Ghana Cocoa Farmers Survey 2004: Report to Ghana Cocoa Board

Francis Teal and Andrew Zeitlin
Centre for the Study of African Economies
University of Oxford

and

Haruna Maamah
ECAM Consultancy, Ltd.
Accra

March 11, 2006

Contents

1 Executive Summary

This report presents the results of the second wave of the Ghana Cocoa Farmers Survey, a panel dataset collected in August of 2004. This survey builds on an earlier round carried out in 2002 by Marcella Vigneri.

The substantial increase in cocoa output during this period represents a success for cocoa policy. Four key dimensions of this expansion are highlighted:

1. The means of expanding output has differed across regions. In particular, the Western Region, where the price of land has been relatively low, has been characterized by *extensive* expansion. Expansion in the Ashanti and Brong Ahafo Regions has been relatively *intensive* in nature.

2. Increases in non-labor inputs such as fertilizer have played a significant role in the expansion. Increases in output and increases in fertilizer appear to have gone hand in hand.

3. The expansion of output was accompanied by a large increase in number of person-days on the farm. Interestingly, work on farms existent as of 2002 seems to have been done by fewer people. This may reflect the diversion of labor onto new farms. It is also borne out in demographic data, which suggest that there has been significant movement out of existing cocoa households. The cocoa boom therefore seems to have fueled investments in education and in employment opportunities in other sectors.

4. The study focuses on two policy variables, the provision of spray machines by the government and the degree of market power exercised at the village level by Licensed Buying Companies (LBCs). In particular it is argued that:

   (a) Efforts to extend the coverage of the government spray program, particularly in the Western Region, are likely to result in higher output levels. It appears that the greatest returns are due to improving the breadth (i.e., the participation rate) rather than the depth (i.e., completion rate) of the spraying regime.

   (b) Structural policies designed to reduce the monopsonistic market power of LBCs may result in improved service provision and higher productivity levels in the villages affected.

These findings are then supported by production function analysis. This allows, *inter alia*, testing of hypotheses regarding returns to scale. Estimates are obtained for the economic returns to fertilizer use and to government programs such as spray machines in a way that would allow explicit cost-benefit analysis. Incorporation of detailed data on spray machine use available only in the latest round of surveys is instructive as to where the returns to intervention in this market would be greatest.

It is argued that the collection of such data should be of mutual benefit to the Cocoa Board and to economic researchers. In particular, two potential benefits of future work are suggested:

1. The long-term growth prospects in the cocoa sector are dependent on whether the increase in output represents the lifting of constraints on farmers' production possibilities or whether it simply reflects a short-term response to the windfall gains of strong cocoa prices. Distinguishing between these effects is only possible if farmers can be tracked across a variety of price conditions.

2. Resampling carried out as part of a third round of data collection would allow a current and geographically representative cross-section of cocoa farmers. This is required for a full understanding of factors driving cocoa expansion in the sector as a whole.

2 Purpose and background

> "Structural rigidities in agriculture remain intractable largely because of the predominance of subsistence farming. Cash crops are subject to prohibitively high transaction costs, as evidenced by large differentials between producer and consumer prices, arising from monopolistic and inadequate distribution systems. Agricultural policy, notwithstanding massive donor support, has failed to address these market impediments."
>
> – Ghana Poverty Reduction Strategy Paper (2003, p. 32)

The Ghana Poverty Reduction Strategy Paper's lament of the shortcomings of agricultural policy interventions suggests a pessimistic view of the prospects for rural growth. Inherent rigidities are seen as effectively blocking the path to development in the rural sector. However, if ever there were a time of sufficient price incentives to overcome such structural obstacles, then this has occurred in the last few years. Indeed, faced in the period 2001/02 to 2003/04 with a 45% increase in producer prices, cocoa producers in Ghana have responded by more than doubling output. This recent dynamism is illustrated in Figure 1.

These developments suggest as many questions as they do answers. One broad set of questions asks *who* was responsible for the increase in output. Of particular interest will be the distribution of these gains across geography, income, and other dimensions of rural Ghana. The current research is also particularly well suited to answer the *how* and the *why* of the supply response. The 'how' may be interpreted as primarily a technological question, focusing on the economic implications of extensive or intensive growth and the relative contributions of labor and nonlabor inputs (fertilizer, pesticides, etc.).

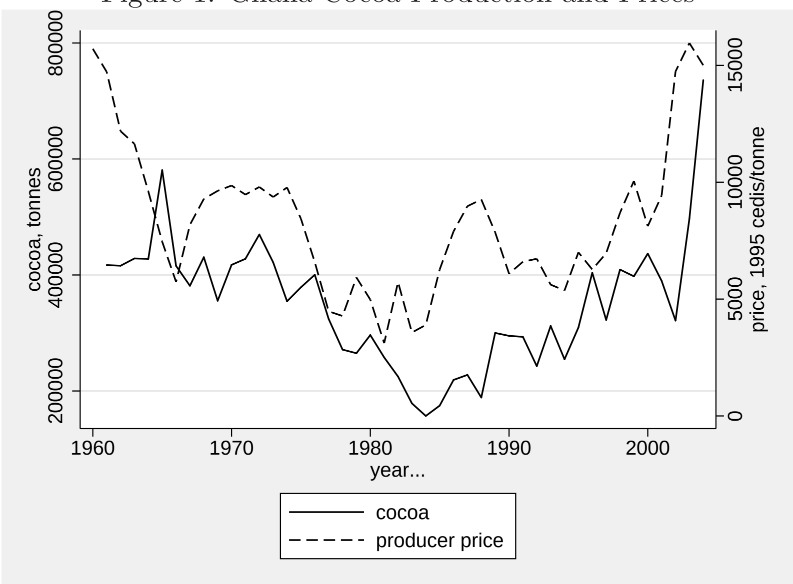
In contrast the 'why' is an economic question and is directly relevant to the description of the Ghana Poverty Reduction Strategy Paper (PRSP). Has the price increase simply provided sufficient incentive to overcome the transaction costs and other rigidities alluded to in the PRSP? Or have the sources of market imperfections actually been alleviated? These contrasting explanations imply either transitory or lasting increases in cocoa output, respectively.[1] It is a stylized fact of economic theory that the failure of one market may lead to inefficient outcomes in others. Consequently, it is a considerable challenge to discern the source of improvements in the functioning of markets. But if there has been a role played by such factors

[1] Clearly, the extent to which a purely price-driven explanation suggests a transitory effect depends on one's long-term forecast of the cocoa price. We do not speculate in this regard, but only point out that there is at least a possibility of reversion to the long-term mean cocoa price.

as government provision of spraying services, then it is entirely possible that there have been complementary effects in other input markets.

Indeed, the possibility of complementarities (or, for that matter, the possibility of substitution) among inputs presents an interesting challenge in the analysis of agricultural production. As famously noted by Marschak and Andrews (1944), the economist is confronted with a fundamentally different problem than the agronomist, since the economist cannot directly control what goes into the ground but rather must take the economic substitutions made by actors with objectives broader than—indeed, potentially divergent from—agricultural output.

Figure 1: Ghana Cocoa Production and Prices



Source: Output, International Cocoa Organization (2003/04);
Prices, Ghana Cocoa Board.

2.1 Previous research

In order to address these questions the present work builds upon a previous round of research conducted at the CSAE by Marcella Vigneri, now at the Food and Agriculture Organization. The Ghana Cocoa Farmers Survey (GCFS) 2002, which Vigneri collected with financial support from the U.K. Department for International Development, studied the 2001/02 cocoa season. The first wave of the GCFS is presented in Vigneri, Teal, and Maamah (2004) and analyzed in Vigneri (2005).

There have been several prior studies of Ghana's cocoa sector. For purposes of assessing both the representativeness of the current survey and changes in production technology over time, it will be particularly useful to compare the findings of the present work with those of Teal and Vigneri (2004). Those authors analyze data from the Ghana Living Standards Survey. Comparison with their findings is undertaken in Section 4, where it is found that—in terms of several key characteristics—the results are broadly similar.

2.2 Data and methods

The findings presented in this paper are based on the two rounds of the Ghana Cocoa Farmers' Survey, collected by the CSAE. The second round of this data collection was carried out by Andrew Zeitlin and funded by the Global Poverty Research Group. Survey rounds took place in the late summer of 2002 and in the early fall of 2004. In each case, questions about cocoa production pertained to the preceding cocoa season—that is, September to August of 2001/02 and 2003/04.

The original 497 farmers surveyed in 2002 were selected as a representative sample of cocoa producers in the Ashanti, Brong Ahafo, and Western Regions of Ghana. Selection was based on those who identify cocoa farming as an income source in the Ghana Living Standards Survey 4 of 1998/99. Farmers come from 25 villages, with the most in the Western Region, followed by Ashanti and Brong Ahafo (see Table 1 below for details). Of the original sample, 443 were successfully re-interviewed in 2004. Additional farmers were added to the sample on the basis of random sampling within the selected villages in order to allow for consistent geographic weighting in future studies. However, attention here will be focused on the members of the panel, that is, those farmers who were interviewed in both rounds. The resulting geographic distribution is given in Table 1.

Table 1: Regional distribution of interviews

Region	2001/02	2003/04	Panel
Ashanti	125	125	113
Brong Ahafo	102	116	98
Western	265	274	232
Total	492	515	443

Source: GCFS 2002, 2004

Sample selection presents a potential source of bias, e.g., if the probability of being reinterviewed is correlated with variables of interest. The data gathered as part of the GCFS 2004 allow us to say a few things about the reasons for sample attrition. Of the 49 farmers lost from the first wave, 7 had died, 12 had migrated, 1 was hospitalized, 1 was inebriated, and 1 was in police custody. Family or others capable of explaining the absence of the others could not be found. While these supplementary data by no means alleviate entirely the concern for sample selection, they suggest may allow for a more full treatment of this issue at a later date. In particular, the implied rate of migration of entire households (as distinguished from an observed out-migration of household members) is of independent interest.

Even in the absence of any attrition from the sample, it is nonetheless important to consider the nature and limitations of the sampling method. The weighting for the stratified sample established in the first round of the survey was based on the most recent, nationally representative information about the cocoa-farming population in Ghana. This was given by the Ghana Living Standards Survey (GLSS) 4. Carried out in 1999, the GLSS 4 provides information on households receiving income from cocoa in the 1997/98 crop season. As a consequence of this, the GCFS sample will not reflect increases in the farming population of the regions considered between the 1997/98 and 2001/02 seasons. Further, because these geographic weights were maintained for the second round of the GCFS, the second wave will not reflect changes to the farming population during the intervening two years.

In summary, this caveat implies that it will be possible only to capture (and perhaps explain) reasons for increases in the production of cocoa farmers who were already receiving income from cocoa production as of 1999. This is likely to understate changes in the industry, as an influx of new farmers—even those with trees in the ground but not yet bearing fruit as of 1999—would not be captured. However, the detailed nature of this microeconomic dataset *does* allow an assessment of the reasons for and means to changing production among this population. It is argued that this will provide important insights into the incentives and constraints facing the sector as a whole.

This report suggests answers to the questions of by what means and for what reasons output has expanded between the 2001/02 and 2003/04 cocoa seasons. A significant intensification in the application of both labor and nonlabor inputs has contributed to the output expansion, along with the underlying increase in cultivated lands. It is argued that the increase in price has provided sufficient incentive to dramatically intensify both labor and nonlabor inputs, even on lands of lesser quality. Our analysis further suggests a role for two aspects of the policy environment. First, government supply of a complementary good in production—namely, the government spraying program—has increased returns to *all* variable inputs. Second, competition among Licensed Buying Companies, which varies in degree from one village to the next, appears to have spurred technological innovation.

To support these claims, we proceed as follows. In Section 3 we describe the central trends in cocoa production in the time period considered. This part of the analysis documents not only changes in output and yields but also changes in variable input use and in the demographic characteristics of cocoa-farming households. Both the productive inputs employed and even the households themselves demonstrate remarkable changes during this time. A cocoa production function is then estimated in Section 4. This uncovers in a more rigorous manner the relative contributions of various inputs. Production function analysis is immediately useful in that it allows an accounting of the sources of growth in the cocoa sector—the 'how' of recent cocoa expansion. This approach is extended to provide tests of the returns to scale in the industry as well as to investigate the effects of government interventions and market structure on farmer productivity. A brief summary of conclusions and directions for future work is provided in Section 5.

3 Trends in cocoa production

Some basic lessons can be drawn from an analysis of changes in key factors. This section investigates three dimensions of cocoa production. First, we describe changes in output and its relationship to the expansion of cocoa-farming land during the time considered. Descriptive statistics suggest that there has been a large increase in cocoa output even among existing farmers, though this does not account for all of the aggregate gains; and that yields are still quite varied across regions and according to farm size within regions, a fact which is suggestive of some degree of market failure. Second, we consider changes in inputs employed over this period. The large increase in fertilizer use, in light of the (unconditional on other inputs) correlation between fertilizer and yields, suggests an important source of output increase. Labor days have also greatly increased, though there is evidence of surplus labor, and there is prima facie evidence for the broad coverage of the government's spraying program but for room for improvement in the depth of that program—with completion of the full course remaining uncommon. Third, we show that changes in output have been accompanied by striking movements in household composition, and that these movements are in some sense contrary to standard predictions of labor-market theory.

3.1 Output, yields, and land

Among the 443 cocoa farmers observed in both waves of the survey, output has increased markedly, but not at the rate suggested by the macro data. Statistics describing cocoa production and its relationship to land and yields are given in Table 2. Mean output has increased by 31% and median output by 38%. Yields have increased by approximately 19% in both medians and means. These correspond to 11% and 15% increases in the median and mean of land in full production, respectively.

Figure 2 shows that these gains have been highly uneven—both in their magnitude and in their nature—across regions. Output has increased most

Table 2: Descriptive statistics: Output, land, and yield

year	variable	median	mean	sd
2001/02	cocoa(kg.)	812.50	1230.94	1350.64
2003/04		1125.00	1609.38	1642.12
2001/02	farm size (ha.)	4.95	6.98	6.71
2003/04		6.07	8.23	8.42
2001/02	cocoa land (ha.)	4.45	6.23	5.92
2003/04		5.26	7.50	7.73
2001/02	bearing cocoa land (ha.)	3.64	5.04	5.09
2003/04		4.13	6.23	7.08
2001/02	cocoa land in full production (ha.)	3.51	4.87	4.89
2003/04		3.89	5.59	6.19
2001/02	output per ha. full production	231.66	313.62	267.27
2003/04		277.99	373.17	314.39

Notes: *bearing* land defined as that where trees not too young to produce; *full production* further excludes trees too old to produce at typical levels.
Source: GCFS 2002, 2004

markedly in the Western Region. This has primarily been driven by relatively large increases in the area of bearing cocoa land in that region. Increases in yields, on the other hand, have been fairly consistent across regions.

This growth on the extensive margin in the Western Region suggests that there may yet be room for growth along the intensive margin: it may still be possible to increase variable inputs such as fertilizer use in the Western Region and bring gains to the yields that exist there. Given the local prices of inputs such as labor and fertilizer, the lower the price of land, the more likely it will be that farmers find it optimal to increase output extensively.

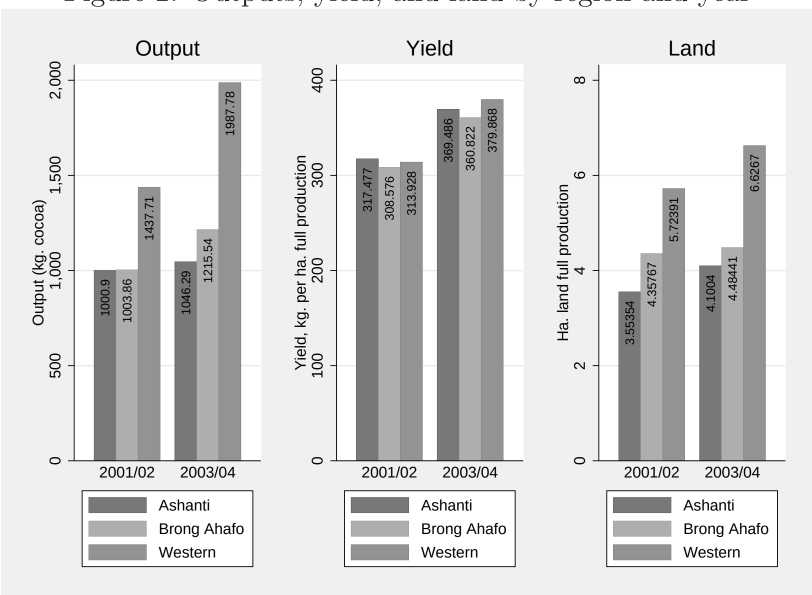
This view is supported by the fact that the price of land (proxied by the reported values per ha. of farmers' plots) is lowest in the Western Region. Median land value in 2003/04 was 16.5 million Cedis per hectare in the Western Region, compared with 20.6 and 21.2 million Cedis in Ashanti and Brong Ahafo, respectively. More precisely, we find the weighted mean of land values (weighted by farm size) to be 26 million, 35 million, and 22 million Cedis in the Ashanti, Brong Ahafo, and Western Regions. Of course, these figures should be regarded with caution, since it is generally quite difficult to get accurate reports of such saleable assets, but the interregional differences may nonetheless be informative. This is taken as suggestive evidence to support the idea that land prices constrain extensive expansion in the Ashanti and Brong Ahafo regions.

There are two important reasons for consideration of the dispersion of yields and not just their averages. These are suggested in Figure 3, which highlights the relationship between total farm size for each region and year. First, the large variance in yields even within a particular region and year gives an indication of just how far most producers fall from the technological frontier. This is related to the concern expressed in the Ghana Poverty Reduction Strategy Paper (2003, p. 69), which argues that "estimates of achievable yields... show that there is a wide gap between actual and potential yields." Similarly, Teal and Vigneri (Teal and Vigneri 2004) argue find that yields in Ghana are typically far below the experimental frontier as established in Ghana and elsewhere.

Secondly, the relationship between yields and farm sizes may be helpful in ascertaining not just the potential for production but also the means to achieving this frontier. The relevant economic question asks why farmers do not always apply techniques giving such high yields. One potential answer was suggested above: namely, that where land prices are low, farmers find it optimal to increase output extensively rather than intensively. However, if the returns to intensive production techniques are high—and the variation in yields does suggest that there is great scope for improvement—then farmers who do not apply those techniques must either face higher input prices or be constrained by limited credit, etc., such that they are prevented from accessing inputs even when the economic returns justify their use.

Indeed, there exists a large economic literature dating back to Sen (1962) that seeks to explain the "inverse relationship" between farm productivity and size. It is commonly argued (see, for example, a discussion of this argument in Benjamin (1995)) that this can result from market imperfections. For example, if the owners of relatively small plots cannot find sufficient

Figure 2: Outputs, yield, and land by region and year

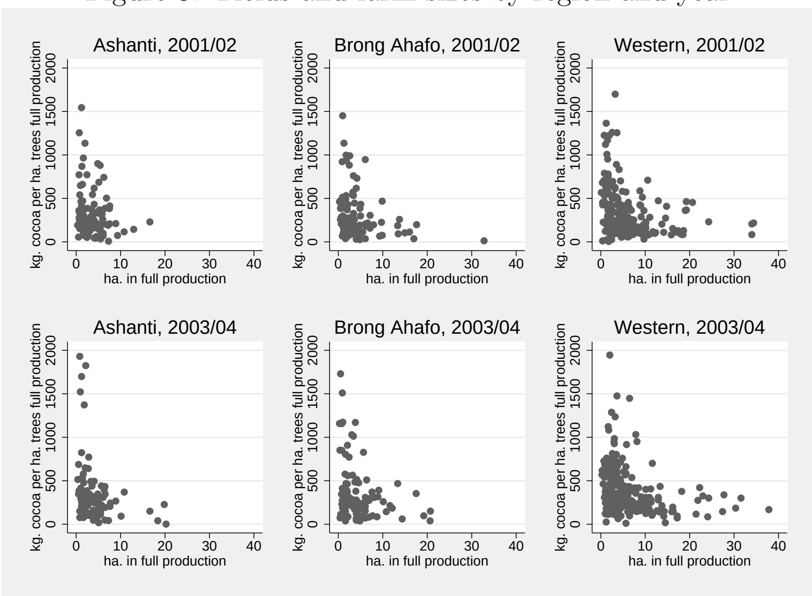


Source: GCFS 2002, 2004.

work as paid laborers on relatively large plots, then they will only be able to use their labor on their own. Since these small farmers are endowed with more labor per hectare, yields will tend to be higher on small farms.

This makes clear that discussion of the means to achieving the productive frontier requires an understanding of how the market for inputs works. Constraints on input use, ranging from regional differences in factor prices to capital-market imperfections, appear to underly the different outcomes observed. With this point in mind, we turn attention to a description of changes in input use in Section 3.2.

Figure 3: Yields and farm sizes by region and year



Source: GCFS 2002, 2004.

3.2 Variable inputs

Tables 3 summarizes the changes in the application of non-labor and labor inputs during the period studied. In this section, we focus particular attention on changes in fertilizer, labor, and spraying practices. These are taken in turn below.

3.2.1 Fertilizer use

The rise in fertilizer use is immediately striking. Not only has there been an unconditional rise in fertilizer use throughout the population, but this is driven in large part by new fertilizer use from farmers who applied no fertilizer in the 2001/02 season. The percentage of farmers using no fertilizer dropped from 90% to 52% during this period.

Table 3: Non-labor inputs

	2001/02			2003/04		
	median	mean	sd	median	mean	sd
fertilizer (50 kg. bags)	0.00	0.45	2.57	0.00	5.14	11.51
% use no fertilizer	–	0.90	–	–	0.52	–
insecticide (l.)	1.00	0.86	0.35	1.00	0.94	0.24
% use no insecticide	–	0.14	–	–	0.05	–
use any spray machine	–	0.82	–	–	0.98	–
use govt. spray machine		n/a		–	0.92	–
use private spray machine		n/a		–	0.88	–
hybrid use	–	0.79	–	–	0.67	–
hybrid share	0.50	0.58	0.39	0.50	0.47	0.41

Notes: Variables defined as the percentage of the sample using a particular input (fertilizer, insecticide, spray machines, and hybrid trees) are presented as means of an indicator variable for the population. Variable *hybrid use* takes value of unity if any hybrids are used on the farm; variable *hybrid share* gives the percentage of hybrid trees on the farm. Source: GCFS 2002, 2004

The scatter plots presented in Figure 4 further suggest that this rise in fertilizer use may have played an important role in the overall output increase. The right panel of Figure 4 shows a strong positive relationship between the intensity of fertilizer application and the resulting yield. While one clearly cannot infer causality from this bivariate relationship—because other variables may be correlated with fertilizer use and it may be those variables that have a causal effect on yields—the regression results presented in Section 4 bear out this idea.

The left panel relates fertilizer intensity to farm size: the clear negative relationship suggests that fertilizer tends to be spread more thinly on large farms. This may be helpful in explaining the 'inverse productivity' relationship referred to above, but some care is required in interpreting these results. Taken at face value, this finding would suggest that by lifting whatever constrains farmers from accessing sufficient quantities of fertilizer for intensive application on large farms, one would be able to come closer to achieving experimental yields on farms.

However, a cautionary note is required here. One must be mindful of the possibility that the effects of fertilizer use are determined by its interaction with soil quality or other unobserved inputs. For example, it is possible that fertilizer only has such a strong impact on output when soil quality is high. If it is also the case that high-quality plots tend to be small ones, then such complementarities between (unobserved) land quality and fertilizer application could lead to overstatement of the benefits of increasing fertilizer use. Stated differently, such an interaction might imply that extensive expansion—by lowering average land quality—actually increases the

Figure 4: Fertilizer intensity by farm size and implications for yield

Notes: Conditional on positive fertilizer use. Includes 2001/02 and 2003/04.
Source: GCFS 2002, 2004.

amount of fertilizer required to achieve experimental yields. This is exemplary of a more general difficulty in interpreting cross-sectional relationships: one must be mindful of the possible correlation between input use and unobserved characteristics, such as land quality or farmer ability. These issues are taken up further in Section 4.

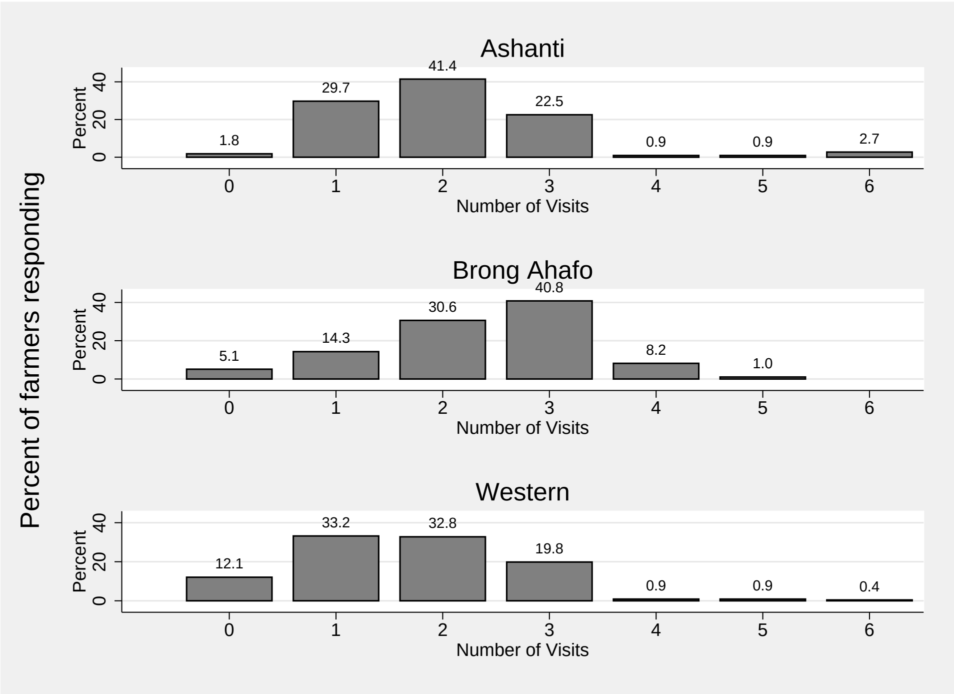
3.2.2 Use of spray machines

The extent of coverage of the government's spray program is of primary interest from a policy perspective. It is necessary to establish the extent of coverage along two dimensions: the extent of access to the program and the rates of completion of the recommended course of treatment. With this information in hand, it is possible to estimate the productive impact of this program, and the implications of the two dimensions considered in suggesting scope for improvement. This analysis necessarily involves a further complexity, as farmers may respond to the provision of this input by supplying more or less of other inputs themselves. The degree of substitution or complementarity has important implications for the net impact of the program. It may be the case that government provision of some inputs has important impacts in *other* input markets, by increasing the returns to complementary inputs.

Data on the government spray program were not collected for the 2001/02 season, so we restrict our attention in this section to consideration of 2003/04.

For that season, Table 3 reports that 92 percent of farmers interviewed had received some assistance from a government spray gang. This is to be compared with 88 percent who used a privately owned spray machine; the net result was that 98 percent of farmers had used some form of mechanized spray machine.

Figure 5: Government spray program: Visits by region, 2003/04



Source: GCFS 2002, 2004.

Regional coverage, as indicated in Figure 5, was somewhat uneven. The mean number of visits by a government spray gang in the Ashanti, Brong Ahafo, and Western Regions was 2.0, 2.4, and 1.7, respectively. The relatively sparse coverage of the spray program in the Western Region is also reflected in a higher percentage of farmers (12.1 percent) who received no visits at all. This stands in contrast with comparable figures of 1.8 percent and 5.1 percent for the Ashanti and Brong Ahafo Regions. Since growth of output in the Western Region has been relatively extensive in nature during this time period, this suggests that there may be scope for improvement by broadening and deepening the coverage of the program in that Region. This argument is supported below in Section 4, where we find in the context of a production function analysis that the *breadth* of coverage—that is, the percentage of people receiving at least one visit from the government—has a particularly significant (economically and statistically) effect on output.

3.2.3 Labor use

A striking finding of the GCFS is the increase in labor usage between 2001/02 and 2003/04: adult labor-days have increased by 129 percent in response to the price increase. Labor supply during this period is documented in Table 4. We present two measures of labor use on the farm: the number of person-days and number of people employed on the farm.[2] In this section, we discuss first the allocation of labor across tasks on the farm, the implications of the increase in labor days for labor productivity, and the relationship between labor use and farm size. The section concludes with a discussion of the divergence between the person-days and the person-count measures of labor supply; this is argued to be suggestive of the linkages between cocoa and other industries.

Table 4: Labor inputs

	2001/02			2003/04		
	median	mean	sd	median	mean	sd
Person-days						
tot. adult days	141.00	319.21	607.61	324.00	626.18	923.01
men days	6.00	42.05	82.80	86.00	164.39	500.73
women days	21.00	43.80	69.50	40.00	92.55	236.13
annual days	0.00	36.23	239.82	0.00	64.57	271.35
contract days	40.00	160.46	427.83	108.50	265.32	434.39
nnoboa days	0.00	29.94	146.08	0.00	22.73	89.29
Worker counts						
tot. adult count	17.00	24.49	30.82	7.00	9.46	10.08
men count	1.00	2.60	3.82	1.00	1.35	3.26
women count	2.00	2.86	2.71	1.00	1.28	3.27
annual count	0.00	1.42	5.85	0.00	0.55	1.51
contract count	7.00	10.90	22.11	3.00	4.42	5.12
nnoboa count	0.00	6.71	13.76	0.00	1.86	5.11
Tasks						
days land prep	0.00	33.29	215.11	60.00	149.02	492.56
days planting	0.00	23.42	68.72	30.00	75.83	144.50
days maint./weeding	80.00	196.90	436.78	126.00	235.13	357.66
days app. inputs	4.00	12.00	38.95	16.00	43.67	110.78
days harvesting	28.00	57.08	119.32	60.00	126.59	262.00

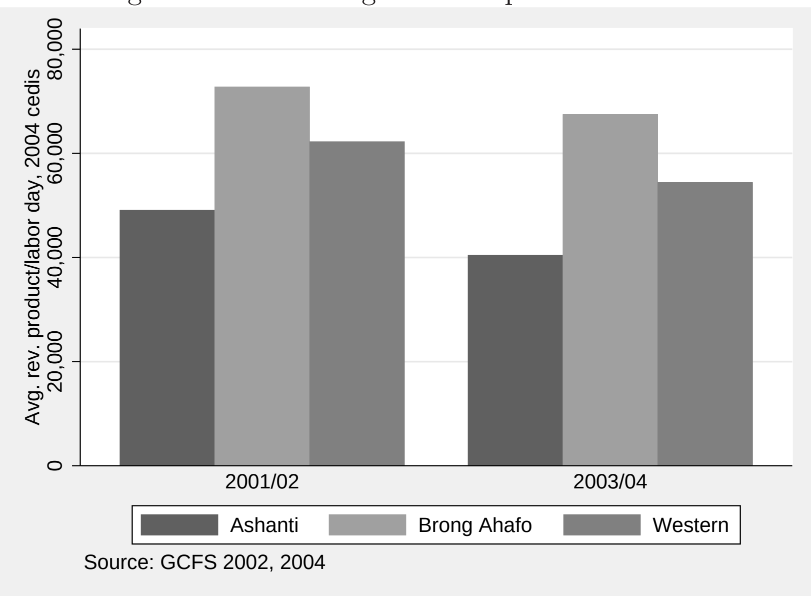
Source: GCFS 2002, 2004.

Table 4 reveals that the increase in labor inputs on the farms studied has been generally consistent across labor types, with the greatest percentage

[2]The counts of people working on a farm are imputed, using the maximum number of people belonging to each category who work on any task.

change in the labor days of household men and the sole decrease in the use of *nnoboa* labor. In light of this, it is remarkable that the average revenue product per labor day, as shown in Figure 6 has remained relatively stable during this time—this suggests a very large increase in the *total* return to labor. Differences across regions reflect both differences in intensities of other inputs as well as land productivity and other unobserved factors. The general trend of an increase in the return to labor is compounded by the fact that the number of workers (as opposed to the number of days worked) has actually decreased during the time period considered. It appears that fewer workers are working more days, a fact consistent with the demographic trends discussed in Section 3.3.

Figure 6: Real average revenue product of labor



Source: GCFS 2002, 2004.

The lower section of Table 4 displays the allocation of labor across tasks. It is indicative of positive prospects for the long-term health of the industry that the greatest (absolute) increases in labor application have occurred with respect to land preparation and planting. The bulk of labor was allocated to maintenance and weeding in both years studied, with land preparation and harvesting constituting relatively larger proportions of labor use in 2003/04. Increases in land preparation and planting are suggestive of positive prospects for the long term health of output. The increase in days spent applying inputs is also large in percentage terms. While this still represents a minority of farmers' time, it reflects the increasing importance of non-labor inputs in cocoa production.

Lastly with respect to labor, it is worth relating the discussion of labor inputs to the question of decreasing returns to farm size. It should be noted that a regression of (the log of) labor days on (the log of) productive cocoa land yields a coefficient of 0.53 with a standard error of 0.04 (804 observations across 2 years). Thus a simple F-test suffices to firmly reject the hypothesis that labor inputs increase proportionally with farm size.[3] To the extent that this reflects a failure of the local labor market—in the sense that farmers cannot hire in or rent out their labor in order to equalize returns across farms—then it may provide an explanation for the relationship between farm sizes and yields (see, for example, Udry (1999)).

3.3 Demographic change

The collection of data on farmer and household demographic characteristics allows a discussion of linkages between cocoa market outcomes and these traits, which are of inherent interest from the perspective of a policy maker concerned with farmer welfare. Such demographic variables combine characteristics of both policy *instruments* and policy *objectives*. Thus, for example the regression analysis presented in Section 4 tests for an impact of primary school completion on cocoa output. In this section, we confine attention to three basic trends in household characteristics.

Table 5: Household Summary Statistics

	2001/02			2003/04		
	mean	sd	obs	mean	sd	obs
HH size	6.9	2.6	443	5.7	2.5	443
Age HH head	51	15	439	53	15	439
HH head male	.82	.38	441	.83	.38	434
HH head primary school educ.	.73	.44	441	.67	.47	433
Pct. income from cocoa	75	22	443	79	21	443
Has other income source	.86	.34	419	.64	.48	443

Source: GCFS 2002, 2004

First, it may be insightful to highlight a few characteristics of the farmers interviewed. They are primarily (approx. 82 percent) male. Farmers were on average 51 years old in the first round of the survey, though clearly the age profile in the panel (without considering new observations) will continue to mature. Approximately 70 percent of farmers have completed primary school; differences in this variable across years may reflect misreporting.

Second, as is consistent with the observed cocoa boom, cocoa has become increasingly important to the panel members as a source of income. This

[3] Specifically, the $F(1, 802)$ test of this restriction gives an F-statistic of 132.27 with an associated p-value of 0.00. Care should be exercised in interpreting this result, since it does not condition on other inputs or characteristics of the farm: this is a purely descriptive statement, rather than a causal one.

growing economic significance is captured by an increase in the percentage of income accruing from cocoa, but perhaps more interesting is the sudden movement toward specialization. 22 percent of the farmers considered reported having another source of income in 2001/02 but did not do so in 2003/04.

Third, there has been a marked decrease in household size within the sample. To some extent this represents normal changes in the life-cycle composition of households (for example, children getting married and moving out). However, evidence presented in Table 6 suggests that the decrease in household size largely reflects migration, both within the country and without.

Table 6: Migration from cocoa households

	Ashanti	Brong Ahafo	Western
number hhs	125	116	275
mean migrants/hh	0.66	0.78	0.73
Geographic destination			
rural Ghana	21%	47%	41%
urban Ghana	65%	49%	58%
international	15%	4%	1%
Selected reasons for migration			
education	35%	36%	47%
employment	39%	47%	32%
apprenticeship	0%	10%	13%

Source: GCFS 2002, 2004

These patterns highlight the connection between rural and urban labor markets. One hypothesis is that the benefits of a good cocoa price are used in part to invest in education or as an opportunity to diversify income sources. Determining the household characteristics that lead to migration and the permanence of this decision remain important areas for further study.

3.4 Licensed Buying Companies

Table 7 describes the findings with respect to the presence and impact of Licensed Buying Companies (LBCs) in the villages studied. From an economic policy perspective, the role of competitive market structures in fostering growth remains something of an open question. The GCFS provides evidence that competition among LBCs exists but varies in degree across villages. This variation will be exploited to test the impact of market institutions on productivity in Section 4.

Farmers reported the presence of slightly more than 3 LBCs in their villages in each year. They sold some amount of cocoa to an average of

Table 7: Institutional Context: Summary Statistics

year	2001/02			2003/04		
variable	mean	sd	obs	mean	sd	obs
no. lbcs in village	3.2	1.5	443	3.3	1.7	443
no. lbcs sold	1.5	1.1	443	1.2	0.44	443
lbc switches		n/a		1.1	1.2	443
lbc hhi	0.50	0.19	443	0.54	0.21	443
lbc loans	0.16	0.37	443	0.22	0.41	443
lbc inputs on credit	0.045	0.21	443	0.11	0.31	443
lbc input subsidies	0.045	0.21	443	0.16	0.37	443

Source: GCFS 2002, 2004

1.5 and 1.2 LBCs in 2001/02 and 2003/04, respectively. This is suggestive of some competition, since it would seem that farmers who sell cocoa to more than one LBC are less likely to be 'captive' clients. This aspect of the relationship between farmers and their buyers is investigated further by the creation of a variable *lbc switches*. This variable is the sum of two components: the number of LBCs that a farmer sold to in 2001/02 but did not sell to in 2003/04, plus the number of LBCs that a farmer sold to in 2003/04 but did not sell to in 2001/02. On average, farmers dropped or added more than one LBC in this period, a fact that suggests considerable turnover in these relationships despite the stable number of LBCs in the villages in this period.

For use in the econometric analysis, a measure of LBC competition that can be constructed on the basis of cross-section data alone is required. Thus the variable *lbc hhi* is defined as a Hirschman-Hirfindahl Index (HHI) of industry concentration at the village level. This measure, which is commonly used in the study of industrial organization, is defined here as the sum of squared market shares in the village, with market shares ranging from zero to one.

$$lbc\ hhi_v = \sum_{j=1}^{N_v} s_{i,v}^2,$$

where $s_{i,v}$ is the market share of the i^{th} LBC in village v, and N_v is the total number of LBCs in that village. A high HHI therefore reflects a monopsonistic industry, where a few LBCs have a lot of power, while a low HHI reflects a relatively competitive environment.

Finally, some indication of possible mechanisms by which LBC policies could affect productivity are suggested by the *loans*, *inputs on credit*, and *input subsidies* variables. These take a value of one if the farmer has received the relevant services from any of his or her LBCs.[4] The data gathered

[4]More precisely, the variable *loans* takes a value of one if the farmer actually received

suggest that LBCs have increased the provision of services across the board.

4 Estimated determinants of cocoa output

To provide a more rigorous foundation for the analysis of changes in cocoa production, we estimate a production function using the GCFS data. This section presents the basic specification employed before discussing the results and their implications for policy-relevant hypotheses.

Following Teal and Vigneri (2004), the production of cocoa, Y, estimated is of the Cobb-Douglas form

$$\ln Y = \beta_0 + \beta_1 \ln(farmsize_{it}) + \beta_2 \ln(laborers) + \beta_3 \lambda_{it} + \beta_4 \ln(A_{it}) + \beta_5 X_{it} + \varepsilon_{it},$$

where λ_{it} gives the fraction of paid workers employed by farmer i at time t, A_{it} is a vector of observed productivity-enhancing inputs (such as use of insecticide, government spray machines), and X_{it} is a vector of local institutional variables and farmer characteristics hypothesized to influence productivity. Controls for region are included in X. Under this specification, coefficients corresponding to logarithmic variables can be interpreted as elasticities (e.g., β_1 gives the percentage increase in output expected from a one percent increase in farm size).

The parameters of this specification are estimated by Ordinary Least Squares (OLS), with results from this procedure presented in Table 8. Column (1) gives the estimates for the entire sample. Columns (2) and (3) present two specifications for analysis of the 2003/04 data, and Column (4) estimates the production function based on 2001/02 data for comparison. Several points of emphasis are highlighted below.

First, it is noteworthy that the coefficients on land, labor, and fertilizer (the primary variable non-labor input captured by the survey) from the pooled estimation (Column (1)) correspond closely to the findings of Teal and Vigneri (2004, Table 5). As in their pooled OLS estimates, we find a coefficient on farm size significantly below unity, suggesting that conditioning on observed inputs is not sufficient to explain the inverse productivity relationship. As in Teal and Vigneri, we find that the coefficient on labor input is relatively small and difficult to identify with statistical significance by use of OLS. In the work of those authors and in that of Zeitlin (2005), it is found that Instrumental Variables (IV) approaches can identify a stronger effect of additional labor. This suggests that the small point estimate on labor inputs may be caused by attenuation bias, as this input tends to be measured with relative imprecision.

Fertilizer used has a strong economic and robust statistical effect in the analysis. Indeed, this result closely parallels the finding of Teal and Vi-

this service, while the latter variables take a value of one if the farmer cites the potential receipt of this service as a reason for selling to the LBC concerned.

Table 8: OLS estimates of Cobb-Douglas production

	pooled	2003/04	2003/04 detail	2001/02
	(1)	(2)	(3)	(4)
Const.	5.623***	5.516***	5.482***	5.242***
	(.374)	(.423)	(.448)	(.524)
ln productive ha. cocoa	.602***	.540***	.551***	.622***
	(.050)	(.060)	(.056)	(.055)
ln count workers	.092**	.088	.091	.183***
	(.038)	(.071)	(.076)	(.057)
fraction paid workers	.005	-.134	-.117	.153
	(.128)	(.177)	(.164)	(.125)
ln fertilizer (50 kg bags)	.233***	.231***	.230***	.329**
	(.065)	(.072)	(.075)	(.145)
no fertilizer used	-.351***	-.300***	-.308***	-.564***
	(.064)	(.103)	(.102)	(.184)
insecticide used	.112	.042	.061	.135
	(.122)	(.186)	(.262)	(.105)
any spray machine used	.143	.680*		.014
	(.132)	(.357)		(.100)
govt and private spray mach used			-.554	
			(.462)	
private spray mach			.514	
			(.503)	
govt spray mach.			.777*	
			(.449)	
govt completed 3+ sprayings			-.155	
			(.125)	
LBC HHI	-.667*	-.904**	-.950**	-.367
	(.374)	(.380)	(.381)	(.445)
pct land owned	-.027	-.021	-.033	-.073
	(.135)	(.115)	(.116)	(.197)
share hybrid trees	-.074	.032	.055	-.174
	(.076)	(.098)	(.101)	(.106)
black pod	-.188**	-.083	-.071	-.349**
	(.090)	(.085)	(.088)	(.172)
mistletoe	.057	-.007	.013	.081
	(.075)	(.101)	(.101)	(.085)
hhh male	.191**	.173*	.175*	.165
	(.087)	(.092)	(.095)	(.120)
age hhh	.001	-.002	-.002	.004
	(.003)	(.003)	(.003)	(.004)
primary educ hhh	.012	.035	.042	.006
	(.078)	(.086)	(.089)	(.118)
BA	-.008	-.020	.017	.083
	(.267)	(.280)	(.284)	(.314)
Western	.148	.202	.200	.192
	(.192)	(.212)	(.206)	(.230)
Obs.	812	396	394	416
R^2	.486	.522	.522	.478
F statistic $(N, 1)$	.826	3.506	2.932	.725
p-value	.372	.074	.1	.403

Notes: Dependent variable is $\ln(kg.cocoa)$. F-test and p-value refer to test of hypothesis (constant returns to scale) that $\beta_{land} + \beta_{labor} + \beta_{fertilizer} = 1$. Heteroskedasticity-robust standard errors reported, clustered at village level. Source: GCFS 2002, 2004 gneri with respect to input expenditure (those authors find fertilizer to have a coefficient of 2.46 and standard error of 0.036 in their pooled regression). It should, however, be noted that the Cobb-Douglas specification is not well suited to handling inputs that often take a value of zero. Since one cannot take the logarithm of fertilizer used when this takes a value of zero, in such cases we use an indicator variable to demonstrate the effect of using zero fertilizer and impose a value $\ln(fertilizer) = \ln(\text{median}(fertilizer|fertilizer > 0))$ when no fertilizer was used. Thus, by construction, the coefficient on the variable '*no fertilizer used*' identifies the average difference in output between farms that used no fertilizer and farms that used the median positive amount of fertilizer observed. When this variable input is accounted for in the pooled regression, we cannot reject the hypothesis of constant returns to scale in land, labor, and fertilizer, as indicated by the F statistic and p value. This hypothesis is accepted for 2001/02 but rejected at the 10% confidence level for both specifications used with the 2003/04 data alone.

Additional information gathered in the GCFS 2004 and used in Column (3) allows a close analysis of the impact of the government spray program. For this year, the data allow separate identification of the use of private and government mechanized spray machines. To capture the possibility of complementarities between these two types, we include distinct indicator variables taking a value of one if government machines, private machines, or both are used. We find a significant effect from use of government spray machines relative to the counterfactual in which the farmer would not otherwise have used a spraying machine. For simplicity, consider as a baseline the question of the productive impact of receiving fewer than 3 visits from a government spray gang. Then to compare the average productivity of farmers using no spray machines of any kind with one who uses both government and private machines, one should sum the coefficients on these three variables. Going from no spray machine use to both types gives a point estimate of 0.737—a 74% productivity increase. On the other hand, the effect on expected percentage increase in productivity resulting from being visited by a government spray team for a farmer *conditional* on using a private machine in any case is estimated as $0.777 - 0.554 = 0.223$. Likewise, the net effect of using a private machine given that one already receives the assistance of a government spray team is negative and insignificant: $-0.554 + 0.514 = -0.040$. While the results do show a positive effect from going from zero to one or more visits from the government spray team, a positive effect of receiving three or more visits (representing completion of the intended course) cannot be identified.

The pooled analysis in Column (1) suggests a negative and significant effect of the incidence black pod (measured here as a binary indicator variable), even after controlling for spraying efforts. This would seem to suggest an area of concern for the Cocoa Board. This concern should be mitigated to some extent by the fact that the BP effect is not significant in the 2003/04 cross section—the pooled effect appears to be driven by the results for 2001/02.

As measured by the *LBC HHI* variable, market structure seems to have significant implications for productivity. The relevant estimated coefficient in Column (1) implies that a decrease in market concentration by a single standard deviation would have a 113% increase in productivity, other factors held fixed. Econometric complications involved in making this assessment and mechanisms by which such a competitive effect may obtain are the subject of ongoing research; see Zeitlin (2005) for details. That analysis shows that this effect appears robust to econometric concerns, though the mechanism through which this effect operates (to name a few possibilities: improved credit, incentives, inputs access, or technology choice) remains the subject of ongoing study.

Finally, it is interesting that there are no statistically significant differences between regions when output is conditioned on the factors considered above. This suggests that differences in method of cultivation (e.g., intensive vs. extensive methods) across these regions can be accounted for in terms of observable inputs.

The rather complex argument required above to interpret the effects of spraying highlights an more general caveat to the interpretation of these results. The partial correlations uncovered by multivariate OLS analysis embody a *ceteris paribus* assumption: they describe the effect on output of increasing a particular input *assuming all other inputs remain constant.* As discussed above, this is generally not a reasonable assumption for policymakers to make. Policy interventions typically alter incentives for farmers. In response to changes in the provision of spraying services, for example, farmers may decide to increase complementary inputs (or decrease others, for that matter). Thus the policymaker must be careful to anticipate the likely reaction of economic agents. Given a particular policy intervention, these may be teased out by careful use of a panel dataset such as the present one. Moreover, some of the inputs considered here may be endogenous to output—e.g., because they are correlated with some unobserved aspect of land quality—and so produce biased estimates. This suggests a further argument for a fixed effects and/or instrumental variables approach to estimating cocoa production; for further exploration of these techniques in the present context see Zeitlin (2005).

5 Conclusions and future objectives

This report has presented results from the second wave of the Ghana Cocoa Farmers Survey. Descriptive statistics reveal—in contrast to the characterization by the Ghana Poverty Reduction Strategy Paper—a dynamic rural environment, one which has responded to structural incentives such as changing relative prices. Key policies questions such as the provision of government spray services and the significance of LBC competition can be analyzed by use of these data.

Important questions remain, suggesting two directions in particular for future research. First, there remain questions of whether rural market failures, which may underly the low yields on relatively large farms, have been overcome. Since the present analysis has taken place over the course of a price boom, it is difficult to distinguish improvements in market function from short-run responses to the windfall of a real price increase. This can be overcome by continuing to track the panel of farmers across different price conditions.

Second, the survey work to date has been limited in its ability to provide a geographically representative sample of the regions studied. As discussed in Section 2.2, this is because the cocoa-farming population had essentially outgrown the initial sample. An understanding of the role of new farmers as a source of growth for the industry can be achieved by using more recent census information to re-weight the sample in order to improve its geographic coverage.

It is hoped that this report serves to highlight key areas and the value of such data for cost-benefit analysis and economic policymaking more generally. Future collaborative efforts on the collection and analysis of data have the potential shed further light on these questions of great mutual interest.

References

Benjamin, D. (1995): "Can unobserved land quality explain the inverse productivity relationship?," *Journal of Development Economics*, 46, 51–84.

Government of Ghana (2003): "Ghana Poverty Reduction Strategy, 2003–2005: An Agenda for Growth and Prosperity," 1/11/2005, http://siteresources.worldbank.org/ghanaextn/Resources/Ghana_PRSP.pdf.

International Cocoa Organization (2003/04): "Quarterly Bulletin of Cocoa Statistics," Vol. XXX, No. 4.

Marschak, J., and J. William H. Andrews (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12(3/4), 143–205.

Sen, A. K. (1962): "An aspect of Indian agriculture," *Economic Weekly*, 14(4–6), 243–266.

Teal, F., and M. Vigneri (2004): "Production Changes in Ghana Cocoa Farming Households under Market Reforms," Centre for the Study of African Economies, Working Paper No. 2004-16.

Udry, C. (1999): "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture," in *Trade, Growth, and Development: Essays in honor of Professor T. N. Srinivasan*, ed. by L. K. Raut, and G. Ranis, pp. New York and Oxford Elsevier Science, North–Holland, 1999; 363–93. Econ Growth Center, Yale U.

Vigeneri, M., F. Teal, and H. Maamah (2004): "Coping with Market Reforms: Winners and Losers Among Ghanaian Cocoa Farmers," Report to the Ghana Cocoa Board.

Vigneri, M. (2005): "Trade Liberalisation and Agricultural Performance: Contrasting Macro and Micro Evidence from Ghana," DPhil thesis, University of Oxford.

Zeitlin, A. (2005): "Market Structure and Productivity Growth in Ghanaian Cocoa Production," Paper submitted for Cornel/ISSER/World Bank Conference on Shared Growth in Africa.