competition

Figure 11. The Role of Stocks in Buffering Shocks
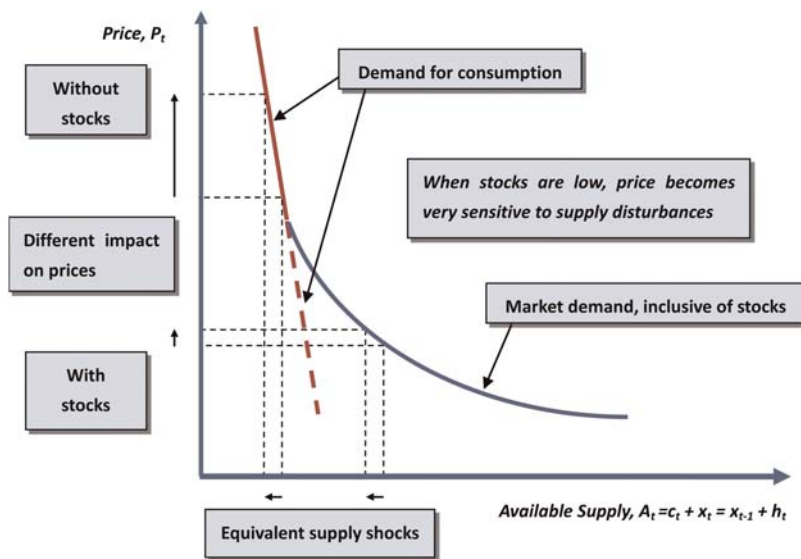
between storers prevents them from making greater profits.) This means that the current price of a unit stored must be expected to rise by a sum that equals the cost of storage k and the interest charge at rate r on the value of the unit stored. Given available supply, A_t, storers carry stocks x_t from year t to year $t + 1$ following a version of the age-old counsel to "buy low, sell high" represented by the competitive "arbitrage conditions":

Price$_t$+Storage Cost $= \frac{1}{1+r}$ (Expected Price$_{t+1}$), if stocks exceed minimum levels,

Price$_t$+Storage Cost $\geq \frac{1}{1+r}$ (Expected Price$_{t+1}$), if stocks equal minimum levels.[17]

As shown in Figure 11, when price is high and discretionary stocks are zero, the market demand is identical to the consumption demand. Those who consume grains such as rice, wheat, or maize as their staple foods are willing to give up other expenditures (including health and education) to continue to buy and eat their grain, so the consumption demand is very steep and unresponsive to price ("inelastic"); large changes in price are needed if consumption must adjust to the full impact of a supply shock. In 1972/73, for example, a reduction in world wheat production of less than 2% at a time when discretionary stocks were almost negligible caused the annual price to more than double, as indicated in Figure 2. Figure 11 also shows how, when stocks are clearly above minimum

The World Bank Research Observer, vol. 27, no. 2 (August 2012)

working stocks, storage demand, added horizontally to consumption demand, makes market demand much more elastic (less steeply sloped) at a given price.

The responsiveness of this aggregate consumption demand to price is difficult to estimate, for several reasons. One is that, in empirical demand studies at the level of the individual consumer, it is difficult to distinguish consumption from storage (including stocks held by consumers) as prices fluctuate, and when the two get confounded the estimated response overstates the consumption response. Secondly, at the aggregate level, years with high prices and negligible discretionary stocks are too rare in samples typically available (less than one hundred years) to establish, by themselves, the steepness of the consumption demand. Estimation of the dynamic storage model offers the opportunity to use data from all available years in determining consumption demand. However, the storage model has been difficult to implement empirically. One major hurdle is, again, the lack of reliable stock (or consumption) data for global markets. (In recognition of this, grain statistics refer to "disappearance" rather than consumption.) Work that pioneered the econometric estimation of this model in the 1990s, assuming no supply response, finessed the data problem by estimating the model on prices alone.[18]

Recent econometric application of a model in this tradition to prices of a set of commodities suggests that consumption demand for food responds very little to changes in the price of major commodities; the slope of the consumption demand curve for major grains may be even steeper than previously believed.[19] To compensate for the low price response of consumption, more of the commodity is stored and stocks run out less frequently. The storage implied by the model smoothes prices, replicating the kind of price behavior observed for major commodities.

By acquiring stocks when consumption is rising and price is falling, storers can reduce the dispersion of price and prevent steeper price slumps. Disposal of stocks when supplies become scarcer reduces the severity of price spikes. If the supply of speculative capital is sufficient, storage can eliminate negative price spikes *but can smooth positive spikes only as long as stocks are available*. When stocks run out, aggregate use must match a virtually fixed supply in the short run. Less grain goes to feed animals and the poorest consumers reduce their calorie consumption, incurring the costs of malnutrition, hunger, or even death.

Storage induces positive correlation in prices and is least effective when harvests are positively correlated; storage is ineffective in smoothing price changes caused by persistent increases in demand such as a mandated increase in biofuel production. Note also that the storage demand shown in Figure 11 would shift upwards, pulling total demand with it, if the supply variance rose or interest costs fell.

If producers can respond to incentives with a one-year lag, that response is highly stabilizing for consumption and price. Their competitive adjustments of planned production increase the effectiveness of adjustments of stocks in smoothing consumption and price. When supplies are large, for example, expected returns to production are low, so producers cut back production in response to lower returns, and hold more stocks.

6. The Counter-intuitive Effects of Price-Band Buffer Stock Programs

Many different policy interventions have been used in attempts to reduce grain price volatility or support price levels. These include controls or sanctions on private "hoarding" or "speculation," buffer stocks, buffer funds, strategic reserves, use of options and futures, marketing boards, and price floors, all of which obviously affect storage incentives. Other measures that can also affect storage are trade barriers, export taxes, interest rate policies, and production controls.

In the past, prominent economists supported market stabilization using a price band bounded by the floor and ceiling prices to reduce the "boom and bust" gyrations typical of commodity prices (Keynes 1942, Houthakker, 1967, Newbery and Stiglitz 1981). Since 1931 there have been more than 40 international commodity agreements. The products covered include wheat, sugar, rubber, coffee, cocoa, olive oil, tea, and jute. In the 1930s international commodity agreements were explicitly designed to address the severe problems of over-supply and low prices associated with the Great Depression by restricting exports and raising prices. They had some degree of success until the over-supply problem was eliminated by the onset of the World War II. The United States from the 1930s until the 1970s operated price support schemes involving buffer stocks of major commodities and in the European Union storage-related programs to support and stabilize prices have been part of its Common Agricultural Policy.

A major element of the economic doctrine heralded as the "New International Economic Order" by the United Nations Conference on Trade and Development (UNCTAD) was negotiation of international commodity agreements (ICAs).[20] Important programs were directed at sugar, coffee, cocoa, tin, and rubber. The first two of these, like the pre-war agreements, managed storage only indirectly via commitments to control exports, but the others involved attempts to control prices using versions of price-band schemes. When the price fell to the floor of the band, acquisitions were to be made; when the price reached the ceiling, stocks were, if available, released from the stockpile by the program's management. A later Australian wool reserve price scheme acted more like a floor price scheme with a variable release

price and a buffer stock. Because of the distinctive nature of Australian wool, this program was effectively a global program in its effect on the market.

International agreements involving commodities, including rubber, cocoa, and tin, have often combined the floor price with a higher "ceiling" or "release" price, a plausible way to protect consumers from the most extreme effects of price spikes. Policy makers find such "price band" policies attractive because they seem simple and easy to explain. An appealing intuition is that such a program keeps the price around the middle of the price band most of the time, and affects the market mainly in unusual periods, if the band is judiciously chosen. But numerical examples made possible by advances in computing and dynamic programming, not available in the early 1970s, show that this is not true.[21] As illustrated in Figure 12 using a simple numerical market model[22], for a program with a floor that is 87.5% of the mean price of $100 and a ceiling set at 112.5%, the program greatly reduces the probability of spikes above the ceiling. But the probability that the price will be at or above the ceiling is greatly increased, to 30%, and there is a probability of about 15% that the price will be at the floor. Relative

Figure 12. Price Probabilities Under a Price Floor and a Price Band
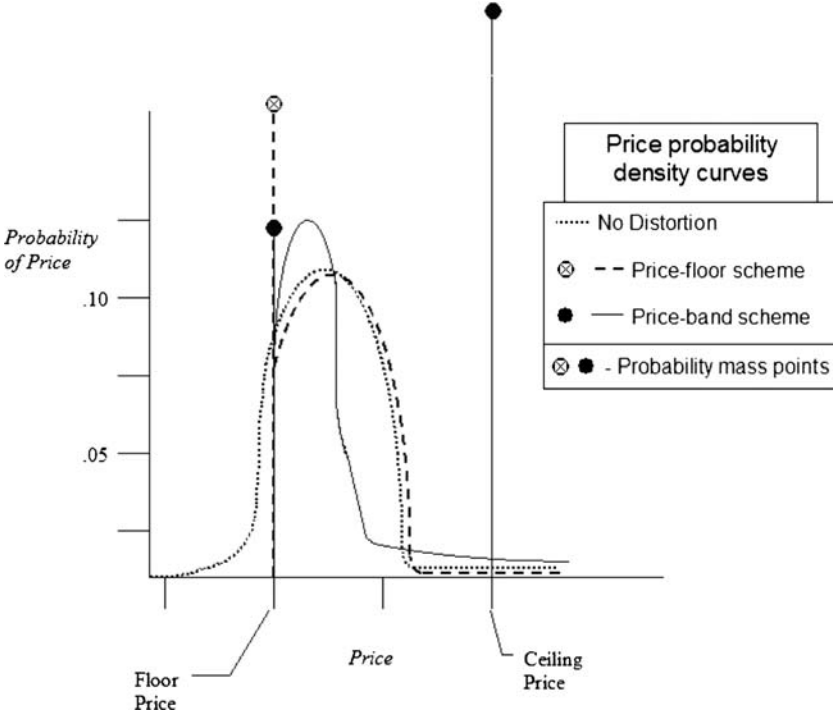
to a free market with storage, there is a much lower probability that price will be located between the mid-point of the band and the top.

Most of the time, the market appears to be "challenging" either the floor or the release price. The price ceiling discourages production and storage and increases volatility of the price as the latter approaches the ceiling. Paradoxically, the price is much more likely to be near the center of the band under the free market, or where program stocks are made available for release to private storers and consumers at the floor price.

Another serious consideration is budget cost. When a program chooses a price floor p^F that is no higher than the free-market mean (adjusted for a perfectly estimated trend if necessary) or a price band where the mean of the floor and ceiling price equals the free-market mean, the program has frequently been assumed by economists (see, for example, Newbery and Stiglitz 1981) to be "self-liquidating"—that is, financially sustainable, based on the fact that expected net balances should equal zero and the intuition that the summed funds from purchases and sales after several years of operation should be close to their initial values. But this intuition is wide of the mark even for a simple floor price scheme in a market with no underlying trend.[23]

The fund may in the short run accumulate great profits, appearing to affirm the manager's skill and to belie the skepticism of "theoretical" economists, inducing pressure to raise the price floor. Such pressures can be very difficult to resist.[24] Even if the manager can commit to the original rules, any given operating reserve will be depleted in finite time.

In practice, postwar experience has affirmed that the "finite time" within which such programs fail is disconcertingly short, often less than a decade or two. Recent failures in programs for tin and wool, among others, have shown that the largest price effect of these interventions can be the severe price collapse that accompanies their inevitable failure.[25]

When such price support programs do fail, there is generally a public consensus that the intervention price was wrongly set; management is often blamed for faulty trend forecasting. There is scant recognition that failure is inevitable at any relevant intervention price even if the fundamentals are stationary. Higher floor prices merely advance the time of reckoning. Price band programs tend to fail sooner because they tend to accumulate stocks at a faster rate.

The attraction of price bands might well be at least in part due to the failure to appreciate the potential of competitive storage. To illustrate the latter, it is necessary to use a numerical dynamic model of competitive storage. Figure 13[26] illustrates three probability densities for prices conditional on current prices at, respectively, 74%, 94%, and 114% of the mean generated by a numerical model of competitive storage. In this example, if price is 94% of the mean, there is virtually no chance it will be below 70% of the mean the next year. If, after a string of

Figure 13. Probability of Price in Period $t + 1$ When Current price P_t is 74%, 94%, or 114% of the mean price of $100
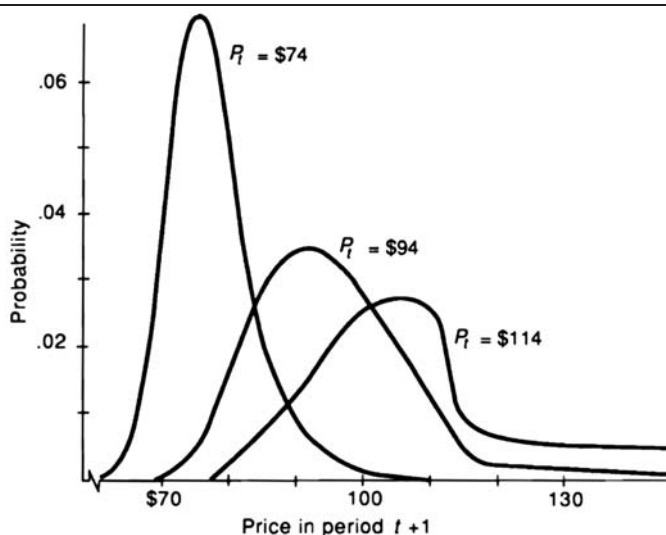
good harvests, the price does eventually fall to 70% of the mean, there is virtually no chance it will fall below 60% (or rise above 110%) the following year. Note also that if the price is 114% of the mean the figure indicates a much larger chance of a lower price than a higher price the following year. There is a modest right tail indicating the probability of a price at least 14% above the mean but the model is acting much like an imperfectly effective price-band program with a floor around 65% and a ceiling around 114% of the mean price.

In sum, much of the stabilizing benefits of a price-band scheme are furnished by competitive private storage in a free market in which there is no fear of punitive measures against "hoarding" or other perceived offenses. Price-band schemes in theory are bound to fail if the bands are not adjusted to reduce losses. In practice, failure comes fairly quickly. If, on the other hand, bands are adjusted to reduce accumulation of losses, the program tends to mimic what the free market can provide. Price-band schemes are unsustainable and expensive, in theory and in practice, and can be hugely destabilizing when they fail.

7. Public Policy for Grain Supply and Food Security

Since ancient times, national leaders have recognized a responsibility to ensure adequate domestic availability of staple foods. For example, the Ch'ing Dynasty in

China maintained a nationwide granary system with responsibilities that included moderation of seasonal fluctuations and famine relief.

Intervention in markets for staple foods is still prevalent, even in modern capitalist economies. Why is this so? Surely an undistorted free market could equalize the marginal value of a given grain supply across alternate uses, including placement in storage?

In a free market, only those who have the necessary resources or "entitlements" can acquire food. The needs of the destitute may not affect prices at all. Whether or not governments have any sympathy for the plight of the poor, only the most totalitarian are able to ignore pressures from consumers mobilized by concerns for their own consumption needs. In response to this temporarily powerful constituency, governments often force traders who have accumulated grain to surrender those stocks to the government or directly to consumers, often without compensation. Such so-called hoarders are typically vilified, and sometimes also punished or even killed. In such emergencies, the argument that the "hoarders" might be the sole source of supply if the next crop fails gets scant consideration.[27]

Anticipation of such treatment understandably discourages private storage for distribution at a high price in time of need. Even if a government commits not to confiscate stocks (or otherwise penalize hoarders) in emergencies, a commitment against all intervention that would discourage speculation is not credible. Hence governments often choose to supplement private storage with publicly acquired stocks or storage subsidies. (Even if the government manages all market stocks, it is difficult to prevent consumers from storing some domestic supplies.) When public stocks are released to consumers (other than those with no money at all for food) they will, to some extent, have a negative effect on prices. Anticipation of this price effect reduces private storage incentives below those offered by a free market. Hence it is natural to expect that governments will intervene actively when supplies are plentiful to increase grain stocks and thereby help ensure supplies for the needy and/or stabilize the market.[28]

Before assessing specific grain market interventions, it is useful to be aware of the following facts:

(1) Any activity or policy that does not change consumption in a market does not affect prices in that market. On the other hand, if a policy decreases price, it increases consumption and decreases stocks. If planned production is responsive, it also decreases when the price drops, unless the spot price is so high that there are currently no discretionary stocks.

(2) If they fail to address the fundamental source of disturbance (for example, disease, war, arbitrary policy initiatives or weather), "stabilization" policies must actually destabilize some key variables (stocks or public budgets, for example) as they stabilize others (such as price).

(3) There is no evidence that any chosen group of experts, no matter how well qualified and motivated, can reliably determine when a competitive market is acting in a way not justified by fundamentals. The general proposition that designated experts can outperform the market in forecasting or trading might have been plausible in the time of Keynes, but a large body of empirical evidence to the contrary has accumulated in the intervening decades. The best-informed international organizations concerned with food markets for the poor (including the World Bank) wisely make no claims of superior forecasting capacity.

(4) In any intervention, net efficiency gains to society as a whole are typically dwarfed by redistribution of gains and losses between producers and consumers. Those who most enthusiastically and effectively support storage interventions naturally tend to be the ones who expect to gain from those policies. To comprehend these distributional effects, it is necessary to recognize the dynamic nature of the problem and the importance of private responses to public actions.

With the above points in mind, let us consider several recently discussed policy initiatives:

A Proposed International Coordinated Global Food Reserve

The recently evident failure by many grain exporters (especially in the rice market) to commit to offer uninterrupted market access to their supplies has highlighted the desirability of commitment-reinforcing mechanisms for international grain market participants. One such mechanism, an international coordinated global food reserve, has recently been proposed.[29] The rationale for this reserve is to reassure importers that they could rely on exporters to supply them in time of need. The proposal is sketched as an agreement by members of a "club" that would include members of the G8 + 5 plus major grain exporters such as Argentina, Thailand, and Vietnam. Members would commit to holding specified amounts of public grain reserves in addition to reserves held by the private sector. The public stores would be used for emergency aid as directed by the World Food Programme.

A Proposed Global Virtual Reserve

A related proposal is for a global "virtual reserve." Nations that are members of the "club" would commit funds amounting to US$12−20B to be provided, if necessary, to the high-level technical commission for operations in the futures markets. One version of the proposed intervention characterizes it as a dynamic price-band system[30] operated by a "global intelligence unit" that apparently is assumed to have superior forecasting ability, and can reliably detect when the market price has departed from levels supported by fundamentals.

By operating via long futures positions, the scheme would aim to induce a buffer stock indirectly, by raising future prices and thereby inducing increased private stockholding. This virtual scheme, if large enough to move markets (and if allowed under the rules of relevant commodity markets), would require ready access to large and in fact indeterminate amounts of margin financing, and be subject to manipulation by traders. This initiative ignores a major achievement of empirical econometrics in economics and finance in the decades since UNCTAD advocated buffer stock programs as part of its New Economic Order, namely the accumulation of evidence against the proposition that a group of "experts" can reliably outguess the market. If, as we have every reason to believe, its global intelligence unit does not in fact have superior forecasting ability than the market as a whole, it will lose money on average, and will eventually exhaust its budget, like schemes with similar ambitions dating back many years. One example, reviewed in Peck (1976), is the Federal Farm Board's intervention in the United States' cotton and wheat markets using futures contracts to try to stabilize prices in the face of a bear market during the Great Depression. This stabilized American wheat prices for a year or so before essentially owning the United States' wheat stocks and losing $188 million—a great deal of money in the 1930s—and being disbanded. Regional supplies were severely distorted even within the United Stated market, creating shortages in some localities and gluts in others, an unanticipated collateral effect of relevance to modern proposals for price interventions. For a multilateral program, another major challenge for such a commitment-reinforcing program is to ensure commitment by the participants themselves to honor their obligations when markets are under stress.

In another interpretation reflecting written sketches by von Braun and Torero (2009) and Robles, Torero, and von Braun (2009), the operator would not attempt to operate a price band, but would stand ready to take naked short positions (not backed by stocks or prospective harvests) when a disequilibrium price surge is reliably detected. The idea appears to be that this action would convince speculators to sell their discretionary stocks, and thus reduce prices. Apart from the problematic and unverified assumption of superior information, one must recall that, as noted above, all recent grain price spikes have occurred when there were almost no stocks available for speculators to have held and later released.

Futures Market Regulation

In any grain price crisis, futures and options traders get blamed sooner or later. This happened in the United States, for example, in the last century when many forms of futures and options trading were banned and it is happening again now.[31] This time, the critiques come with novel twists.

The major criticism focuses on the entry of new money from (1) index funds holding persistent long positions (contracts to purchase grain in the future at a set price) and managing those positions by rolling the hedges over to later maturities or increasing or decreasing their positions to maintain portfolio allocation shares, and (2) speculative investors such as hedge funds. The argument is that these long positions have added buying pressure, raising prices for the physical commodity above the levels justified by supply and demand.

For United States futures markets, the facts tend to contradict the assumptions underlying this critique.[32] First, for soybeans and maize in particular, short-hedging by producers, merchants, and processors grew more from 2006 to 2008 than did long speculation. For wheat, the increase in long speculation was greater but the relative magnitudes stayed within normal ranges.[33] Second, the commodities for which index investment grew most over the two years saw no significant price increases. Third, commodities neglected by index funds (such as rough rice and fluid milk) experienced large price increases, as did commodities with no futures markets at all (apples, edible beans). Fourth, index funds, if operating as advertised, rebalance as grain prices rise, reducing long positions to maintain portfolio shares, and thus stabilizing prices somewhat like a more flexible variant of a price-band policy. Fifth, empirical work has shown no significant evidence that position changes by speculators help forecast price changes in these markets.[34]

Finally, if long futures market positions exacerbated price spikes 2007/08 they must have reduced consumption and increased commodity stocks. But stocks were around minimal feasible levels. To the extent that speculators might have influenced the market by increasing stocks in previous years, their unwinding of those positions should have increased consumption and moderated price, hardly undesirable effects.

Policies to Prevent Irrational or Manipulative Bubbles

The reality that overall grain availability increased prompted a second and quite different rationalization of the crisis in the grain markets: there were irrational or manipulative bubbles attributable to "greedy" speculators that burst in the spring and summer of 2008. In 2007, one story goes, prices got out of line in the grain markets and supplies were withheld in anticipation of greater profits later. The sharp reversals of grain price trends in different months of 2008 are viewed as confirmation of this interpretation: the "bubbles" proved unsustainable, as bubbles always are, and burst. Given the recent history of financial markets, an explanation dependent on greed and irrationality is both plausible and appealing.

Unfortunately, recent research on models of commodity markets like the one represented in Figure 12 but with slightly different, though hardly unconventional, demand behavior has shown that irrational bubbles are difficult if not

impossible to distinguish from rational investment behavior by nonmanipulative market participants, just as greedy investors are appear to be indistinguishable *ex-ante* from regular profit maximizers.

There is another reason to discount the need to prevent bubbles. If a bubble occurred in a grain market in 2007/08, to affect price it must have increased stocks. But, as previously noted, stocks were at or close to minimum levels. Where were the increased stocks to be found as prices rose to their peaks? Moreover, had such stocks existed, would it have been prudent, *ex-ante*, to force the release of scarce stocks if there were no guarantee that the next harvest would be better?

Controls on the Investment of Excess Global Liquidity

A related set of arguments points to the entry of holders of new and cheap capital into commodity futures markets in the past few years as a key cause of grain price spikes. One part of the argument has some plausibility and is favored by respected researchers in international finance. A brief sketch goes as follows. A large pool of global capital accumulated largely in China was invested in the United States housing market until that market collapsed. Hoards of these global dollars, seeking new targets, were dumped into the commodity markets through hedge funds and other investment vehicles. These new dollars caused commodity prices to soar.[35]

All but the last sentence is plausible. The real cost of capital to major financial and commodity markets was low until the United States financial sector descended into disarray and international dollar surpluses were a part of this phenomenon. As previously noted, lower interest rates tend to be associated with higher stocks, higher current prices, and lower futures prices. But the facts regarding key agricultural commodity market behavior just quoted fail to imply any causal relation between the cash inflow and commodity price spikes. This is not surprising. No one has demonstrated that this cash increased grain stocks when, as previously noted, stocks were around minimal feasible levels for normal market operations. As previously noted, if the cash inflow did not increase stocks, it cannot have reduced consumption or raised the market price in the short run. If it did increase stocks earlier, their release before the price spiked must have moderated the price increase and smoothed consumption.

## 8. Recent grain price spikes: A reappraisal

If international income growth, population growth, futures market speculation, irrational bubbles and global financial flows do not explain the recent grain price spikes, what does? Why were they so large? Were they caused by the oil price

surge shown in Figure 4? Were they irrational bubbles, unrelated to fundamentals, after all?

An important part of the answer is that the spikes, appropriately deflated, were not unusually large. Look again at Figures 2 and 3. There were comparable spikes around 1996—smaller for wheat, larger for maize. Another glance at Figure 4 shows that those spikes were clearly unrelated to oil prices, which were stable around that time. They could hardly have been caused by index fund investment—one of the two major indexes was not even in existence then.

A more promising line of investigation is suggested by Figure 14, which shows world stock-to-use ratios for the sum of the three major grains (corn, wheat, and rice).[36] Around 1996, the world aggregate stock-to-use ratio was much higher than recently. But the world figure was distorted by the huge holdings of China, whose exports were negligible in that period. If China's effect is removed, the ratio around 1996/97 looks as tight as observed in 2007/08. The lack of stocks in both episodes left the market susceptible to large price spikes from small supply disturbances. One possible objection to this assertion is that the ratio was about as tight around 2002-2004 and yet the price changes observed then were much smaller. But in that period, in contrast to the other episodes, China made substantial exports of maize and rice, increasing available supplies in the global grain market. Thus the recent history of grain markets supports two conclusions. First, the price spikes of 2008 are not as unusual as many discussions imply. Second,

Figure 14. World Stock-to-Use Ratios for the Sum of the Three Major Grains (Corn, Wheat, and Rice)
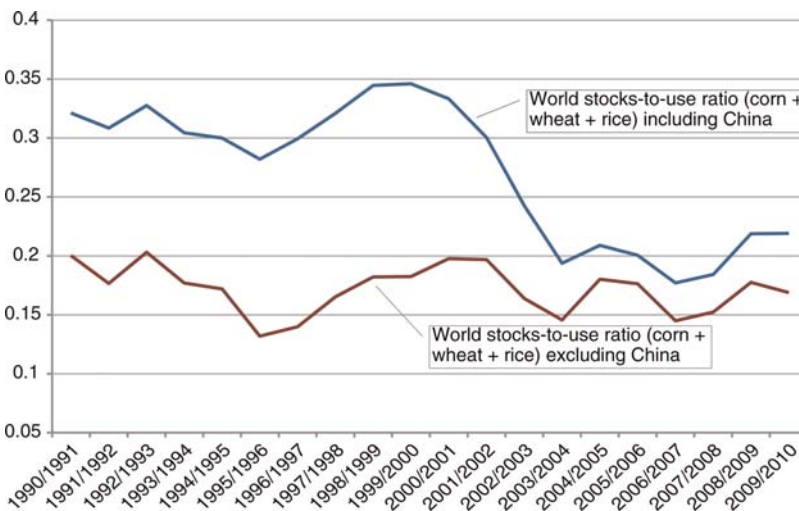
(*Source:* USDA Foreign Agricultural Service–Production Supply and Distribution Online).

the balance between consumption, available supply, and stocks seems to be as relevant for our understanding of these markets as it was decades ago.

9. Policy Responses to Ensure Adequate Consumption

The evidence reviewed above points to the key role of stock levels in recent market volatility. The following policy options address this age-old problem in several ways, with the general objective of mitigating the effects of volatility on the most vulnerable consumers. Some are variations of programs that have been implemented in the past. Others informed by recent experience, or motivated by the new challenge posed by use of agricultural resources for liquid fuel production, are more novel and less time-tested, but well worth considering.

Emergency Food Reserves to Stabilize Consumption of Vulnerable Groups

If such a reserve is successfully targeted at a small part of the aggregate consumer market, it should not have a major effect on prices in the broader market. Operation of disaster relief programs typically requires reserves to be on hand to ensure a smooth and timely response to food supply emergencies and related humanitarian disasters.[37] One would anticipate that this type of stock would be used for local and regional food shortages, often in landlocked countries or failed states. Such shortages are usually unrelated to global market conditions, and the stock is of smaller magnitude than needed for a global price stabilization scheme, so the exporter commitment problem previously discussed is less serious, though still a serious issue. Recent difficulties involving lags in food aid responses and mismatches between years when aid is plentiful and years when it is needed might be alleviated by such a reserve. On the other hand, care must be taken to minimize disincentives caused by the price-depressing effects of food distribution for the local farmers and merchants who are the first line of defense against famine for such countries[38].

The reserve could be useful in improving the speed and flexibility of short-run responses to local food crises. But its operation presents many challenges familiar to administrators of aid programs. For example, measures should be taken to ensure that transport will be available for delivering this aid, especially for landlocked countries such as those in Africa that have recently encountered food crises. It seems likely that direct assistance to the neediest, where feasible, would be more effective than attempting to reduce prices by supplying extra grain to regular food markets. Public employment programs for those needy who are able to work have been successful in cases where it has been possible to keep the reward for work low enough to be unattractive to those with other employment

alternatives.[39] A modest emergency reserve of this type could be crucial for improving responses to local humanitarian crises. However, its impact would be negligible on the global market volatility that is the focus of this paper.

National Strategic Reserves to Stabilize Consumption

Such reserves are designed to ensure adequate national consumption in those (hopefully infrequent) occasions when a country finds itself cut off from its regular access to food imports. Thus they will affect national prices in emergencies, but should not eliminate incentives for private stockholding.[40] One reason that grain prices have not declined further from recent peaks is that many countries are rebuilding or expanding their grain reserves in reaction to the export bans and export taxes observed recently.[41] Such actions appear almost inevitable at the national level given the inability of exporters to commit to being reliable suppliers in emergencies. According to a recent report, the United Arab Emirates, presumably capable of offering a logical food-for-oil deal, were unable to obtain blanket assurances from Pakistan that grain produced from the Emirates' planned agricultural projects in that country would not be subject to export controls.[42] Futures contracts eliminate counterparty risk but can expose countries to location-basis risk and sudden large margin calls. Further, a futures market might be shut down or exports banned; both actions were taken in India in 2007 at a time when the situation in its grain markets fell far short of emergency conditions.

A key question is how large the reserve should be. The answer must depend on the facts of each case, including the diversity of food supplies, dependability of traditional suppliers, and cost of the program. Such stocks tie up capital for the substantial intervals between releases and can be expensive to maintain, especially in humid tropical countries.[43] Their efficient management also uses scarce human capital and temptations for corruption can easily arise.

If the public stock's management can commit to hold the stocks for release only in circumstances in which private stocks would be exhausted, the disincentives to storage by the private market can be reduced. For a landlocked country, this type of emergency situation might be the second year of a severe drought. For an importer, it might be the second year of a global shortage. In such real emergencies, releases of stocks via direct distribution outside the market can be targeted to ensure that all consumers receive what is minimally needed, as previously discussed for the case of the small emergency reserve. A release policy designed to operate via its effect on the general market price is likely to be more costly and less effectively targeted to those in need.

Thus the national storage activity discussed here is appropriately directed at a stockpile of a certain size deemed appropriate to meet security goals rather than aimed at modification of the behavior of prices. In contrast, many international

commodity agreements and some programs proposed recently are targeted at market-wide price behavior rather than targeted consumption goals.[44]

Besides measures affecting storage activity directly, other policies might be considered to reduce market volatility and/or increase market access. Some of these have substantial merit; others do not. We now turn to several of these, starting with the more promising.

Improvements in Availability of Critical Information

One striking feature of recent chaos in grain markets is the paucity of timely data on available stocks in each country and particularly in Asia. Earlier and more accurate data can reduce volatility, improve planning, and encourage international confidence and cooperation. Until now, key national participants have treated their stock data as a national secret and a source of commercial advantage. Policies that facilitate communications between private traders have great potential for preventing famine in isolated markets. Fortunately, improvements in Global Information Systems are improving global access to information on weather, production, and stocks without the need for international collaboration on data sharing. Aker (2008) has shown that spatial price and supply variation in Niger during the recent famine was moderated by the adoption of cell phones by key traders, as they became available.

Commitments to Refrain From Using Export Restrictions

Recent experience in the rice market has demonstrated the hazards associated with reliance on imports to satisfy needs for a staple commodity. Exporters and importers have a joint interest in keeping trade open when prices are high so they can together reap the full benefits of the smoothing role of trade, which can exceed what can be achieved via storage. But commitments to do so are difficult to achieve and can easily collapse due to pressure from politically powerful urban consumers. One useful policy change to improve the commitment capacity of exporters would be a reform of WTO disciplines on export bans and export taxes consistent with existing rules against import tariffs and quotas. Whether such a reform is feasible is a question I leave for others to decide.

Creation of Options to Divert Grains From Biofuel and Feed Uses in Emergencies

Modern food markets are, in an important sense, more inherently stable than their predecessors. Now, a significant portion of the domestic supply food grains and oilseeds is used for biofuel in many countries, and for large-scale animal feeding in many more. The increasing non-food uses for grains increases the

pressure on food supplies, but it also offers a new source of emergency supplies in food crises. In such circumstances, it should be possible to ensure diversion of some feed grains and oilseeds from use as animal feed or biofuels feedstocks to domestic use as food distributed to vulnerable consumers, without undue hardship to the generally more prosperous consumers of substantial quantities of energy or meat. (Commitments for *international* diversion are much more problematic.) Similar contracts have been used, for example, to ensure secure urban water supplies in the United States by diversion from irrigation during droughts, including diversion of irrigation water to urban consumption. (See O'Donnell and Colby 2009 for a guide to such contracts, and Hansen and others 2008 for other references.) On a different time scale, interruptible electric power contracts are commonly used for industrial customers willing to relinquish claims on electricity when net supply is low.

The food supply authority could purchase call options on grain from biofuel producers, most likely directly, with performance guarantees, as trade volume is unlikely to support an organized exchange. Diversion would be triggered by specified indicators of food shortages, and the biofuels supplier or animal feeder would commit to make a corresponding reduction in output (rather than substitute other food grain as feedstock). Delivery specifications could be designed to ensure the grain will get to where it is needed in a market emergency. All parties can gain from implementing such contracts.

If biofuels mandates inhibit such diversion, they should be altered to allow for use of such options. Better yet, biofuels mandates should be made conditional on food prices or availability. But the conditional mandates are not sufficient to protect consumers. If petroleum prices soar, biofuel demands could trump those of poor food consumers. The proposed options would protect consumers in such circumstances.

If biofuel feedstocks are sourced from permanent stands of miscanthus or other perennial grasses with low feed value, rather than from annual grains, this potential flexibility could be lost. If biofuel conversion of such inedible crops becomes more efficient, producers may well be tempted to increase the area planted to them. In that case, the threat of biofuels to food supply security could become much more serious than it is at present, and diversion of animal feed would become more important.

10. Conclusions

The storability of grains causes the price response to a change in supply to vary with the level of available supply. The major grains—wheat, rice, and maize–are highly substitutable in the global market for calories. When their aggregate

supply is high, a modest reduction can be tolerated with a moderate increase in price by drawing on discretionary stocks. But when stocks decline to a minimum feasible level, similar supply reduction can cause a price spike. In a free market, poor consumers with little wealth may be forced by high prices to spend much of what resources they have on food and reduce consumption at great personal cost. Others reduce consumption very little even when prices soar.

In 2007/08 the aggregate stocks of major grains carried over from the previous year were at minimal levels due largely to substantial mandated diversions of grain and oilseeds for biofuel and strong and sustained increases in income in China and India. Lack of stocks rendered the markets vulnerable to unpredictable disturbances such as regional weather problems, the further boost to biofuel demand from the oil price spike in 2007/08, and the unprecedented extension of the long Australian drought. However, supplies were sufficient to meet food demands without jumps in price, had exporters not panicked, leading to a cascade of export bans and taxes that cut off importers from their usual suppliers.

If in future food shortages more serious supply problems arise, there is little doubt that export bans will recur. Governments that recognize an obligation to protect poor consumers or are sensitive to pressure from consumers will intervene. Exports will be taxed, cut, or banned, distorting private storage incentives and cutting off importers' access to supplies. Given these realities, there is a case for public interventions when supplies are more plentiful in anticipation of future crises.

Deflated prices of food grains follow long-run downward trends interspersed by episodes of steep price increases immediately followed by even more precipitous price falls. Relative to other episodes of grain price spikes, volatility in the real grain price the past few years has not been particularly high. There is no evidence of a change in the global grain price regime.

Their experience in the grain markets in the past few years has encouraged many governments to build or expand national grain reserves. If such reserves are aimed at ensuring minimal levels of consumption, they should be designed to meet the needs of vulnerable consumers by nonmarket distribution in emergencies. Decisions about their size should reflect both the advantages of secure supplies and the substantial costs of acquisition, storage, and administration.

The recent food price spikes have led to several proposals for international intervention in commodity markets. One suggests that creation of a small emergency reserve to respond quickly to regional emergencies would help speed up responses by international organizations in aiding groups in distress. The free market cannot be relied upon to service this need, for such groups lack the resources to bid for the food they require. Since regional emergencies often involve landlocked nations, contingent transport contracts may be useful to ensure adequate and timely distribution of stored grain.

A large international grain reserve, held at optimal locations and controlled jointly by national governments to mitigate global food supply crises could economize on stocks and storage costs in providing a globally adequate amount of storage and help maintain the valuable stabilizing role of free international trade in grains during emergencies. Unfortunately, such an ambitious scheme appears to be infeasible without improved means of guaranteeing continued international collaboration by the participants during food emergencies. Stronger WTO disciplines on export tariffs and adoption of disciplines on export bans are would increase incentives for collaboration, but are unlikely to be persuasive in serious food price crises.

Other recent responses to the events of the last few years include proposals for a combination of international physical reserves provided by members of a group of national participants and "virtual" reserves to control speculative price behavior in grain markets. In at least one version, the interventions would be naked speculative short positions taken when a global intelligence unit using special knowledge unavailable to the market decides, using criteria not identified, that prices do not reflect "fundamentals." Similar proposals made many years ago were easier to take seriously. In the last half century, a large body of work including theoretical and empirical analyses has shown how difficult it is, even for top experts, to be sure that markets are out of equilibrium and that proposed price interventions will do more good than harm. Naked short speculation to stabilize prices is very risky and indeed could quickly lose vast sums of money, especially if positive initial results increase the confidence of management, encouraging decisions that lead to greater financial exposure.

Use of price-band rules to operate international or domestic market stabilization schemes is less simple than often assumed and less effective in ensuring food security for those most at risk. The price tends to hover at or near the upper or lower band, private storage is reduced or eliminated, and production is discouraged just when it is most needed. Theory predicts, and experience confirms, that these programs inevitably fail even if there is no underlying trend in price.

The recent history of the markets for major grains highlights the need for greater caution in adopting policies that subsidize or, worse, mandate further diversion of grains or grain-producing land to biofuel. Abrupt increases in diversion to biofuels can induce serious price spikes when stocks are low, threatening the security of grain for consumption by the world's most vulnerable consumers, and continued diversions will lead to increases in the levels of food prices; poor consumer will pay a price for biofuels consumption by others.

On the other hand, the reality that substantial quantities of grains and oilseeds will continue in the near future to be converted into biofuel or animal feed in many countries suggests a new strategy to ensure that the most vulnerable consumers have access to sufficient food (as distinct from the goal of stabilizing

market price). Governments of nations with substantial domestic biofuels production or livestock feeding industries should seriously consider the purchase of diversion options form producers of biofuels or meat products, much like dry year options in water markets. These option contracts would give government the right but not the obligation to acquire, in serious pre-specified food supply emergencies, domestic grains or oilseeds that would otherwise be allocated to biofuel production or animal feed. These grains could then made available as food for consumers (or substituted for other typed of feed grains more attractive to human consumers). All parties could gain from such diversion options, which could be written as contracts with specific biofuels producers; they are not necessarily dependent on the existence of a commodity exchange.

Notes

1. See Slayton (2009) for a colorful account of the rice market in this period.
2. http://news.bbc.co.uk/go/pr/fr/-/2/hi/business/7379368.stm (last accessed July 23, 2010).
3. In June 2009, wheat prices surged to their highest levels since October 2008.
4. Although we must focus on aggregate numbers here, it is important that they mask a tremendous amount of variation between countries, due to trade barriers, exchange rate movements, domestic price and tax policies, and transport costs. As trade barriers, tariffs and transport costs have changed abruptly, the scope of various international markets has also been redefined. Furthermore, in large or landlocked countries international prices often face widely varying prices; for many consumers, international prices and policies discussed here have little relevance, as noted below.
5. See Abbott and others (2008, 2009), Mitchell (2008), Timmer (2008), and Gilbert (2008).
6. Though Brazil is a major biofuel producer (using sugar cane), its production reportedly has not diverted large acreages from grain production.
7. Bloomberg.com, April 17, 2008 (http://www.bloomberg.com/apps/news?pid=20601082&sid=a2QZ._5PDbEs, last accessed July 9, 2009).
8. There have been conflicting reports on the extent to which the announced ban actually reduced the size of Indian exports. But here is no doubt that the ban created great anxiety among importers. As noted, lack of reliable information on quantities is a bane of global grain markets.
9. Vietnam had announced a ban on new sales in July 2007 (Slayton 2009). Thailand and the United States remained in the market as exporters.
10. See Timmer (2008). I have no information that Japan has actually sold these stocks.
11. Near minimum stock levels, small additional fractions of stocks are placed on the market only when the incentive is very high. These stocks may be in relatively inaccessible locations, given current transport costs, or perform valuable roles in keeping the system operating efficiently, such as avoiding the use of half-empty railcars. The small feasible changes in these stocks are ignored here; they have negligible effects on food supply or price volatility. For model of the supply of these stocks, see Bobenrieth, Bobenrieth and Wright (2004).

12. Paul (1970). Deterioration is not important for grains stored in appropriate environments but can be serious in hot and humid environments.

13. In contrast, storage of extra water in a reservoir may incur virtually no extra cost until it reaches full capacity, beyond which extra storage is infeasible in the short run. Above-ground storage of petroleum is similarly limited.

14. Transaction costs associated with adding or removing stocks are assumed to be negligible.

15. Discounting by the cost of capital also makes the timing of benefits and costs to producers, traders and consumers important in determining who gains and who loses from policies affecting storage activity. See Wright and Williams (1984).

16. The first paper to pose the solution to this problem in a modern analytical fashion is Williams (1936). The first satisfactory solution following the approach proposed by Williams appeared more than two decades later in the pioneering dynamic model of Gustafson (1958). A solution method for storage models with responsive supply and rational expectations was first presented in Wright and Williams (1984). See also Williams and Wright (1991, chapter 3).

17. That is, the arbitrage equations for risk–neutral competitive storers who maximize expected profits can be written as

$$P(A_t - x_t) + k = \frac{1}{(1+r)} E_t[P(x_t + \tilde{h}_{t+1} - \tilde{x}_{t+1})], \quad if\ x_t > 0;$$

$$P(A_t - x_t) + k \geq \frac{1}{(1+r)} E_t[P(x_t + \tilde{h}_{t+1} - \tilde{x}_{t+1})], \quad if\ x_t = 0,$$

where k is marginal physical storage cost, E_t denotes the expectation conditional on information available in year t, and $\tilde{h}_{t+1}$ and $\tilde{x}_{t+1}$ are random variables.

18. Deaton and Laroque (1992, 1995, 1996), Chambers and Bailey; Miranda and Rui. The conclusion they draw from their estimation using pseudo maximum likelihood is that the storage model cannot reproduce the serial correlation observed in prices of major commodities.

19. Cafiero and others (2011) show that Deaton and Laroque's negative conclusion regarding the ability of their model to fit the data is due to numerical inaccuracy in the implementation of the estimation model. Bobenrieth, Bobenrieth, and Wright (2010b) present a maximum likelihood estimator for the storage model and apply it to the world sugar market.

20. See Gilbert (1996, 2005) and Gardner (1985) for excellent surveys of international agreements.

21. There are important interactions between band width, private storage within the band, the supply response, the expected rate of accumulation of losses, and the maximum level of stocks. See Williams and Wright (1991, chapter 14).

22. See Williams and Wright 1991, p. 404 for a similar figure. Supply elasticity is one 1.0 with a one-year lag, consumption demand is linear with price elasticity at the mean equal to -0.2, interest rate is 5% and coefficient of variation of harvest is 0.1.

23. To see this, consider the simple case in which demand is linear and planned production is constant so the mean price is exogenous. Assume further that the harvest has a symmetric stationary two-point distribution, that there is no private storage, and that p_p is set at the mean price—the price when consumption equals mean production. Imagine a "buffer fund" scheme whereby the government pays (p_p) for each unit sold at each time t. Negative payments are receipts by the government. The fund's monetary balance, B_t, with initial value B_0, follows a random walk. Given an infinite horizon, the balance passes any finite negative bound in finite time and the probability that it is zero at any future date is the same as the probability that it is never zero before that date and quickly becomes negligible (see Feller [1967, lemma 1, p. 76]). Similarly, a price floor backed by a buffer stock generates a fund balance that hits zero with probability one in finite time (that is, "infinitely often"). If a price ceiling is added, the expected time to a zero balance is shorter.

24. The history of the Australian reserve price scheme for wool (a more complex version of a floor price scheme) is a salutary example where short-run success boosted the confidence of management in its own judgment, leading to decisions that hastened the later catastrophic failure.

25. See Bardsley (1994), Gilbert (1996), and Haszler (1988).

26. See Figure 6.8 in Williams and Wright (1991), p. 171. Consumption demand is linear with price elasticity at the mean -0.2, supply elasticity is zero, coefficient of variation of harvest is 0.1, and interest rate is 5 percent.

27. In the United States, long-run speculators, whose futures positions provide the incentive for storage by short-hedgers, have recently endured a great deal of negative attention, regardless of a lack of evidence of excessive stocks.

28. For more extensive discussions of the rationale for public intervention in storage markets, see Wright and Williams (1982b) and Williams and Wright (1991, chapter 15).

29. von Braun and others (February 2009).

30. von Braun and others (March 2009), p. 3.

31. See for example United States Senate Subcommittee on Investigations, *Excessive Speculation in the Wheat Market*, June 24, 2009, which conflates a real and persistent issue, the failure of convergence of spot and futures prices at delivery, with broader conclusions regarding the role of speculation in market volatility.

32. See Irwin and others (2009).

33. See Verleger (2009) for related findings for the market for crude oil.

34. See the Granger causality tests in Sanders, Irwin, and Merrin (2008).

35. See Caballero and others (2008) for a version of this argument focused principally on the oil market.

36. This figure and the associated argument draw on the work of Dawe (2009).

37. An example of such a reserve forms the first part of a recent three-point proposal by von Braun and others (2009). It sketches an outline of a small "independent emergency reserve" of about 5% of the current annual food aid flow of 6.7 wheat-equivalent metric tons. This would be a decentralized reserve managed by the United Nations World Food Program and held in existing national storage facilities at strategic locations with essentially a call option on the grain deposits at pre-crisis prices.

38. Even if we ignore this difficult issue, optimization of the details of location and operation presents a challenging spatial-temporal problem that merits further attention. See Brennan, Williams, and Wright (1997) for a spatial-temporal model of an exporting region that gives some hint of the issues involved in modeling imports of food aid for a geographically dispersed population.

39. See, for example, Subbarao (2003).

40. The United States Strategic Petroleum Reserve has a similar purpose. Even though its interventions are infrequent, its operation does appear to have reduced private stocks by about half the amount stored in the reserve, consistent with the *ex-ante* analysis of Wright and Williams (1982b).

41. Recent reports indicate that Saudi Arabia, Egypt, Iran, China, Russia, Jordan, Mozambique, Morocco, and Malawi are among the countries placing grain in national reserves (Marc Sadler, personal communication, April 30, 2009).

42. Oxford Analytica, *Global Strategic Analysis*, April 20, 2009.

43. Stocks would be "rolled over" with no net release as frequently as needed to maintain quality.

44. One reason might be that the (generally urban) consumers who most influence the government often are not those most in need.

References

Abbott, P.C., C. Hurt, and W.E. Tyner. 2008. *What's Driving Food Prices?* The Farm Foundation, Oak Brook, IL.

————. 2009. *What's Driving Food Prices?* March 2009 Update. The Farm Foundation, Oak Brook, IL.

Aker, Jenny C. 2008. "Does Digital Divide or Provide? Information Technology, Search Costs and Cereal Market Performance in Niger." Bureau for Research and Economic Analysis of Development (BREAD) Working Paper No. 177.

Bardsley, P. "The Collapse of the Australian Wool Reserve Price Scheme." *The Economic Journal* 104 (September 1994): 1087–1105.

Bobenrieth, E.S.A., J.R.A. Bobenrieth, and B.D. Wright. 2008. "A Foundation for the Solution of Consumption-Saving Behavior with a Borrowing Constraint and Unbounded Marginal Utility." *Journal of Economic Dynamics and Control* 32: 695–708.

————. 2004. "A Model of Supply of Storage." *Economic Development and Cultural Change* 52(3): 605–616.

————. 2002. "A Commodity Price Process with a Unique Continuous Invariant Distribution Having Infinite Mean." *Econometrica* 70(3): 1213–1219.

Brennan, D. "Price Dynamics in the Bangladesh Rice Market: Implications for Public Intervention." *Agricultural Economics* 29(2003): 15–25.

Brennan, D., J.C. Williams, and B.D. Wright. 1997. "Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage under Backwardation." *The Economic Journal* 107(443): 1009–1022.

Caballero, R.J., E. Farhi, and P.O. Gourinchas. 2008. "Financial Crash, Commodity Prices, and Global Imbalances." National Bureau of Economic Research Working Paper 14521, December, Washington, D.C.

Cafiero, Carlo, Eugenio S.A. Bobenrieth, Juan R.A. Bobenrieth, and Brian D. Wright. 2011. "The Empirical Relevance of the Competitive Storage Model." *Journal of Econometrics* 1623: 44–54.

Cafiero, C., E.S.A. Bobenrieth, J.R.A. Bobenrieth, and B.D. Wright. 2010a. "A Maximum Likelihood Estimator of the Commodity Storage Model with an Application to the World Sugar Market." Working paper, University of California, Berkeley.

————. 2010b. "The Empirical Relevance of the Competitive Storage Model." *Journal of Econometrics*, Special Issue.

Dawe, D. 2009. "The Unimportance of "Low" World Grain Stocks for Recent World Price Increases." Agricultural and Development Economics Division (ESA) working paper 09-01. Food and Agricultural Organization, Rome.

Deaton, A., and G. Laroque. 1992. "On the Behaviour of Commodity Prices." *The Review of Economics Studies* 59(1): 1–23.

————. "Estimating a Nonlinear Rational Expectations Commodity Price Model with Unobservable State Variables." 1995. *Journal of Applied Econometrics* 10, Special Issue: The Microeconometrics of Dynamic Decision Making (Dec. 1995): S9–S40.

————. 1996. "Competitive Storage and Commodity Price Dynamics." *The Journal of Political Economy* 104(5): 896–923.

Feller, W. 1968. *Introduction to Probability Theory Third Edition*, Vol. 1. New York: J. Wiley and Sons, Inc.

Gardner, B.L. 1979. *Optimal Stockpiling of Grain*. Lexington, KY: Lexington Books.

———. 1985. "International Commodity Agreements." Processed.

Gerber, N., M.V. Eckert, and T. Breuer. "The Impacts of Biofuel Production on Food Prices: A Review." ZEF−Discussion Paper on Development Policy, December 2008.

Gilbert, C.L. 1996. "International Commodity Agreements: An Obituary Notice." *World Development* 24(1): 1−19.

———. 2005. "International Commodity Agreements." Processed.

———. 2008. "How to Understand High Food Prices." Processed.

Gustafson, R.L. 1958. "Carryover Levels for Grains: A Method for Determining Amounts That are Optimal Under Specified Conditions." USDA Technical Bulletin 1178.

Hansen, K., R. Howitt, and J. Williams. "Valuing Risk: Options in California Water Markets." *Am. J. Agr. Econ.* 90(5): 1336−1342.

Haszler, H.C. "Australia's Wool Policy Debacle: Efficiency, Equity and Government Failure." Ph.D. Dissertation, School of Business, La Trobe University, Bundoora, Victoria, Australia, 1988.

Houthakker, H.S. 1967. *Economic Policy for the Farm Sector*. Washington, D.C.: American Enterprise.

International Energy Agency. "OECD Stocks." *Oil Market Report*, November 2008.

Irwin, S.H., D.R. Sanders, and R.P. Merrin. "Devil or Angel? The Role of Speculation in the Recent Commodity Price Boom (and Bust)." Southern Agricultural Economics Association Meetings, Atlanta, January 31−February 3, 2009.

Keynes, J.M. 1942. "The International Regulation of Primary Products." Reprinted in *J.M. Keynes Collected Works*, Volume 27. London: Macmillan, 1982.

Miranda, M.J., and P.G. Helmberger. "The Effects of Commodity Price Stabilization Programs," *American Economic Review* 78: 46−58.

Mitchell, D.O., and C.L. Gilbert. "Do Hedge Funds and Commodity Funds Affect Commodity Prices?" *DECnote*, International Economics Department of World Bank, February 2007.

Mitchell, D.O. "A Note on Rising World Food Prices." Policy Working Paper 4682, Development Prospects Group, World Bank, July 2008.

Newbery, D.M.G., and J.E. Stiglitz. 1981. *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk.* Oxford: Clarendon.

McNicol, D.L. 1978. *Commodity Agreements and Price Stabilization*. Lexington, MA: Lexington Books.

Mitchell, Donald. 2008. A Note on Rising Food Prices. The World Bank Prospects Group, July. 34.

O'Donnell, M., and B. Colby. "Dry−Year Water Supply Reliability Contract: A Tool for Water Managers." University of Arizona, Department of Agricultural and Resource Economics, October 2009. http://ag.arizona.edu/arec/pubs/facultypubs/ewsr-dyo-Final-5-12-10.pdf (last accessed August 3, 2010).

Paul, A.B. 1970. "The Pricing of Binspace: A Contribution to the Theory of Storage." *American Journal of Agricultural Economics* 52(1): 1−12.

Peck, A.E. "The Futures Trading Experience of the Federal Farm Board." Futures Trading Seminar, Chicago Board of Trade, 1976.

Robles, M., M. Torero, and J.V. Braun. "When Speculation Matters." *IFPRI Issue Brief* 57, February 2009.

Sanders, D.R., S.H. Irwin, and R.P. Merrin. "The Adequacy of Speculation in Agricultural Futures Markets: Too Much of a Good Thing?" Paper presented at the 2008 NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, December 2008.

Schmidhuber, J. "Impact of an Increased Biomass Use on Agricultural Markets, Prices and Food Security: A Long-Term Perspective." *International Symposium of Notre Europe*, November 27–29, 2006.

Slayton, T. "Rice Price Forensics: How Asian Governments Carelessly Set the World Rice Market on Fire." Center for Global Development Working Paper Number 163, March 2009.

Subbarao, K. "Systemic shocks and Social Protection: Role and Effectiveness of Public Works Programs." Washington, D.C.: World Bank, 2003.

Timmer, C.P. 2008. "Causes of High Food Prices," *Asian Development Outlook Update*. Asian Development Bank, Manila, Philippines.

Trostle, R. *Global agricultural supply and demand. Factors contributing to the recent increase in food commodity prices.* WRS-0801, Economic Research Service, Washington D.C., July 2008.

United States Department of Agriculture. 2010. Foreign Agricultural Service-Production Supply Distribution-online commodities data reports. http://www.fas.usda.gov/psdonline/.

United States Senate Subcommittee on Investigations. *Excessive Speculation in the Wheat Market*. June 24, 2009.

United Nations, Food and Agriculture Organization. 2009. "Declaration of the World Food Summit."

Varangis, P., D. Larson, and J.R. Anderson. "Agricultural Markets and Risks: Management of the Latter, Not the Former." *Policy Research Working Paper* 2793, The World Bank, Development Research Group, Rural Development, February 2002.

Verleger, P.K. 2009. "The Great Glut: The Influence of Passive Investors." *The Petroleum Economics Monthly*, XXVI(3).

von Braun, J. "Food and Financial Crisis: Implications for Agriculture and the Poor." *IFPRI Food Policy Report*, December 2008.

———. "Threats to Security Related to Food, Agriculture, and Natural Resources–What to Do?" International Food Policy Research Institute, 2009.

von Braun, J., A. Ahmed, K. Asenso-Okyere, S. Fan, A. Gulati, J. Hoddinott, R. Pandya-Lorch, M.W. Rosegrant, M. Ruel, M. Torero, T.V. Rheenen, and K.V. Grebmer. "High Food Prices: The What, Who, and How of Proposed Policy Actions." *IFPRI Policy Brief*, May 2008.

von Braun, J., and M. Torero. "Implementation of Physical and Virtual International Food Security Reserves to Protect the Poor and Prevent Market Failure." *International Food Policy Research Institute*, October 2008.

———. 2009. "Implementing Physical and Virtual Food Reserves to Protect the Poor and Prevent Market Failure." *IFPRI Policy Brief* 10.

von Braun, J., J. Lin, and M. Torero. *Eliminating Drastic Food Price Spikes–A three pronged approach.* International Food Policy Research Institute, March 2009.

Williams, J.B. 1936. "Speculation and the Carryover." *Quarterly Journal of Economics* 50, 436–455.

Williams, J.C., and B.D. Wright. 1991. *Storage and Commodity Markets.* Cambridge, UK: Cambridge University Press.

Wright, B.D. "The Effects of Ideal Production Stabilization: A Welfare Analysis under Rational Behavior." *Journal of Political Economy* 87(5): 1011–1033.

———, 2008. "Speculators, Storage and the Price of Rice." *ARE Update* 12 no. 2, Giannini Foundation, University of California, Davis.

Wright, B.D., and J.C. Williams. 1982a. "The Economic Role of Commodity Storage." *Economic Journal* 92(59): 6–614.

———. 1982b. "The Roles of Public and Private Storage in Managing Oil Import Distributions," *Bell Journal of Economics* 13:341–53.

———. 1984. "The Welfare Effects of the Introduction of Storage." *Quarterly Journal of Economics* 99(1): 169–182.

Wright, B.D. 2011. "The Economics of Grain Price Volatility." *Applied Economic Perspectives and Policy* 33(1): 32–58.

Wright, B.D., and C. Cafiero. 2011. "Grain Reserves and Food Security in the Middle East and North Africa." *Food Security* (Suppl 1): S61–S76.

Using Contingent Valuation in the Design of Payments for Environmental Services Mechanisms: A Review and Assessment

Dale Whittington and Stefano Pagiola

As the use of payments for environmental services (PES) programs for conservation has grown in developing countries, the use of stated preference methods, particularly contingent valuation (CV) surveys, to estimate the maximum amount that users of environmental services ("buyers") would be willing to pay has also increased. This paper reviews 25 CV studies conducted in the context of PES programs (CV-PES) and assesses their quality and usefulness for designing PES programs. Almost all these studies attempt to estimate the demand of downstream water users for up-stream watershed protection and, more generally, for improved water services. Most studies were methodologically uninspired and generally low-quality applications of stated preference methods, with limited policy relevance. The quality and usefulness of CV-PES studies could be substantially improved at only a modest increase in costs. JEL codes: Q51, Q57

Introduction

Payments for environmental services (PES) programs are an increasingly popular policy instrument in developing countries, especially for watershed protection. Most PES programs involve downstream water users, such as municipal water supply utilities or hydroelectric power (HEP) producers, paying upstream

landholders to undertake activities to protect a watershed. Upstream landholders may be paid to stop deforestation, undertake afforestation, reduce soil erosion on agricultural lands, or cease slash-and-burn agriculture. The potential benefits to downstream water users include improved quality, quantity, and reliability of water supplies; reduced risk of severe floods; and increased preservation of natural areas for future generations.

The price to be paid for environmental services is a critical aspect of any PES program. The viability of any PES program requires that the maximum amount that users of environmental services ("buyers") would be willing to pay for improvements in those services exceeds the minimum amount that providers of those services ("sellers") would be willing to accept. PES program designers have often turned to stated preference methods, particularly contingent valuation (CV) surveys, to estimate either or both of these values.[1] As the use of CV in this context grows, it becomes important to assess how well this method is being applied and how its results can best be used.

In this paper, we review CV studies conducted in the context of PES programs (CV-PES), almost all of which attempt to estimate the demand of downstream water users for upstream watershed protection and, more generally, for improved water services. Our objective is to assess the quality of these CV-PES studies and their usefulness for designing PES programs. We begin by briefly reviewing the use of PES in developing countries (section 2). We then discuss the possible uses of CV studies in PES program design (section 3). Section 4 discusses nine indicators of good practice that we use to assess the quality of CV-PES studies. Although many of the issues that a well-designed CV study must consider are not unique to CV-PES studies, the PES context introduces several special considerations, which we discuss in section 5. In section 6, we review the existing CV-PES studies and assess their overall quality. We then discuss the limitations of the results from this literature (section 7) and conclude by summarizing the implications of our findings (section 8).

Payments for Environmental Services

PES is a market-based approach to conservation financing that is based on two principles: those who benefit from environmental services (such as users of clean water) should pay for such services and those who contribute to generating these services (such as upstream land users) should be compensated for providing these services (Wunder 2005; Pagiola and Platais 2007; Engel et al. 2008). PES programs can thus be conceptualized as an attempt to strike a Coasian bargain between service users and providers, internalizing what would otherwise be an externality. PES programs are attractive because (i) they generate new financing

that would not otherwise be available for conservation, (ii) they are likely to be sustainable because they depend on the mutual self-interest of service users and providers and not on the vagaries of government or donor financing, and (iii) they are likely to be efficient in that they conserve services whose benefits exceed the cost of providing them and do not conserve services when the opposite is true.

There are two basic types of PES programs (Pagiola and Platais 2007; Engel et al. 2008): user-financed PES programs, in which service providers are paid by service users, and government-financed PES programs, in which providers are paid by a third party, typically a government. User-financed programs are preferred in most situations. They are most likely to be efficient because service users provide not only financing but also information regarding the value of services, they can readily observe whether they are receiving the desired services, and they have strong incentives to ensure that payments are used effectively. Conversely, government-financed programs typically cover much larger areas but are less likely to be efficient because governments have no direct information regarding service value or whether services are being provided and they must respond to numerous pressures that are often unrelated to the program's objectives.

In developing countries, user-financed PES programs have been most common for water services, where users are easy to identify and receive well-defined benefits (Pagiola and Platais 2007).[2] The dominance of payments for water services within PES programs is likely to continue. Because of the nature of the services involved, water programs are much easier to implement than, for example, payments for biodiversity services (Pagiola and Platais 2007).[3]

There are now numerous PES programs that involve direct payments by various types of water users at a variety of geographic scales. Municipal water supply systems have been the most frequent participants in PES programs, at a variety of scales ranging from large cities, such Quito, Ecuador (Southgate and Wunder 2009), and medium-size towns, such as Heredia, Costa Rica (Barrantes and Gámez, forthcoming), to small rural towns, such as San Pedro del Norte, Nicaragua (Obando 2007).

HEP producers are also well represented in current PES programs. In Costa Rica, for example, many public-sector and private-sector HEP producers pay to conserve the watersheds from which they obtain water, generating payments of about US$0.5 million and conserving about 18,000 hectares (ha) annually (Pagiola 2008; Blackman and Woodward 2010). In Venezuela, the CVG-Edelca power company pays 0.6 percent of its revenue (about US$2 million annually) to conserve the watershed of the Río Caroní, where 70 percent of the country's HEP is generated (World Bank 2007). Some irrigation systems have also participated in PES programs, for example, in Colombia's Cauca Valley (Echavarría 2002).

Government-financed PES programs can, in principle, target any environmental service deemed to be of social importance. In practice, these programs have focused primarily on water services. The main window of Mexico's Payments for Forest Environmental Services program targets water services (Muñoz et al. 2008). China's Sloping Lands Development Program focuses exclusively on areas at risk of erosion (Bennett 2008). Costa Rica's Program of Payments for Environmental Services currently defines its eligible areas primarily on biodiversity criteria because of early financial support from the Global Environment Facility, but the program is evolving toward a greater focus on water services (Pagiola 2008). Some governments use public resources for PES programs aimed at biodiversity conservation, but such funding is very limited. The area enrolled under the biodiversity window of Mexico's Payments for Forest Environmental Services program is less than one-tenth of that enrolled under the water services window.

Uses of CV surveys in PES Program Design

Payments to service providers in a PES program must be less than the value of the service to users (or it would not make economic sense to provide payments) but more than service providers' cost of supplying the service (or providers would not supply the service). The objective of a CV-PES study could be to determine the maximum amount that a user would be willing to pay suppliers, the minimum compensation that sellers would accept to change their behavior by undertaking different land use activities, or both. To date, the vast majority of CV-PES studies have focused on estimating the buyers' willingness to pay (WTP) for improved environmental services; only a few CV-PES studies have examined service providers' willingness to accept (WTA) payments to modify their behavior.[4] In this paper, we focus on the WTP studies.

One reason that most CV-PES surveys focus on the WTP of service users is that estimates of the cost of service provision by upstream landholders are often relatively easy to obtain by other means. These estimates consist primarily of the opportunity costs of displaced land uses plus any out-of-pocket costs (for example, for planting trees). The rental value of land in an upstream watershed can also serve as a useful proxy for the costs of service provision.[5] The value of improved service provision to users, however, is typically more difficult to observe because prices for such services are administratively determined and often heavily subsidized. Thus, these prices (water tariffs) do not reflect the real value of the services to users.

CV can play several possible roles in PES program design. The most obvious role is to help assess whether PES programs are feasible. By providing estimates of

either WTP for services or WTA to provide them, CV-PES studies can help determine whether there is "room for a deal." A related objective is to assess whether a PES program would improve welfare. In this case, the WTP estimates are used in a cost-benefit analysis to estimate potential welfare increases resulting from improved service provisions. This role is particularly important in the case of government-financed PES programs. CV-PES studies can also provide guidance on the price to be charged to service users. Finally, CV-PES studies can reassure policy makers that implementing a PES program is politically feasible by indicating that users would indeed be willing to pay for the benefits that they would receive.

CV-PES studies can be administered at different stages of PES program design. A survey intended to determine whether a program is feasible would best be administered early in the process whereas a survey aimed at establishing appropriate prices would be most useful late in the process. This decision has implications for the information available for the construction of the stated preference scenario, as discussed below.

Indicators of Good Practice in CV Applications in the PES Field

Conducting CV-PES studies requires adherence to good practices that are necessary in applications of CV studies in all sectors. CV consultants can refer to numerous excellent manuals and books (Mitchell and Carson 1989; Arrow et al. 1993; Louviere et al. 2000; Bateman et al. 2002; Champ et al. 2003; Alberini and Kahn 2006). Best practices in the design and implementation of CV surveys are constantly evolving. What we must do to ensure high-quality results in any particular context is a matter of judgment and is subject to budgetary constraints. In this section, we briefly describe the nine indicators of good practice that we subsequently use to assess our CV-PES study sample.

The nine indicators that we use here are not meant to be comprehensive. Moreover, we recognize that CV-PES study researchers may not always have the time or budget to implement all of these best practices. The National Atmospheric and Oceanic Administration (NOAA) Panel Guidelines (Arrow et al. 1993) form the basis for some, but not all, of our indicators. We have selected these nine indicators because they are relatively easy and straightforward to assess by reading the CV-PES studies and because they cover a range of design and implementation issues.

Using Methods to Reduce Hypothetical Bias

The main criticism that economists level at CV studies is that WTP estimates are inflated because respondents do not face an actual budget constraint

(hypothetical bias) and because they are prone to say "yes" too easily, perhaps just to please the interviewer (enumerator bias). These sources of bias are serious threats to CV-PES study results (Whittington 2010). However, CV researchers have developed several ways to reduce this yea-saying tendency, including (i) cheap-talk scripts (Cummings and Taylor 1999; List 2001; Carlsson et al. 2005), (ii) ballot boxes to simulate voting behavior (Carson et al. 1994; Krosnick et al. 2002; Harrison 2006), (iii) recalibration of results using data from real experiments (Blackburn et al. 1994), (iv) time-to-think experiments (Whittington 2002), and (v) drop-off protocols (Subade 2007). Using any of these methods to reduce the risk of hypothetical bias is an important indicator of the quality of a CV-PES study.

Asking Debriefing Questions

CV researchers typically follow up a respondent's answer to the valuation question with a series of "debriefing questions." The NOAA Panel Guidelines (Arrow et al. 1993) called for debriefing questions, referring to them as "Yes/No Follow-ups." If respondents say "yes" and agree to pay the offered amount (bid) in the CV scenario, the interviewer follows up with questions about why the respondents agreed to pay. If the respondents say "no," that they will not pay, then the interviewer follows up with questions about why they are not willing to pay. The purpose of debriefing questions is to attempt to determine whether respondents have interpreted and answered the valuation question in the way that the researcher intended. Respondents can offer legitimate and illegitimate reasons for both "yes" and "no" answers to the valuation question(s). A well-designed CV-PES study will include debriefing questions to separate legitimate from illegitimate answers to the valuation question(s).

Asking Uncertainty Questions

CV researchers routinely attempt to gauge the level of confidence—or certainty—that respondents have in their answers to the valuation question (Alberini et al. 2003; Li and Mattson 1995; Loomis and Ekstrand 1998; Whitehead et al. 1998; Samnaliev et al. 2006; Akter et al. 2008). A high level of certainty in respondents' answers may be an indicator that, in fact, they will pay the offered bid amount. Answers to uncertainty questions can be used during the analysis of the survey data to decide how many of the respondents who said "yes" to the valuation questions should actually be treated as definite "yes" votes. The NOAA Panel Guidelines called for including a simple "don't know" or "not sure" response. Other approaches have been used to assess respondents' uncertainty (for example, Wang 1997). Some CV researchers prefer to embed the uncertainty questions

directly into the available responses to the valuation questions (Ready et al. 1995). We consider any approach to obtaining information about respondents' uncertainty toward their answer to the valuation question to be an indicator of a high-quality CV-PES study.

Determining Whether Respondents Are "in the Market"

When a dichotomous choice, referendum question is used to elicit respondents' WTP, the researcher will typically want to carefully distinguish respondents who do not value the service at all from those who will not pay the offered price but may be willing to pay something. Policy makers are often interested in the raw data on the number of respondents who are not willing to pay anything. If there are many "zero WTP" respondents, spike models may be the most appropriate econometric framework for analyzing the covariates of respondents' answers to the valuation questions (Hanemann and Kristrom 1995; Kristrom 1997). Several approaches are used in the literature to identify "zero WTP" respondents. The approach that we prefer is to begin the valuation questions with the discrete price offer. Respondents who say "yes" are clearly willing to pay something and are "in the market." If respondents say "no," then it is natural to follow up by asking, "Would you pay anything?" If respondents again say "no," sometimes a second follow-up question is posed: "Would you take the service for free?" However, in our assessment, we consider the inclusion of any sequence of questions to determine whether respondents are in the market to be an indicator of a high-quality CV-PES study.

Using Visual Aids to Explain the CV Scenario

In well-crafted CV survey instruments, respondents are presented with a hypothetical management plan (policy intervention) and a choice as to whether they would be willing to pay a specified amount of money for the plan to be implemented. The NOAA Panel Guidelines called for an "Accurate Description of the Program or Policy" and for "adequate information" to be provided to respondents about the program being offered (Arrow et al. 1993, 10). One way to accurately convey the details of the hypothetical management plan and the results of its implementation is to use pictures, maps, diagrams, figures, and tables (Labao et al. 2008). Visual aids are not always required, but their use in a survey protocol suggests that the researcher is seriously concerned that respondents understand the CV scenario. In CV-PES studies, there are many possible uses of visual aids to convey relevant information. For example, if the management plan requires upstream landowners to change their land use practices, photographs could be used to show the current state of erosion in the upstream watershed and what the

land would look like after afforestation. Diagrams could be used to show how downstream water quality would improve. Conveying such information to urban residents without visual aids could be very difficult. We consider the use of visual aids during the presentation of the CV scenario to be an indicator of a high-quality CV-PES study.

Using Split-Samples to Test for the Robustness of Results

The NOAA Panel noted that "common notions of rationality" impose requirements on CV survey results (Arrow et al. 1993, 11). For example, respondents are usually assumed to be willing to pay more for more of a service than for less of it. CV researchers may ask different split-samples of respondents their WTP for different levels or "scope" of the service to be provided to demonstrate that respondents' answers to the valuation questions are consistent with common notions of rationality. Such scope tests are not always straightforward because there is often little a priori guidance on how much such estimates should differ. We consider the use of scope tests and other split-sample experiments to test for the reliability and accuracy of the WTP results to be an indicator of a high quality CV-PES study.

Testing Whether Income is Positively Correlated with WTP

Demand theory suggests that WTP for normal goods increases as income increases. Other things equal, we expect high-income respondents to have higher WTP than low-income respondents. If this is not true, respondents may not be answering the valuation questions as the CV researcher intended. We thus expect a high-quality CV-PES study to report whether income is positively correlated with respondents' WTP.

Addressing Intrahousehold Allocation

Intrahousehold allocation issues pose complex research design decisions for CV researchers (Adamowicz et al. 2005; Whittington et al. 2008; Prabhu 2010), including whether respondents are supposed to answer the valuation questions for themselves or for the entire household and whether to interview the husband, the wife, or both. The simplest approach is to use the household as the sampling unit and to interview whoever is identified as a household decision maker, usually either the husband or the wife. However, when a household's decision making is best characterized as cooperative bargaining, this simple approach is likely to be inadequate. We consider an explicit effort to address such

intrahousehold allocation issues in the determination of who to interview and in the construction of the CV scenario to be an indicator of a high-quality CV-PES study.

Obtaining Informed Consent

Obtaining informed consent from respondents is necessary to ensure that they can choose whether to participate in the survey (Whittington 2004). An informed consent form is presented to potential respondents before an interview. This form informs the respondents about the research objectives, the sponsoring agency, and any potential risks to their household or others. The form assures the anonymity of the respondents and provides them with someone to contact if problems occur (this person cannot be directly affiliated with the research project). If compensation is offered to respondents, it should be clear on the form that this compensation will be paid even if they decline to participate. Offering respondents an informed consent form certainly does not solve all of the potentially problematic ethical issues involved in conducting CV studies, but it is a step in the right direction. We consider an effort to obtain respondents' informed consent to be an indicator of a high-quality CV-PES study.

Special Challenges in the Design of CV-PES studies

The nine indicators of good practice described in the previous section are broadly applicable to CV studies in all sectors. In addition, there are specific challenges in the design of CV-PES studies. In PES programs, payments from downstream water users are collected and used to pay upstream landholders to undertake land uses that are expected to improve water services. There are several sources of uncertainty in this context.

The first challenge arises from the difficulty of predicting how specific upstream land uses will affect downstream water quality and quantity in a particular watershed. The scientific evidence to establish this relationship is often weak. Downstream users thus bear a risk that benefits will be lower than anticipated (Pagiola and Platais 2007). Undertaking detailed ex ante hydrological studies reduces this risk but cannot completely eliminate it.[6] The impact of this risk can also be mitigated in well-designed PES programs by including monitoring and evaluation systems that enable adjustments to be made to landholder contracts to ensure that downstream users receive the benefits for which they are paying. In user-financed PES programs, users also have the option of ending payments if they are unsatisfied with the services that they receive, thus limiting possible losses.

CV-PES study designers need to decide how much of this scientific uncertainty to explain to respondents during the interview. Broadly speaking, there are two ways to proceed. One approach is to attempt to convey to respondents the true degree of scientific uncertainty about the consequences of upstream actions and to try to ensure that respondents incorporate an understanding of these risks in their responses to the valuation questions. In this case, the WTP estimates will incorporate the information that the policy outcomes are uncertain. Survey designers could also describe the features of a PES program that can help mitigate the risk. The other approach is for survey designers to try to estimate WTP for specific policy outcomes contingent on the success of the watershed protection activities. In this case, the downstream users' WTP estimates are policy relevant only if planners are confident that the upstream watershed protection activities being considered will result in outcomes at least as good as the respondents were told to assume in the CV-PES study.

A second challenge concerns the description of institutional uncertainty. PES programs require money to be collected from service users, administered by an institution, and then used for the intended purposes. In many developing country situations, respondents may be skeptical that any monies that they provide will actually be paid to upstream landholders or that landholders will respond as expected. Respondents may refuse to participate in a PES program not because of scientific uncertainty or because they place low values on service improvements but because they lack confidence in the institutions. Researchers could address this institutional uncertainty by acknowledging in the questionnaire that many people feel this way and specifically instructing respondents to suspend their lack of trust in institutions and assume that the money would be handled honestly and provided to the upstream landowners as promised. This challenge is especially serious in the case of CV-PES studies that are administered early in the program design process. Here, too, the value of the service could be estimated contingent on its successful delivery.

A third challenge faced by survey designers is that respondents may have preferences for more than just improved service delivery. Downstream users may also care about protecting upstream watersheds because they provide wildlife refuges, forests for recreation, and nonuse environmental benefits. Upstream landholders may be poor, and downstream respondents may place a premium on helping them. If people care about upstream land uses for reasons other than downstream service improvements, omitting these other reasons from the information set provided to respondents may result in underestimations of WTP. An important question concerns the amount of detail that should be provided to respondents about how the PES program would work in upstream areas—who would be paid and to do what. One extreme is not to tell respondents anything about the management plan (or even the PES program itself) and to simply measure their demand for

specific improvements in the downstream water services without telling them *how* these improvements will come about. The other extreme is to tell respondents a good deal about the management plan and what landholders would have to do to receive the proposed payments.

The different CV-PES studies can thus be classified according to the degree to which they recognize and address the challenges of explaining scientific and institutional uncertainty and describe the elements of PES program management plans that can affect respondents' preferences. At one extreme, studies could provide information about all three of these aspects. In a simpler approach, the survey could provide information about some, but not all, of these aspects. Finally, respondents could simply be presented with a scenario that asks them to value specific improvements in downstream water services without being told about either the management plan or the uncertainty in the outcome.[7]

An Assessment of the Quality of Existing CV-PES Studies

To assess the quality of existing CV-PES studies, we sought studies that were conducted specifically in the context of actual or hypothetical PES programs for watershed protection. We collected 25 such studies, listed in Table 1.[8] Many of the applications that we review are in the gray literature; only a few have been published in refereed journals.[9] Two of the studies are master's theses. Almost half of the papers are only available in Spanish. Several researchers recur frequently among the contributors.

All of the studies in our review were conducted in the past decade—not surprisingly, because PES program have only been in use since the late 1990s. Almost all of the studies are from Latin America—again, not surprisingly, because most existing PES programs have been implemented there (Southgate and Wunder 2009; Camhi and Pagiola 2009). In fact, ten studies are from Mexico[10], and five are from Costa Rica. Only one is from Africa, and two are from Southeast Asia (both from the Philippines).[11] None of the studies is from the Middle East or other parts of Asia. In almost every study, the downstream parties were urban water users. There are only a few studies on irrigators' WTP to preserve their water supplies (Lopez et al. 2007; Shultz and Solis 2007), and only one study of electricity users' WTP to protect watersheds where HEP is generated (Alpízar and Otarola 2007).

In assessing the studies, we were limited by the information provided by the papers and reports. Many CV-PES studies did not include the survey instrument or report the CV scenario, nor was the approach described in sufficient detail for us to fully assess the quality of the fieldwork and the results. Even the gray

Table 1. Characteristics of CV Surveys Used for Analyzing Payments for Environmental Services Programs

Location (country, site)	Policy or hypothetical[a]	Date of study[b]	Size of sample[c]	Sources[d]
Bolivia				
Comarapa (town)	P	nd	221	Shultz and Soliz 2007 [PR]
Comarapa (lower watershed)	P	nd	188	Shultz and Soliz 2007 [PR]
Colombia				
Chaina	P	2006	300	Moreno-Sánchez et al. 2012 [PR]
Costa Rica				
Cartago	?	2003	413	Alpízar and Otarola 2007 [BC]
Dos Novillos watershed	H	2005	398	Kaplowitz and Lupi 2008 [UN]
Esparza	?	2005	365	Alpízar and Madrigal 2007 [UN]
Reventazón watershed	H	2006	300	Ortega-Pacheco et al. 2009 [PR]
Turrialba		2002	200	Berggren and Stahl 2003 [ST]
Ecuador				
Cotacachi	H	2002; 2004	274	Rodriguez et al. 2009 [PR]
Ghana				
Weija	H	2008	89	Peprah 2009 [ST]
Honduras				
Copan Ruinas	P	nd	285	Cisneros et al. 2007 [UN]; Madrigal and Alpízar 2007 [UN]
Siguatepeque	?	2002 (?)	337	Cruz and Rivera 2002 [UN]
Mexico				
Bahías de Huatulco, Oaxaca	P	2007	376	González-Ortiz 2007 [CR]
Coatepec and nearby towns, Vera Cruz	P	2007	197	Puente-González 2007 [CR]
Colima-Villa de Alvarez, Colima	P	2007	422[e]	Pizano-Portillo 2007 [CR]
El Cielo-Ciudad Victoria area, Tamaulipas	P	2007	432	Campos-Benhumea 2007 [CR]
Monterrey, Monterrey	P	2007	384	Saldivar-Valdés 2007 [CR]
Saltillo, Coahuila	P	2007	180	Arias-Rojo 2007 [CR]
Santa María de Huatulco, Oaxaca	P	2007	381	González-Ortiz 2007 [CR]
Six small towns, Quintana Roo	P	2007	377	Contreras-Benítez 2007 [CR]
Tapalpa watershed, Jalisco	?	2005 (?)	243	López et al. 2007 [PR]
Upper watershed of Rio Balsa, Mexico	P	2007	837[f]	Vargas-Pérez 2007 [CR]
Nicaragua				
San Dionisio	H	1998	153	Johnson and Baltodano 2004 [PR]
Philippines				
Metro Manila	H	nd	2232	Calderon et al. 2006 [PR]
Tuguegarao City	?	2006	401	Amponin et al. 2007 [WP]

a. P = Policy; H = Hypothetical.

b. Dates marked '(?)' are not stated in the report but are inferred from the context; 'nd' indicates that no date was provided.

c. Values refer to completed interviews. Response rates are rarely reported, so it is generally not possible to determine the original sample size.

d. BC = Book chapter; CP = Conference paper; CR = Consultant report; PR = Published in a peer-reviewed journal; ST = Student thesis; WP = Formal working paper series; UN = Unknown.

e. In addition, the researchers surveyed 356 commercial water users.

f. The sample included 168 households in the watershed, 353 households in the city, and 316 households in the suburbs.

literature reports, which do not face the length restrictions of journal articles, often failed to provide sufficient detail regarding their methodology.[12]

Many CV-PES studies were undertaken as part of the design of proposed or actual PES programs or to examine working PES programs. We describe these as "policy" studies. Other applications are for purely hypothetical PES programs. Among the policy studies, almost all were undertaken during the design phase of PES programs, but one (Moreno-Sánchez et al. 2012) examined the possible expansion of a working program.

Most of the studies appear to have used in-person interviews in respondents' homes. All studies used a monetary numeraire to measure WTP.[13] With the notable exception of the Calderon et al.'s (2006) study from Manila, Philippines (n = 2232), the sample sizes of the CV-PES studies were relatively small.[14] Almost all of the studies (18) used dichotomous choice questions, mostly single-bounded, whereas four used a payment card and one asked an open-ended valuation question.

The most common payment vehicle was the household water bill, but a surprisingly large number of studies did not specify a payment vehicle. Respondents were simply asked whether they would pay a given amount without being told how this amount would be collected. In some cases, neither the elicitation procedure nor the payment vehicle was reported in the study.

The mean or median household WTP of water users for improved services is not reported in all studies. Figure 1 summarizes the available results from studies of WTP for improved domestic water supplies. Estimates range from US$0.42 per month for households living in five small communities in Nicaragua (Johnson

Figure 1. Estimated Household Willingness to Pay for Improved Water Supplies in CV-PES Studies
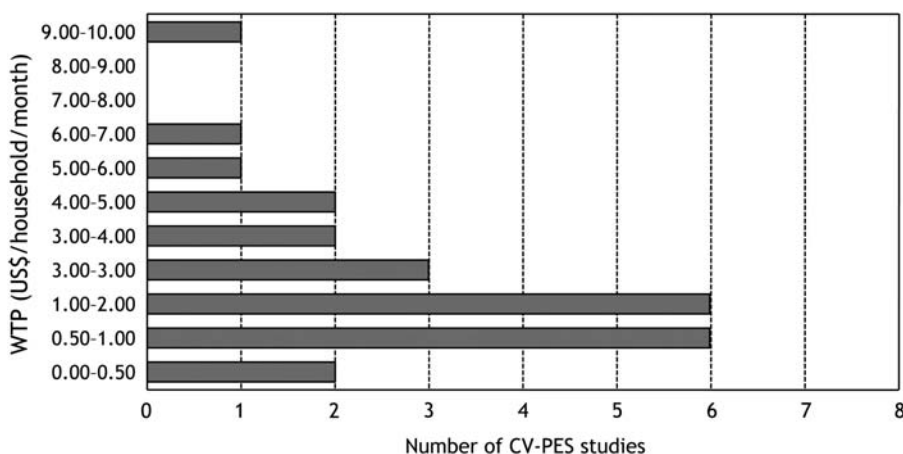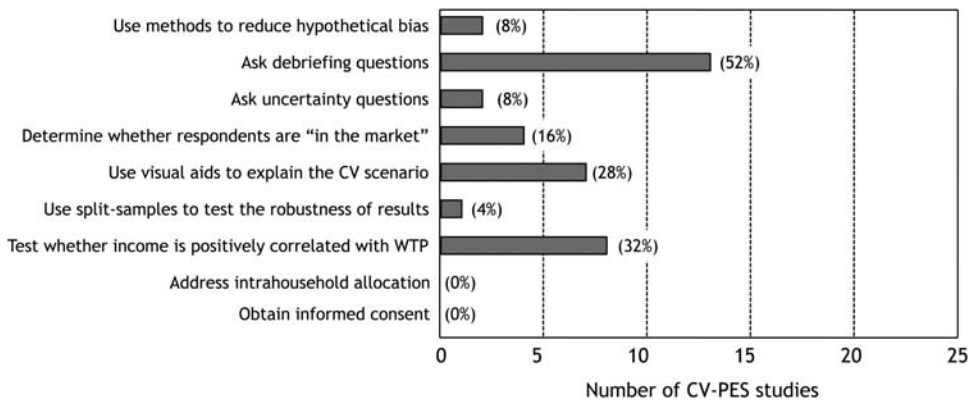
Figure 2. Number of CV-PES Studies That Included Each Indicator of Study Quality

and Baltodano 2004) to US$6.90 in Turrialba, Costa Rica (Berggren and Stahl 2003) and about US$10 in Jalisco, Mexico (Lopez et al. 2007).[15] Two-thirds of the WTP estimates are less than US$3 per household per month. In one of the most carefully executed studies, Calderon et al. (2006) reported a mean WTP for households in Manila of US$0.50 per month. However, these estimates are not strictly comparable because they refer to different degrees of improvements.

Figures 2 and 3 indicate how the CV-PES studies that we reviewed fared in terms of our nine indicators of good practice. Figure 2 shows the number (and percentage) of studies in the sample that used each of the nine indicators of good quality. The three indicators found most often in the CV-PES studies were debriefing questions (52 percent),[16] tests of whether income was positively correlated with income (32 percent), and the use of visual aids in the presentation of the CV scenario (28 percent). Only two studies (8 percent) used any of the currently available techniques to minimize hypothetical bias: Calderon et al. (2006) and Amponin et al. (2007) used 'cheap talk' scripts. Very few studies asked questions to assess respondents' uncertainty (8 percent) or used split-sample experiments to test for the robustness of respondents' answers to the valuation question (4 percent). None of the studies explored intrahousehold allocation issues or discussed obtaining informed consent. Figure 3 presents a simple count of the number of studies in our sample that used different numbers of quality indicators (from zero to nine). For example, seven of the twenty-five CV-PES studies did not have (or did not report) using any of the nine quality indicators, eight studies had only one of the nine attributes, and only two of the twenty-five studies had six or more indicators of quality (neither was from Latin America). The mode was one indicator; the mean was 1.6 indicators.

Figure 3. Number of Quality Indicators Included in Reviewed CV-PES Studies
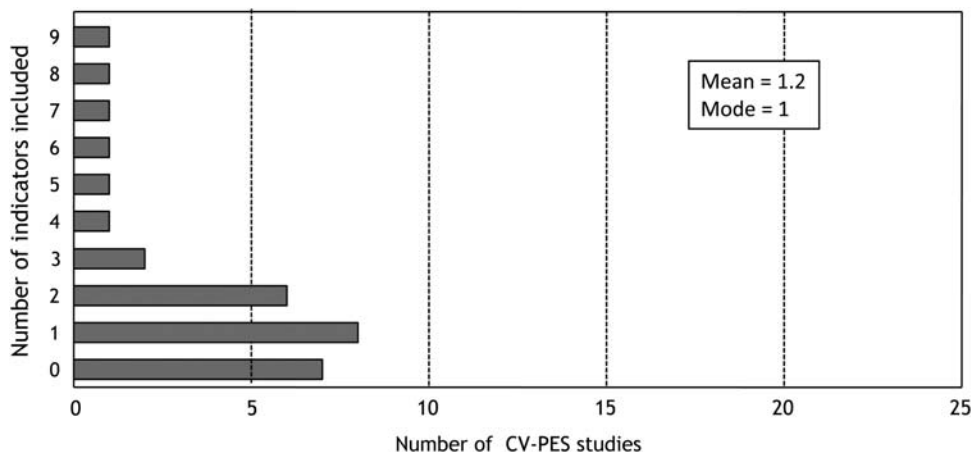
The best studies provide figures, tables, and photographs to respondents to help them understand the choice task, but only seven studies in our sample reported doing so. Many studies provided respondents with little or no information about how service improvements would be achieved; indeed, many studies provided almost no information about what type of service improvement respondents would receive and simply asked respondents for their maximum WTP for "water service improvements." None of the studies attempted to convey to respondents the uncertain outcomes associated with upstream watershed protection activities, nor did any of the studies ask respondents to suspend their possible skepticism about institutional uncertainty. In most studies, respondents were not told about either the management plan or the scientific and institutional uncertainty associated with the management plan and downstream outcomes. Calderon et al.'s 2006 study provides a good example of an information set in which respondents were told about watershed protection activities upstream and the downstream consequences but not about the risk that some of these outcomes might not materialize.

There are two especially revealing indications of the wide variation in the quality of the CV-PES studies in our sample. First, many studies failed to identify a statistically significant relationship between respondents' answers to the valuation questions and household income (or wealth). This result is quite unusual in well-executed stated preference studies. Second, the choice tasks presented to respondents varied widely in their clarity and policy relevance. Some of the valuation questions were not incentive compatible, meaning that respondents had an incentive to misrepresent their preferences, and were inappropriate in the PES context

of collective action. For example, an open-ended valuation question that asked respondents their maximum WTP for upstream watershed protection would not be incentive compatible (Carson and Groves 2007; Whittington, 2002).

The state of the art in conducting stated preference studies is constantly evolving, and some of these CV-PES studies are now a decade old. While it would, of course, be unfair to impose current standards on the older studies, the NOAA Panel's recommendations for CV studies (Arrow et al. 1993) are now almost two decades old, and most CV-PES studies in our review do not meet these standards. Thus, we believe that it is accurate to characterize most CV-PES studies as methodologically uninspired and generally low-quality applications of stated preference methods.[17]

How Useful Are the Results of the Stated Preference Studies for Policy Purposes?

PES programs are not always financed by levying additional charges on water users. In many cases, payments for conservation activities are financed from the savings resulting from lower treatment costs or the avoided costs of building new infrastructure (Pagiola and Platais 2007). In such cases, WTP surveys would not be necessary; if utility payments to upstream service providers were lower than the cost savings, there would be no need for water users to pay more. Because estimates of cost savings would be based on assessments of existing conditions, they would generally be preferred to estimates based on stated preferences. Understanding WTP could be useful when current spending is required to avoid future costs, or when substantial investments are needed to improve water services (or avoid their degradation). None of the studies in our sample, however, appears to have considered alternative approaches to estimating the benefits of watershed conservation before undertaking CV-PES studies.

In cases where CV-PES studies are called for, their potential usefulness depends in part on their accuracy and reliability. As discussed in the previous section, the quality of many studies raises questions in this regard. However, even if the WTP estimates from CV-PES studies are accurate and reliable, they are only one input into a negotiation process between upstream landholders and downstream water users. Almost all of the papers are silent on how their results can be used in PES program design.

The authors of some studies seem to argue that a PES deal is feasible if the summation of downstream users' WTP is greater than the upstream landholders' minimum WTA to implement the watershed protection plan. In fact, for a PES program to be feasible, three conditions must hold. First, the potential revenue

collected from downstream users for the PES program must exceed the minimum payments required by upstream landholders to participate. Water service providers are not perfectly discriminating monopolists, so it is not possible to collect revenues equal to the summation of the maximum WTP of all downstream users. Only one CV-PES study attempted to use the WTP estimates to calculate the revenue that might be collected (Calderon et al. 2006). Although CV-PES studies could provide some of the raw data needed to support PES program design, there is little evidence that these data are being used correctly to estimate potential revenues. Second, the payments from downstream water users must be less than the costs of alternative means that achieve the same service improvements. In the language of negotiations, the PES deal for the downstream users must be better than their "Best Alternative to a Negotiated Agreement." Third, the transaction costs of collecting payments from service users and making payments to service providers must be less than the difference between the WTP and the WTA. These three conditions together imply that there is potential for a PES program if the potential revenues from downstream users are greater than the sum of the payments necessary to compensate landholders and the program's transaction costs *and* if they are less than the costs of alternative means of delivering service improvements.

The results from CV studies of downstream users alone are not sufficient to demonstrate this condition. None of the authors of the papers included in our sample supplemented their CV-PES study results with additional information regarding the costs of alternative means of achieving equivalent service improvements or the compensation needed by participating upstream landholders, to examine the feasibility of a potential PES negotiation, and none provides information about transaction costs (indeed, most do not even mention them).

There is a strong inclination for authors to simply claim that their results are policy relevant without demonstrating how these estimates of demand for improved services or upstream watershed protection, or both, can be used to make better decisions. In some cases, authors may make such claims because the CV-PES study was undertaken primarily for academic purposes, with the authors' search for policy relevance occurring after the research was finished, when they sought to market their findings to policy makers. However, some of these CV-PES studies were, in fact, undertaken for clients. For example, the national forest commission (CONAFOR), which administers Mexico's Payments for Forest Environmental Services program, commissioned most of the studies from Mexico, and several of the studies from Honduras were undertaken under the FOCUENCAS project that implemented a pilot PES study.

The CV-PES studies are largely silent on how the estimated WTP amounts can be used to revise water tariffs and to collect the revenues needed to make payments to upstream landholders. In some CV-PES studies, respondents were asked

an open-ended maximum WTP question; in others, respondents were presented with a fixed increase in their monthly water bill. None of the CV-PES studies offered respondents a higher volumetric charge for their water or asked them how much water that they would want to purchase at this higher price.[18] Alpízar and Madrigal (2007) simply divided the estimated WTP by the average water use to estimate WTP in volumetric terms, but this method ignores decrease in use that would result if the unit price increased. In rural communities in Latin America, volumetric tariffs are relatively rare, but in medium-sized and large municipalities, volumetric charges (often in the form of increasing block tariffs) are often used. If volumetric charges are used, the only reasonable way that CV-PES study results can be used for tariff design is to estimate the amount that can be added to a fixed-charge component in the tariff structure. This addition to the fixed charge is how the extra fee for a PES program should be described to respondents in the CV-PES studies, but it may not be the most appropriate way to modify the tariff (Boland and Whittington 2000).

Some authors use the estimated WTP to calculate the consumer surplus and simply offer the estimated WTP as a maximum total payment that should be collected from water users (Cisneros et al. 2007; Alpízar and Madrigal 2007). While using the estimated WTP for this calculation is technically correct, it does not provide program developers with much concrete guidance. Alpízar and Madrigal suggest charging 50 percent of the estimated WTP to "divide the consumer surplus equally between service users and service providers" (2007, 17). However, this approach does not incorporate information about the service provider's minimum WTA and does not necessarily result in a feasible or fair deal. From the perspective of a two-party PES negotiation, both parties might perceive a negotiated settlement to be fair if it approximately splits the difference between a provider's minimum WTA and a buyer's maximum WTP.[19]

How can the WTP estimates be used to estimate the amount of this increase in the fixed charge? One approach would be to use the CV-PES study results to estimate the monthly charge that would pass a public referendum (for example, 50 percent approval), perhaps with a supermajority (for example, 66 or 75 percent). From both an economic and a political perspective, utilities may not want to implement tariff reforms that would result in dramatically reduced household water use—or in substantial numbers of households disconnecting from their network. None of the CV-PES studies asked respondents what their household would do if the proposed monthly fee were implemented even if they personally said they would not pay. A household that voted "no" to a proposed increase in its monthly water bill might disconnect from the water system. Alternatively, the household might pay the proposed increase in the fixed charge and suffer a welfare loss (Whittington 2002). This uncertainty about how households would behave in response to a tariff increase may be one explanation for what occurred in Heredia,

Costa Rica, where CV was used to estimate households' WTP but actual fees were set far below the estimated WTP (Barrantes and Gámez, forthcoming).

The policy relevance of WTP estimates from these CV-PES studies for the redesign of water tariffs is limited by another factor. In many instances, the existing water tariffs generate revenues below the costs of system operation and maintenance and far below the cost of capital replacement. In such a situation, households' total water bills may still be quite modest, even with an added fee for upstream watershed protection, and they may be more likely to approve the fee. Their WTP for a PES program might have been quite different if the water utility had already implemented a water tariff structure that recovered a higher percentage of the total costs of service. The estimates of incremental WTP for watershed protection may be highly contingent on the low initial water tariff. The utility may have some room to increase water tariffs and still maintain public support, but this slack could be quickly used up by any increase in the monthly water bills, for whatever reason. In other words, utilities could not increase the water tariffs in an attempt to recover more capital costs and then rely on the CV-PES study results to justify raising the tariff again to pay for upstream watershed protection.

In principle, the fact that several CV-PES studies have been conducted in a policy context could facilitate a comparison of ex-ante WTP estimates with ex-post payments. Such a test of the accuracy and reliability of WTP results is, however, not always possible.[20] In Copán Ruinas, Honduras, for example, initial payments under the local PES programs were made using funds provided by a donor rather than from charges to water users (Madrigal and Alpízar 2007). The putative WTP was thus left untested. Water use charges for PES programs have actually been implemented in several cases where prior CV-PES studies were undertaken, including Heredia, Costa Rica (Barrantes and Gámez, forthcoming) and Saltillo, Mexico (Pagiola 2010). Unfortunately, we have been unable to obtain copies of these studies.[21] In both of these cases, the introduction of water charges has been unproblematic. The test of predicted as compared with actual behavior is not very stringent, however, because the payments assessed to water users have usually been far below the estimated WTP. In the case of Heredia, for example, the actual charge was only about 10 percent of the estimated WTP (Barrantes and Gámez, forthcoming). Perhaps the most interesting case is that of Saltillo, where a study showed positive WTP. A purely voluntary payment mechanism was created, in which water users could, at their discretion, add an amount to their water bill. In 2009, 31,000 households (about 10 percent of water users) made voluntary contributions to the program, totaling M$1.2 million (Pagiola 2010). In a more recent case, the PES program in Chaina, Colombia, increased its charges from US$0.50 per household per month to US$1.00 per household per month on the basis of the results of a CV-PES study (Moreno-Sánchez et al. 2012).

Discussion and Recommendations

Our objective in this review of CV-PES studies is quite modest. We have read many of the existing CV-PES studies and have reported on their quality. We do not know the actual impact of the CV-PES studies on the design of PES programs, nor do we know how useful decision makers have found CV-PES study results, except in a few cases. It is possible that they were satisfied with the work despite our assessment that the quality of most studies was quite low judged against the state of the art.

However, there is little reason that the quality of CV-PES studies cannot be substantially improved at only a modest increase in costs. The primary impetus will probably come from the purchasers of CV-PES studies—the clients of the CV-PES study consultants—who should demand higher quality products for their money.[22] A necessary first step will be improved terms of reference (TORs) for CV-PES studies.

We do not recommend that TORs require CV-PES study consultants to rigidly adhere to the NOAA guidelines (Arrow et al. 1993) or other such protocols. Agencies commissioning CV-PES studies have various information needs and different budget constraints. However, we believe that it is reasonable for TORs to include at least the following four elements.

First, clients should be involved in the selection of the information set(s) to be presented to respondents in CV scenarios. CV-PES study consultants should provide clients with alternative information sets and discuss the pros and cons of each before the CV-PES study is launched. Clients should expect that CV-PES study consultants will use photographs, figures, tables, and perhaps video clips to communicate information to respondents and should ask to review such information before the survey is launched.

Second, TORs for CV-PES studies should require researchers to demonstrate that they have considered (i) alternative means to reduce hypothetical bias, (ii) alternative payment vehicles for collecting monies from service users, and (iii) the choice of respondent within a household (that is, who to interview).

The TORs should request that CV-PES study consultants discuss the pros and cons of different options for these three design issues and justify their recommendations. Although it may be difficult for many clients with little experience with stated preference techniques to effectively review such decisions, advisory panels or outside consultants may be engaged to provide suggestions for improvements or alternative perspectives.

Third, the TORs should request that CV-PES study consultants provide estimates of the potential revenues that could be obtained from downstream water service users if different prices or charges were implemented and that they specify the options that they propose for adjusting tariff structures. Decision makers

typically want to understand their options, and CV consultants should be asked to link their studies and recommendations more closely to the actual decisions that need to be made in the design of pricing and tariff structures.

In addition to improved TORs, it would also be helpful for agencies involved in PES programs and commissioning CV-PES studies to have access to information about what others are doing through a Web-based clearinghouse. It would be relatively simple and inexpensive for an international organization (for example, a nongovernmental organization) or one of the regional environmental economics networks (for example, the Latin American and Caribbean Environmental Economics Program, Economy and Environment Program for Southeast Asia, South Asian Network for Development and Environmental Economics) to post studies and survey instruments so that clients and researchers could easily the research of others and how they have tackled some of the challenges discussed in this paper.

An old joke in economics concerns a drunk who searches for his lost house keys under a streetlight, not because that is where he lost them, but because that is where the light is. The use of stated preference techniques in the design of PES programs often has a strong hint of this. Our impression is that the use of the CVS is often driven by the perceived ease with which it can be applied in the PES context rather than because it is the best tool for the job. Properly designed, carefully conducted CV-PES studies can, in many cases, provide useful insight for the design of PES programs, but they are certainly not required in all instances.

Notes

Dale Whittington is Professor of Environmental Sciences & Engineering and City & Regional Planning, University of North Carolina at Chapel Hill, and Professor, Manchester Business School. Dale_Whittington@unc.edu. Stefano Pagiola is a Senior Environmental Economist, Latin America and Caribbean Sustainable Development Department, World Bank, spagiola@worldbank.org

1. Contingent valuation (CV) is one type of stated preference method to estimate the willingness to pay of downstream users and the willingness to accept payment of upstream landowners. In this paper, we focus on CV rather than stated preference methods more generally because the vast majority of stated preference applications in the PES field use CV. Many of our observations and conclusions are equally applicable to other stated preference methods.

2. Our discussion in this paper focuses on the use of PES in developing countries. See Salzman (2005) for a discussion of some applications in industrialized countries.

3. Programs aimed at sequestering carbon are a distant second, in terms of the number of mechanisms and area covered, after water services (Camhi and Pagiola 2009). This position may change in the future, however, if markets develop for Reduced Emissions from Deforestation and forest Degradation (REDD).

4. We found only two CV studies that examined upstream landholders' WTA payments to participate in a PES mechanism (Southgate et al. 2009; Lundine 2005). Porras and Hope (2005) use conjoint analysis to examine farmers' WTA payments in the Arenal watershed (Costa Rica).

5. In fact, in San Pedro del Norte, Nicaragua, payments to participating farmers were explicitly based on land rental values (Obando 2007).

6. Such studies were rarely undertaken during the design of most existing PES programs, but they are common in the design of new PES programs.

7. In fact, a large number of CV studies in the literature attempt to measure households' WTP for improved water services (for more details on this literature, see Whittington et al. 2009). Even though studies in this literature were not conducted in a PES context, their results are potentially useful for PES program design—as long as respondents were not told that the improvements in service quality would occur by some other means.

8. Note that some studies have been the subject of more than one publication and that some publications cover several studies. We count studies rather than publications.

9. Whittington (2010) notes that most stated preference applications now conducted in less-developed countries never make it to refereed journals, for two reasons. First, most support ongoing policy work and were never intended for distribution to a wide, academic audience. Second, most journals have increasingly stringent publication standards for stated preference articles. A simple reporting of empirical findings of straightforward, professional applications of the methods is of little interest to most editors, however useful it may be for policy work. Many well-executed studies thus never reach a wide audience.

10. Nine of the studies conducted in Mexico were contracted as part of an effort to help jumpstart local PES mechanisms to complement the national PES program whereas the other two were academic studies.

11. Bennagen et al. (undated) also conducted CV studies for a hypothetical PES program in the Philippines, including separate surveys of domestic water users, irrigated rice farmers, and tourists. However, we omitted this paper from our review because it provides no description whatsoever of its methodology.

12. Because of this, we may be underestimating the extent to which the studies in our sample use particular indicators. However, failure to provide sufficient methodological information to enable readers to assess a study's quality could itself be considered an indicator of a poor-quality study.

13. This is reasonable in that almost all PES mechanisms take monetary payments from service users and make monetary payments to service providers. Asquith et al. (2008) describe one of the exceptions: a case in Bolivia in which providers receive beehives and training in honey production as compensation for conservation activities.

14. Some studies were conducted in small communities with correspondingly small sample frames, so small samples sizes do not necessarily indicate inadequate sample sizes.

15. The WTP estimates reported here are in US$ for the year the study was conducted. They have not been normalized to a base year. Expressing these results as a percentage of household income or of current water bills would probably be more meaningful, but few studies provided the information necessary to compute these indicators.

16. Most studies only asked a single debriefing question to respondents who refused to pay: "Why?"

17. We attempted to split our CV-PES study sample to determine whether more recent studies (since 2005) showed improvements over earlier studies, but we did not find any significant differences. The sample is too small, however, for any definitive conclusions in this regard.

18. In our opinion, this was the correct decision, but it is important to recognize that the information collected cannot be used to predict how water users would respond to a change in the volumetric component of a water tariff.

19. Note that for a multiparty negotiation, splitting the maximum WTP of all users between users and upstream providers might not be a deal that would receive majority support from either the user or the providers.

20. Similarly, Griffin et al. (1995) found that the actual behavior of households in Kerala, India, where piped water distribution systems were installed, was predicted accurately by their responses to an ex-ante CV survey.

21. A separate CV study was conducted in Saltillo by Arias-Rojo (2007) after the PES program had been instituted.

22. This assumes that there is a client. Studies of hypothetical PES programs would not generally have a client per se. Even many policy studies, however, may not have a formal client. The CV-PES study in Chaina, Colombia, for example (Moreno-Sánchez et al. 2009), was undertaken by the researchers on their own initiative and was only later presented to the PES program operators (Moreno-Sánchez pers. comm., June 2011). The case studies contracted by CONAFOR in Mexico asked for estimates of the potential benefits of watershed conservation but did not specify that CV should be used. Indeed, the consultants were specifically cautioned against using CV unless other approaches were not feasible or indicated.

References

Adamowicz, W., M. Hanemann, J. Swait, R. Johnson, D. Layton, M. Regenwetter, T. Reimer, and R. Sorkin. 2005. "Decision Strategy and Structure in Households: A 'Groups' Perspective." *Marketing Letters* 16(3–4): 387–99.

Akter, S., J. Bennett, and S. Akhter. 2008. "Preference Uncertainty in Contingent Valuation." *Ecological Economics* 67:345–51.

Alberini, A., K. Boyle, and M. Welsh. 2003. "Analysis of Contingent Valuation Data with Multiple Bids and Response Options Allowing Respondents to Express Uncertainty." *Journal of Environmental Economics and Management* 45:40–62.

Alberini, A., and J.R. Kahn, eds. 2006. *Handbook on Contingent Valuation.* Cheltenham: Edward Elgar.

Alpízar, F., and R. Madrigal. 2007. *Valoración económica de servicios ambientales hídricos en paisajes intervenidos, cantón de Esparza, Costa Rica.* Turrialba: CATIE.

Alpízar, F., and M. Otárola. 2007. "Estimación de la voluntad de pago de clientes de la JASEC para financiar el manejo ambiental de las subcuencas del sistema hidroeléctrico Birris, Costa Rica." In *Valoración Económica, Ecológica y Ambiental: Análisis de Casos de Iberoamérica*, ed. IUCN, 206-228. Heredia: Editorial de la Universidad Nacional.

Amponin, J.A.R, M.E.C. Bennagen, S. Hess, J. Di, and S. de la Cruz. 2007. "Willingness to Pay for Watershed Protection by Domestic Water Users in Tuguegarao City, Philippines." PREM Working Paper No.07/06, Amsterdam: IES.

Arias-Rojo, H.M. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales "Zapalanimé-Saltillo."* Guadalajara, Mexico: CONAFOR.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." *Federal Register* 58: 4601–614.

Asquith, N.M., M.T. Vargas, and S. Wunder. 2008. "Selling Two Environmental Services: In-kind Payments for Bird Habitat and Watershed Protection in Los Negros, Bolivia." *Ecological Economics* 65(4): 675–84.

Bateman, I., R.T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Lee, G. Loomes, S. Mourato, E. Ozdemiroglu, D.W. Pearce, R. Sugden, and J. Swanson. 2002. *Economic Valuation with Stated Preference Techniques: A Manual.* Cheltenham: Edward Elgar.

Barrantes, G., and L. Gámez. Forthcoming. "The Payments for Water Services Program of Heredia's Public Service Utility." In *Ecomarkets: Costa Rica's Experience with Payments for Environmental Services*, eds. G. Platais, and S. Pagiola. Washington, DC: World Bank.

Bennagen, M.E., A. Indab, A. Amponin, R. Cruz, R. Folledo, P.J.H. van Beukering, L. Brander, S. Hess, A. van Soesbergen, K. van der Leeuw, and J. de Jong. n.d. *Designing Payments for Watershed Protection Services of Philippine Upland Dwellers.* PREM Project Report. Amsterdam: IES.

Bennett, M.T., 2008. "China's Sloping Land Conversion Program: Institutional Innovation or Business as Usual?" *Ecological Economics* 65(4): 699–711.

Berggren, M., and S. Stahl. 2003. "Paying for Environmental Services: A Choice Experiment in Turrialba, Costa Rica." Masters thesis, Department of Environmental Economics, Gothenburg University.

Blackburn, M., G. Harrison, and E. Rutstrom. 1994. "Statistical Bias Functions and Informative Hypothetical Surveys." *American Journal of Agricultural Economics* 74:1084–88.

Blackman, A., and R.T. Woodward. 2010. "User Financing in a National Payments for Environmental Services Program: Costa Rican Hydropower." *Ecological Economics* 69(8): 1626–38.

Boland, J., and D. Whittington. 2000. "The Political Economy of Increasing Block Water Tariffs in Developing Countries." In *The Political Economy of Water Pricing Reforms*, ed. A. Dinar. Oxford: Oxford University Press.

Calderon, M., L. Camacho, M. Carandang, J. Dizon, L. Rebugio, and N. Tolentino. 2006. "Willingness to Pay for Improved Watershed Management: Evidence from Metro Manila, Philippines." *Forest Science & Technology* 2(1): 42–50.

Camhi, A., and S. Pagiola, 2009. *Payment for Environmental Services Mechanisms in Latin America and the Caribbean: A Compendium.* Washington: World Bank.

Campos-Benhumea, J.C. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales "El Cielo-Ciudad Victoria."* Guadalajara: CONAFOR.

Carlsson, F., P. Frykblom, and C. Lagerkvist. 2005. "Using Cheap-talk as a Test of Validity in Choice Experiments." *Economic Letters* 89:147–52.

Carson, R., and T. Groves. 2007. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37(1): 181–210.

Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell RC, S. Presser, P.A. Ruud, and V.K. Smith. 1994. *Prospective Interim Lost Use Value Due to DDT and PCB Contamination in the Southern California Bight.* La Jolla: Natural Resource Damage Assessment.

P.A. Champ, K.J. Boyle, and T.C. Brown. eds. 2003. *A Primer on Nonmarket Valuation.* Dordrecht, the Netherlands: Kluwer Publishers.

Cisneros, J, F. Alpízar, and R. Madrigal. 2007. *Valoración económica de los beneficios de la protección del recurso hídrico bajo un esquema de pago por servicios ecosistémicos en Copán Ruinas, Honduras.* Turrialba: CATIE.

Contreras-Benítez, H.A. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales "Sian Ka'an-Cancún".* Guadalajara: CONAFOR.

Cruz M., F.J., and S. Rivera R. 2002 . *Valoración económica del recurso hídrico para determinar el pago por servicios ambientales en la cuenca del Río Calan, Siguatepeque, Honduras.* Tegucigalpa: ESNACIFOR.

Cummings, R., and L. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89(3): 649–65.

Echavarría, M. 2002. "Water User Associations in the Cauca Valley, Colombia: A Voluntary Mechanism to Promote Upstream-downstream Cooperation in the Protection of Rural Watersheds." *Land-Water Linkages in Rural Watersheds Case Study Series*. Roma: FAO.

Engel, S., S. Pagiola, and S. Wunder. 2008. "Designing Payments for Environmental Services in Theory and Practice: An Overview of the Issues." *Ecological Economics* 65(4): 663–74.

González-Ortiz, M.A. 2007. *Estudio de valoración y demanda de servicios ambientales en el área promisoria de servicios ambientales "Copalita-Huatulco."* Guadalajara: CONAFOR.

Griffin, C., J. Briscoe, B. Singh, R. Ramasubban, and R. Bhatia. 1995. "Contingent Valuation and Actual Behavior: Predicting Connections to New Water Systems in the State of Kerala, India." *World Bank Economic Review* 9(3): 373–93.

Hanemann, W.M., and B. Kristrom. 1995. "Preference Uncertainty, Optimal design, and Spikes." In *Current Issues in Environmental Economics*, eds. P.O. Johansson, B. Kristrom, and K.G. Maler. Manchester: Manchester University Press.

Harrison, G.W. 2006. "Experimental Evidence on Alternative Environmental Valuation Methods." *Environmental and Resource Economics* 34(1): 125–62.

Johnson, N., and M.E. Baltodano. 2004. "The Economics of Community Watershed Management: Some Evidence from Nicaragua." *Ecological Economics* 49:57–71.

Kaplowitz, M., and F. Lupi. 2008. *Local Markets for Payments for Environmental Services in Costa Rica.* East Lansing: Michigan State University.

Kristrom, B. 1997. "Spike Models in Contingent Valuation." *American Journal of Agricultural Economics* 79(3): 1013–23.

Krosnick, J.A., A.L. Holbrook, M.K. Berent, R.T Carson, W.M. Hanemann, R. J. Kopp, R.C. Mitchell, S. Presser, P.A. Ruud, V.K. Smith, W.R. Moody, M.C. Green, and M. Conaway. 2002. "The Impact of No Opinion Response Options on Data Quality: Non-attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66:371–403.

Labao, R., H. Francisco, D. Harder, and F. Santos. 2008. "Do Colored Photographs Affect Willingness to Pay Responses for Endangered Species Conservation?" *Environmental and Resource Economics* 40(2): 251–64.

Li, C., and L. Mattson. 1995. "Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation." *Journal of Environmental Economics and Management* 28:256–69.

List, J. 2001. "Do Explicit Warnings Elicit the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions of Sports Cards." *American Economic Review* 91:498–507.

Loomis, J., and E. Ekstrand. 1998. "Alternative Approaches for Incorporating Respondent Uncertainty When Estimating Willingness to Pay: The Case of the Mexican Spotted Owl." *Ecological Economics* 27:29–41.

López-Paniagua, C., M.J. Gonzalez-Guillén, J.R. Valdez-Lazalde, and H.M. de los Santos-Posadas. 2007. "Demanda, disponibilidad de pago y costo de oportunidad hídrica en la Cuenca Tapalpa, Jalisco." *Madera y Bosques* 13(1): 3–23.

Louviere, J.J., D.A. Hensher, and J.D. Swatt. 2000. *Stated Preference Methods: Analysis and Applications.* Cambridge: Cambridge University Press.

Lundine, J. 2005. "An Economic Estimation of Small Land Owner Willingness to Accept a Reforestation Project." Master's thesis, Department of Agricultural, Environmental, and Development Economics, Ohio State University.

Madrigal, R., and F. Alpízar. 2007. *Diseño y gestión adaptativa de un esquema de pagos por servicios eco-sistémicos en Copán Ruinas, Honduras.* Turrialba: CATIE.

Mitchell, R.C., and R.T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington: Resources for the Future.

Moreno-Sanchez, R., J.H. Maldonado, S. Wunder, and C. Borda-Almanza. 2012. "Heterogeneous Users and Willingness to Pay in an Ongoing Payment for Watershed Protection Initiative in the Colombian Andes." *Ecological Economics* 75: 126–134.

Muñoz-Pina, C., A. Guevara, J. Torres, and J. Brana. 2008. "Paying for the Hydrological Services of Mexico's Forests: Analysis, Negotiations and Tesults." *Ecological Economics* 65(4): 725–36.

Obando, M. 2007. "Evolución de la experiencia de los PSA hídricos en Nicaragua: El caso de la microcuenca Paso de los Caballos, Municipio de San Pedro del Norte, Chinandega." *Serie Técnica No.2/2007.* Tegucigalpa: PASOLAC.

Ortega-Pacheco, D.V., F. Lupi, and M.D. Kaplowitz. 2009. "Payment for Environmental Services: Estimating Demand within a Tropical Watershed." *Journal of Natural Resources Policy Research* 1(2): 189–202.

Pagiola, S. 2008. "Payments for Environmental Services in Costa Rica." *Ecological Economics* 65(4): 712–24.

——— 2010. "Payments for Environmental Services in Saltillo." Washington, DC: World Bank.

Pagiola, S., and G. Platais. 2007. *Payments for Environmental Services: From Theory to Practice.* Washington: World Bank.

Peprah, G. 2009. "Investigating the Feasibility of Instituting Payment for Environmental Services (PES) Scheme in Ghana: The Weija Watershed Case Study." Masters thesis, Geo-Information Science and Earth Observation, International Institute for Geo-Information Science and Earth Observation.

Pizano-Portillo, A. 2007. *Estudio de valoración y demanda de agua en el área promisoria de servicios ambientales Cerro Grande-Colima, Jalisco.* Guadalajara: CONAFOR.

Porras, I., and R.A. Hope. 2005. *Using Stated Choice Methods in the Design of Payments for Environmental Services Schemes.* Edinburgh: IIED.

Prabhu, V. 2010. "Tests of Intrahousehold Resource Allocation using a CV Framework: A Comparison of Husbands' and Wives' Separate and Joint WTP in the Slums of Navi-Mumbai, India." *World Development* 38:606–19.

Puente-González, A. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales "Pico de Orizaba-Coatepec."* Guadalajara: CONAFOR.

Ready, R., J. Whitehead, and G. Blomquist. 1995. "Contingent Valuation When Respondents Are Ambivalent." *Journal of Environmental Economics and Management* 29:181–97.

Rodriguez, F., D. Southgate, T. Haab, and J. Lundine. 2009. "Is Better Drinking Water Valued in the Latin American Countryside: Some Evidence from Cotacachi, Ecuador." *Water International* 34(3): 325–34.

Saldivar-Valdés, A. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales 'Cumbres de Monterrey-Monterrey.* Guadalajara: CONAFOR.

Salzman, J. 2005. "Creating Markets for Ecosystem Services: Notes from the Field." *New York University Law Review,* June: 870–961.

Samnaliev, M., T.H. Stevens, and T. More 2006. "A Comparison of Alternative Certainty Calibration Techniques in Contingent Valuation." *Ecological Economics* 57:507–19.

Shultz, S., and B. Soliz. 2007. "Stakeholder Willingness to Pay for Watershed Restoration in Rural Bolivia." *Journal of the American Water Resources Association* 43(4): 947–56.

Southgate, D., T. Haab, J. Lundine, and F. Rodríguez. 2009. "Payments for Environmental Services and Rural Livelihood Strategies in Ecuador and Guatemala." *Environment and Development Economics* 15:21–37.

Southgate, D., and S. Wunder. 2009. "Paying for Watershed Services in Latin America: A Review of Current Initiatives." *Journal of Sustainable Forestry* 28(3–5): 497–524.

Subade, R. 2007. "Mechanisms to Capture Economic Values of Marine Biodiversity: The Case of Tubbataha Reefs UNESCO World Heritage Site, Philippines." *Marine Policy* 31:135–42.

Vargas-Pérez, E. 2007. *Estudio de valoración y demanda de servicios ambientales hidrológicos en el área promisoria de servicios ambientales 'Amanalco-Valle de Bravo.* Guadalajara: CONAFOR.

Wang, H. 1997. "Treatment of Don't Know Responses in Contingent Valuation Surveys: A Random Valuation Approach." *Journal of Environmental Economics and Management* 32(2): 219–32.

Whitehead, J.C., J.C. Huang, G.C. Blomquist, and R.C. Ready. 1998. "Construct Validity of Dichotomous and Polychotomous Choice Contingent Valuation Questions." *Environment and Resource Economics* 11:107–16.

Whittington, D. 2002. "Improving the Performance of Contingent Valuation Studies in Developing Countries." *Environmental and Resource Economics* 22(1–2): 323–67.

———. 2004. "Ethical Issues with Contingent Valuation Surveys in Developing Countries: A Note on Informed Consent and Other Concerns." *Environmental and Resources Economics* 28(4): 507–15.

———. 2010. "What Have We Learned from 20 years of Stated Preference Research in Developing Countries?" *Annual Review of Resource Economics* 2:209–36.

Whittington, D., W.M. Hanemann, C. Sadoff, and M. Jeuland. 2009. "The Challenge of Improving Water and Sanitation Services in Less Developed Countries." *Foundations and Trends in Microeconomics* 4(6–7): 469–609.

Whittington, D., C. Suraratdecha, C. Poulos, M. Ainsworth, V. Prabhu, and V. Tangcharoensathien. 2008. "Household Demand for Preventive HIV/AIDS Caccines in Thailand: Do Husbands' and Wives' Preferences Differ?" *Value in Health* 11(5): 965–74.

World Bank. 2007. "Venezuela Expanding Partnerships for the National Parks System Project: Project Appraisal Document." Report No.37502-VE. Washington, DC: World Bank.

Wunder, S. 2005. "Payments for Environmental Services: Some Nuts and Bolts." CIFOR Occasional Paper No. 42. Bogor: CIFOR.

ISBN 978-0-19-965562-5

9 780199 655625

SFI® Certified Sourcing
www.sfiprogram.org
SFI-00453