IICA

Consultant Final Report
IICA/EMBRAPA-PROCENSUL II

Ex-post evaluation system for
PROCENSUL II

ESCRITÓRIO NO BRASIL

Consultant Final Report
IICA/EMBRAPA-PROCENSUL II

Ex-post evaluation system for PROCENSUL II

Série Publicações Miscelâneas Nº A4/BR-87-019
ISSN-0534-0591

Ex-post evaluation system for PROCENSUL II

Consultant Final Report
IICA/EMBRAPA-PROCENSUL II

Duncan Thomas

Brasília, abril de 1989

INSTITUTO INTERAMERICANO DE COOPERAÇÃO PARA A AGRICULTURA
EMPRESA BRASILEIRA DE PESQUISA AGROPECUARIA

Thomas, Ducan.
Ex-post evaluation system for PROCENSUL II. Consultant final report IICA/EMBRAPA-PROCENSUL II/por Ducan Thomas.- Brasília : IICA/EMBRAPA, 1989.
23 p. (IICA. Série Publicações Miscelâneas, A4/BR 87-19)
ISSN 0534-0591

1. Avaliação - Sistema. I. Título. II. Série

AGRIS E16
CDU 631.16

APRESENTAÇÃO

A reprodução e difusão dos Relatórios de Consultores, no âmbito restrito das Diretorias das Unidades do Sistema Nacional de Pesquisa Agropecuária, vinculado à EMBRAPA, tem como objetivo principal o de divulgar as atividades desenvolvidas pelos consultores e as opiniões e recomendações geradas sobre os problemas de interesse para a pesquisa agropecuária.

As atividades de consultoría são realizadas no âmbito do Projeto de Desenvolvimento da Pesquisa Agropecuária e Difusão de Tecnologia na Região Centro-Sul do Brasil - PROCENSUL II, financiado parcialmente pelo Banco Interamericano de Desenvolvimento - BID e a EMBRAPA conforme os contratos de Empréstimo 139/IC-BR e 760/SF-BR, assinados em 14 de março de 1985 entre o Governo Brasileiro e o BID.

As opiniões dos consultores são inteiramente pessoais e não refletem, necessariamente, o ponto de vista do IICA ou da EMBRAPA.

A coordenação dos Contratos IICA/EMBRAPA agradeceria receber comentários sobre estes relatórios.

Horacio H. Stagno
Coordenador Contratos IICA/EMBRAPA

INTER-AMERICAN INSTITUTE FOR COOPERATION ON AGRICULTURE
IICA/EMBRAPA CONTRACT

CONSULTANT FINAL REPORT

1. Consultant's full name: *Duncan Thomas*

2. Specialist in: *Agricultural Economics and Research Planning*

3. Title of IICA Project: *2.SB.3*

4. EMBRAPA Program for which consultancy is provided:

PROGRAMA: *PROCENSUL II*

SUBPROGRAMA: *09-AVALIAÇÃO EX-POST*

IICA Project Activity Code: *2.SB.3.09*	Administrative Code: *R 4884 B1B 03109*
Title of Activity of IICA Project corresponding to this consultancy	*Cooperation for the definition of an ex-post evaluation system for PROCENSUL II*

CONSULTANT CONTRACT PERIOD	DUTY LOCATION (Center)
August 9th. to 30 th., 1988	*SEP/EMBRAPA-Sede*
CONTRACT EXTENTION PERIOD (If any)	DUTY LOCATION (Center)

5. Financial support: *PROCENSUL II*

Report to DEP-EMBRAPA and IICA
Ex-post evaluation system for PROCENSUL II

Visit by Duncan Thomas to Brazil, August, 1988.

Possibly one of the most effective ways of strengthening a research program is through collaboration on joint projects. As part of this exercise, I worked with Mariza Barbosa in Brasilia, and together we visited Sonia Teixeira in Goias, Dehli Dossa in Parana and the Instituto Economia Agricola (IEA) in Sao Paulo. In addition, I briefly visited Edgar Lanzer in Florianopolis. Our aim was to identify datasets which might be used for analytical work within the context of this research program, do some preliminary data analysis with these data and work on issues of modelling farmer decision-making with these datasets. There is considerable commonality of the modelling strategies which might be adopted; I discuss these first. There are specific issues associated with each dataset and so these are discussed separately in the second section.

1. Modelling Strategies

(a) Technological innovation

Many of the questionnaires I worked with gather extensive information on the technology used by farmers. It is relatively easy to identify simple relationships among choices of technique and indicators of output, profits or wealth using cross-tabulations. It is also possible to identify the determinants of technological choice using multivariate regression analysis.

For example, one might like to know what variables affect the probability a farmer will adopt a particular technology such as applying lime to soil. This might be modelled in a limited or dummy dependent variable framework, whereby all farmers who adopt the technology are assigned a value of one and the rest zero. This variable can then be regressed on a series of farm level covariates such as socio-economic characteristics of the farmer (education, age, etc.) and farm size or assets owned by the farmer (an indicator of wealth) as well as community characteristics such as prices of outputs and inputs, credit availability (which may be tied to farmer wealth), quantity and quality of extension services available.

In this model, least squares estimates will be consistent but inefficient. A more appropriate statistical model might map the covariate space into a univariate distribution function which ensures the predictions of the dependent variable lie in the space [0,1]. In fact, we observe only whether the dependent variable is greater than some threshold value. If the distribution is assumed to be Gaussian, then we have the familiar Probit model; if it is logistic then the model is the Logit. Empirically, there is

little reason to prefer one over the other and so, often, the choice is one of convenience. Note that there are some theoretical and empirical grounds for expecting the ratios of the least squares coefficient estimates to reasonably approximate the Probit or Logit estimates. They can, therefore, be a useful starting point for analysis.

Since there are a large number of technology variables in many of the datasets, it makes sense to choose a subset of important variables and estimate a series of limited dependent variable models and compare the results.[1] It would, of course, be possible to identify a bundle of inputs as a technology and perform exactly the same analysis with farmers who do and do not use the entire bundle. This would, however, throw away some information as farmers who use some, but not all, of the inputs would be treated in exactly the same way as farmers who use none of the inputs. Assume there are two inputs. Assign a score of one for each input used. Then farmers who use the entire technology will be assigned 2, those who use it in part 1 and the rest 0. These scores have no cardinal value since they could just as easily be 20, 10 and 0. The limited dependent variable framework remains appropriate and is easily extended to the ordered probit or multinomial logit case. Of course, as the number of possible outcomes becomes large least squares becomes the appropriate model.

Finally, it is often observed that farmers adopt two different technologies simultaneously. This might be modelled by using the proportion of land devoted to modern techniques as the dependent variable. Clearly this variable is truncated at zero and unity--and we would expect there to be stacking at these end-points. Assuming Gaussian errors then a simple statistical model which could account for this stacking is the (two-limit) Tobit model.

(b) Production functions

It is straightforward to relate output to inputs through a production function the estimation of which is, by now, quite common. I would recommend adopting a flexible functional form--such as the translog model. Since some inputs and outputs are jointly determined, however, input choice is endogenous; these inputs therefore need to be instrumented in order to purge coefficient estimates of simultaneity bias. An appropriate set of instruments would be variables which affect input choices but have no direct effect on output, such as (perhaps) ownership of assets. Community level data, including prices, availability of marketing, credit, purchasing and extension services might also provide useful instruments.

---

[1] In principle, these decisions should be modelled in a simultaneous framework. If the data turn up some interesting results, this might be a useful avenue for future research.

If the notion of a technical production function is taken seriously, then clearly only farmers choosing the optimal input mix could achieve the production frontier and all other producers must lie within it. In this case, the empirical assumption of symmetric errors, which is necessary for least squares to be unbiased, cannot possibly be satisfied. An appropriate error model would permit only non-negative values -- such as the gamma distribution. In this case, the production frontier would have to be estimated by maximum likelihood methods. It may be of some interest, then, to examine the residuals and identify those farmers who are close to the production frontier (have small residuals) and those well within the frontier (large residuals).

(c) Profit functions

The production function approach is not, in general, well suited to modelling farms with multiple crops and multiple techniques. A more modern approach is based on the farm profit function. Using duality results, it can be differentiated to obtain input demand and output supply functions which are functions of all output and input prices (including the price of labor) as well as farm characteristics. These might include fixed inputs (such as land), farmer characteristics (such as age, education, household composition and wealth) and environmental or community characteristics (such as the availability of transport, credit, marketing and extension services).

Not only can these estimates provide valuable information on the likely effect of price changes on inputs and outputs but also measure the impact of other important policy variables such as extension services on farmer choices. The estimated effects of these kinds of control variables can provide valuable inputs to a wide range of policy discussions.

If the number of visits to a farmer by an extension agent turns out to have a significant impact on farmer input demands or output supplies, then it is hard to determine whether this is because the extension visits are causing the farmer to choose a different mix or because extensionists prefer to visit farmers who choose a particular mix. Similar problems arise with all service type data and with prices. Clearly, to avoid problems of simultaneity bias, it is important that these variables be collected at the community level. In some cases, such as prices, community data can be hard to collect, but they may be created by computing community averages based on the farm level data. In many cases, such as quantity and quality of extension services, the data are readily available and the use of this information should be encouraged. Indeed, as far as I am aware, there are very few good estimates in the agricultural economics literature of the impact of extension services on productivity. Several of the datasets EMBRAPA has at hand could be used to enhance this literature.

One might be tempted to include choice of technique as an explanatory variable in an input demand or output supply function.

There is no doubt that technology and inputs/outputs are related but they are also jointly determined. Clearly, to estimate unbiased coefficient estimates we need to have an instrument with which to predict technological choice. In the datasets I examined, I did not come across any likely candidates for variables which should enter a technology choice function but not the input/output functions.

(d) Demand Systems

In order to measure the effects of agricultural policy on farmer welfare -- and, more generally, the welfare of all members of society -- the analysis needs to be broadened to include consumption patterns. Assuming households maximize utility subject to a budget constraint then it is straightforward to generate a set of commodity demand functions which depend on prices of commodities, income and household (or farmer) characteristics. The case of farmers who produce for their own consumption is a little more complicated since their consumption and production decisions will be inter-dependent. In this case, input demands and output supplies will depend also on prices of consumption goods and consumption will depend on the prices of all inputs and outputs in the production process.

This is a complicated model to estimate and so it is common to assume that complete markets exist such that the production and consumption decisions can be modelled recursively. Notice, however, that the effect of production decisions will still enter consumption choices through the (income) effect of profits from the farm. Failure to include this "profit" effect in the demand system can lead to misleading estimates and inferences.

2. Datasets and Analyses discussed at EMBRAPA stations

(a) Visit to Goiania and discussions in Brasilia

Mariza and I spent two days working in Goiania with Sonia Teixeira; in addition we spent some time working in Brasilia. I also worked with Raimundo Quindere who is an excellent programmer based in Brasilia; he has the soya beans and rice (1985) survey on-line and ready for analysis on the mainframe; we both spent a lot of time working on the beans (1986) survey. The surveys are discussed separately.

(i) Soya beans and rice

Among the surveys I worked with while in Brazil, this is probably one of the most comprehensive on the technology side. About 200 farms in the Center West were covered during 1985. Information are available on land usage since 1980/1, quantity and value of outputs for the past two seasons and technology usage both currently and

with some retrospective information. There is, in addition, information on migration from the farm, sources of information and contact with extension services. Mariza has merged with these data, information on prices of outputs, extension services available and credit facilities available. It may be useful to incorporate indicators of risk into the dataset -- since risks associated with each crop are surely going to affect rates of adoption. One could, for example, compute the standard deviation of output for each municipality using these data, the Agricultural Census or some other production data source. It would also be very useful to obtain a time-series of these risk indicators.

Mariza showed me some tabulations she and Sonia have generated; the data appear to be of high quality and should provide some interesting results within the technology adoption framework. In addition to these tabulations, it may be worth examining the determinants of technology adoption. A useful starting point could be the explanation of lime and fertilizer usage along the lines proposed in section 1. In addition to the farm-level covariates (education and age of the farmer, farm size, etc.) one might include a set of municipality fixed effects. This modelling strategy would allow municipality heterogeneity to enter the adoption function completely generally and would allow the researcher to determine the extent to which failure to include municipality heterogeneity biases estimated effects. It would also be of interest to attempt to explain some of this heterogeneity by replacint the municipality fixed effects with municipality level covariates (prices, extension services, credit facilities, indicators of risk) which might affect rates of adoption.

It would also be possible to extend the analysis to include multiple response models as discussed above. Mariza and Sonia have assigned a score to each farmer based on his choice of technology; this might be used as a dependent variable in an ordered probit or multinomial logit model. The results, however, should be treated somewhat cautiously as it may be that the score assignment is not always appropriate and the dependent variable may be quite noisy.

As a first step towards building a dynamic model of adoption, one might include past information on technological choice and also attempt to explain changes in technological choice. The latter analysis may, however, be limited by the sample size.

The survey includes farmers who produce soya and rice and so it may be of interest to attempt to explain the shares devoted to each crop. Since modern technology is typically adopted with soya, and not with rice, this will probably approximate a model of share devoted to modern or traditional technology. For this model, one might adopt the two-limit Tobit discussed above.

(ii) Beans

The survey instrument for this survey is quite comprehensive and, of those I worked with, is the only one conducted by EMBRAPA which includes a consumption module. Data have been collected on land, value of inputs, quantities of outputs, labor demand (in man days), extensive information on technology, household composition and quantities of consumption. In addition, Sonia said she can easily collect data on services available at the municipality level.

Technology adoption functions -- along the lines discussed above-- can be estimated. Production functions can also be estimated with the quantity data, using asset ownership information for instruments. I discussed both of these with Sonia. Clearly, there are a number of important holes which will have to be filled before the data can be used for the kinds of analyses discussed at the end of the first section.

Firstly, prices of inputs and outputs need to be merged in with the dataset. Apparently farm gate prices of outputs are available from FGV; prices of inputs may be obtained from the Agricultural Census. Mariza is investigating this issue. If these prices are available then a duality-based input demand-output supply system may be estimated.

Secondly, to estimate a demand system, we need to know expenditures rather than just quantities of goods consumed. Some very heroic assumptions are going to have to be made in order to transform the data into a useful form.

First, we need to know prices of goods consumed. We shall have to assume that all consumers purchase the same quality at the same price. As far as I can tell, the only rural price data source for 1986 is FGV -- and thus these prices need to be converted from prices of unprocessed goods to prices of processed goods. For some goods, such as bread, this is quite easy; for others it is rather hard. Alterntaively, it may be possible to compute prices using the ENDEF survey and then assume relative prices have not changed. Finally, again using ENDEF, one might compute the rural-urban price ratio for each commodity (in 1974) and use municipality level urban prices for 1986 to create rural prices for the survey.

Second, the number of commodities included in the module is quite limited. Thus, it is necessary to make a (rather arbitrary) separability assumption so that farm-households are assumed to maximize a sub-utility function (which depends only on consumption of goods in the survey). This felicity function must be strongly separable from the utility associated with all other goods. Some of the commodities in the survey are quite aggregated (meat without bones, fish, fruit). The ENDEF data could be used to compute the representative components of these aggregates and we shall then have to assume those bundles have remained constant over time in order to use these shares to compute price indices.

The use of these data relies heavily on being able to match the data with ENDEF. It turns out that only a small number of municipalities are common to both surveys. Investigation reveals, however, that every micro-region in the beans survey is also represented by at least one municipality in ENDEF. The translation between ENDEF and the beans survey will, therefore, have to be at the micro-region level.

In principle, with these data one could estimate a complete farm-household model along the lines discussed above. It turns out, however, that there are some non-trivial problems with this survey. The survey instrument is not well designed; the data have not been coded well; data entry appears to have been a disaster and enumerators do not appear to have been supervised very closely. In my opinion, at this time, the data are not in a usable form. Although they have been subjected to some cleaning in Goiania this does not appear to have been very systematic or complete. I spent a great deal of time going through the data with Raimundo in an effort to clean up some of the many errors -- and we partially succeeded. Before the consumption data are useful, however, much work will have to be done converting all data to common units and checking for errors -- of which there are many. This is a tedious and time consuming task. It would make much more sense to invest resources in good data-entry and enumerator-checking methods (see below) than to waste programmer time fixing these kinds of errors two or three years later.

(b) Visit to Londrina

Mariza and I spent three days working with Dehli Dossa at the Soya Bean Center. Dehli has collected survey data for 500 soya bean farms during the 87/88 season. Unfortunately, he has not been able to perform any analysis of these data. In addition, he currently has another survey in the field. I shall discuss each survey separately.

(i) Analysis of 1987/88 survey data

(a) The dataset

The primary aim of this survey appears to have been the categorization of farms into groups in order to select a small number of "typical" farmers who would be tracked intensively for a longer period of time. Beyond this characterization, analysis with these data is likely to be limited.

The survey instrument contains a good amount of information on the technology used by the farmer in the production of soya beans and, where there is multi-cropping, some information on the technology used in the production of two other goods. Almost all the data, apart from land and yield are qualitative in nature; no

information on either quantities or values of inputs or value of outputs was collected. No data on income were collected although the possession of some assets is known. Unfortunately, it turns out that among the farmers surveyed there is very little variation in the ownership of these assets and so even the collected data are unlikely to provide much information about wealth. There is very little socio-economic information in the survey and in some cases the information collected is of limited value because of lack of dispersion. For example, education of the farmer was categorized into three groups: literate, primary/secondary education, tertiary education. Since the bulk of farmers have had at least some primary or secondary education and few have had any tertiary education, this variable is unlikely to be very informative. It would have been better to determine the number of years of education completed by the household head. There is also a considerable amount of attitudinal information in the survey; gathering these sorts of data is often time-consuming, the quality of the data are sometimes questionable and it is often hard to incorporate this information in the context of the behavioral models discussed in section 1. In general, I would recommend focusing survey instruments more on facts about farm activities and adopting more quantitative questions.

(b) The sample

The sampling frame was based on a series of aerial photographs which identified soya bean farms in Parana. IBGE then drew a (random?) sample of these farms and this list of 500 farms formed the sample. Of these, it turned out that about 30 farmers were duplicated in the survey because they owned more than one farm; these farmers refused to be interviewed twice. This suggests that using farms as the basic unit for listing is not optimal; it may be better to use a farmer listing from, for example, the 1985 Agricultural Census. In addition, about 100 questionnaires were, according to Dehli, not completed satisfactorily. The final useble sample is, therefore, 370 farms.

The data appear to have been very thoroughly cleaned -- perhaps too conscientiously. In many cases, farms have been dropped from the sample because they did not make sense. This is not necessarily a good idea -- especially since no record of which farmers were deleted has been maintained. It is, therefore, hard to detemine whether the sample corresponds to a random draw at this point. In fact, since Dehli apparently spent more effort correcting information form farms in more 'important' soya bean growing areas; it is quite likely that the usable sample is not representative of all soya bean farmers. Indeed, some useful information from the point of view of modelling will have been thrown away since it is often farmers in marginal areas who provide a good part of the heterogeneity necessary to estimate effects precisely.

In future, I would propose starting out with all listed farms and retaining all data collected -- including whether or not they

refused to be interviewed, were duplicated or the questionnaire was incomplete. This would permit examination of the structure of the final sample as well as a series of checks on the randomness of non-response. Observations can easily be deleted as appropriate at the analysis stage.

The survey has been converted into machine readable form both on the mini- and micro-computers in Londrina. The data are superbly documented and readily retrieved. The programmer, Serafim, has done a spectacular job and should be highly commended.

(c) The analysis

The most important problem with the analysis of this survey is the fact that there is no appropriate software available for the statistical analysis of medium size survey data. Using software I brought on my own micro-computer, I was able to work with Dehli and Serafim and do some simple exploratory data analysis, tabulations and a small amount of regression analysis. Clearly, the return to data collection will be very low if even simple tabulations are beyond the reach of the software available to researchers. I would urge EMBRAPA to ensure that there is adequate software at each station to do at least simple statistical analysis.

While this survey is not optimal for the estimation of regression models, a number of issues can be addressed such as explaining areas under production and choice of technique. We discussed these both theoretically and with examples I had produced with Dehli's data. I think, with Mariza's assistance, Dehli should be abloe to produce some useful characteristics of soya bean farmers in Parana and the determinants of their production and technology choices.

(ii) Survey in the field

A second technology survey is in the field at the moment. There is a general perception that the survey instrument could be improved and, having seen it, I am inclined to agree. The questionnaire has many open-ended questions which are hard to answer, often provide noisy information and are difficult to analyse. The questionnaire also lacks focus -- it is not clear what researchers would learn from several of the questions. I would urge data collection to focus more on objective and factual questions. It might be helpful if, a priori, researchers thought about the range of questions they would like to examine with the data, identified the essential information needed to answer these questions and constructed the survey instrument within this structure.

I have proposed that the focus of this questionnaire be shifted slightly and that a better description (in quantities and values) of all inputs and outputs be collected; more attention needs to be paid to labor inputs (both on and off the farm), socio-economic

characteristics of the household (composition, age and education) and income or wealth. These changes can be incorporated without changing the basic thrust of the questionnaire and, with the deletion of some other questions, without significantly increasing the size of the instrument. It was a real delight to work with Dehli on this instrument as he seems very willing to entertain new ideas and, indeed, seemed to be quite enthusiastic about including a consumption component in the future.

(c) Visit to Sao Paulo

Mariza and I spent one day with Luiz Irias and Jose Alberto Vicente at the Instituto Economia Agricola, (IEA), in Sao Paulo. The IEA have been collecting survey data from farms (greater than 3 Ha) in Sao Paulo state on an annual basis since 1970. Unfortunately, at this time, the data are not available on tape but they are being recoded and re-entered into machine readable form.

A random sample of farmers, stratified on farm size and region, are chosen based on a listing provide by the National Institute for Land Reform. Each farmer is surveyed five times in a year, and, typically, the same sample has been tracked for three years. The survey is currently in its sixth wave which began with 3600 farmers in September, 1986. The preceding wave was the same size and extended for five years beginning in June, 1981.

The survey instrument is very extensive indeed and includes information on all inputs (quantities), outputs (quantities and values), (family and non-family) labor usage (in some cases by task), technology used, assets owned, source of inputs, destination of outputs and credit facilities. The primary purpose of the survey, however, is to generate projections of future output -- and it is impossible to determine the extent to which the surveys were completed. If the data are of high quality, then these surveys could provide a very rich source for this project.

There are, in addition, a series of subjective questionnaires completed by agronomists providing information at the municipality level on prices, wages, land planted and expected harvests. It is claimed that input prices are available. Information on extension services in each municipality is apparently readily available from CATI and prices of output can also be obtained from an external research house. (Mariza said she would check into these.)

IEA should have 1970/1 through 1972/3 and 1983/4 through 1985/6 in machine readable form by January 1989. If the data are available at that time, then it would make good sense to determine at that time whether the instruments have been completed in sufficient detail to warrant their incorporation in this project. I

understand the data have not been used for micro-level analysis except in a small number of specific studies and so it is hard to know, at this point, how useful the surveys will be. In principle, however, they should be amenable to production function estimation, estimates based on duality and, perhaps, some study of technology. In addition, they may permit examination of specific issues such as the impact of different labor intensity or usage by task and the role of family and non-family labor.

In the event that the data are of interest to EMBRAPA, I would propose attempting to exploit the panel nature of the data which would open up another very valuable area of research. It does appear, however, that there may be some non-trivial problems associated with attrition in the samples. If the panel dimension turns out to be a fruitful resource, it may be worth encouraging IEA to focus their data-entry on the 1981-1983 period to complete that panel before coding more of the 1970 surveys.

My sense is that IEA may also be willing to entertain suggestions for additional modules (or replacement modules) in their instruments. Armed with information based on analysing their data, it may make good sense for EMBRAPA to take advantage of this opportunity.

(d) Visit to Florianopolis

I spent one day in Florianopolis working with Edgar Lanzer; it seems likely that the most useful data from the point of view of this project are located here. Edgar and I looked at data collected by the agricultural extension agency in Santa Catarina, ACARESC. The first, PAT, collected very detailed production and consumption information from farmers between 1981/2 and 1984/5. The second program, Gestoas Agricola (Gestoas) has been collecting production data since 1984/5 and is still in progress. For this survey, consumption data were collected only from about thiry farms over two years.

(i) PAT

Production and consumption data have been collected for a small number of farmers on a weekly basis since 1981/2. The production side of the survey is very extensive and includes information on all outputs, sales and consumption (in both values and quantities), inputs (by crop -- again with values and quantities), land owned and used, capital assets owned (age, purchase price), (quantity of) family labor, (quantity and value of) non-family labor. Clearly municipality level prices can be computed. Information on the quantity and quality of extension services available are readily available from ACARESC. Apparently additional community level data can be obtained from EMPASC. The production side would permit adopting the duality approach and estimating a series of

input demand and output supply functions. In addition, it would be very valuable if Edgar were to use these data for some dynamic linear program modelling.

The consumption data are also very detailed. Again, on a weekly basis, the value of auto-consumption (of both processed and unprocessed foods) and value of purchases are reported for a large number of goods. In addition, quantities are reported for food, clothing and education. The only major item missing from the consumption side is an imputed value for housing; nor is there much information about housing characteristics. A complete household roster and some time use information are included every year; in 1981/2, there is very detailed time use information (especially of women) and these data are currently being tabulated by ACARESC. A list of durables owned is reported.

Clearly these data would permit the estimation of a (subgroup) demand system and this can be linked to the production side. In fact, I think there are very few surveys of this type which rival these data for extensiveness and quality of data collection. Several hypotheses -- in particular those related to the completeness of markets -- which are maintained as true in much of the empirical farm-household literature may be amenable to testing with these data.
The PAT data have been tabulated and summarised by ACARESC for the last two years of the program and our examination of these tables suggest the data are of very high quality. It appears that farmers took the survey seriously and completed questionnaires very carefully. On the production side, this is not surprising since the aim of the survey was to provide participating farmers with feedback about their farming practices. The sample was chosen, therefore, by asking for volunteers for the program and is, therefore, certainly not random. It may be worth trying to match the characteristics of the sample with municipality means computed off the Agricultural Census to determine whether the sample is close to being representative -- which it is claimed to be by some of the ACARESC staff.

The sample has a panel component although there was a fair amount of attrition, especially in the last two years of the program. Unfortunately, it was not possible to determine how many farmers had participated in the program for two, three or four years without retrieving questionnaires from the archives. In any case, the data could be treated as a series of independent cross sections, in which case there would be just over 500 farmers. Of these, it seems likely that data from the 320 farmers covered in the last two years are of relatively high quality since the survey instrument was modified to reduce the number of open-ended questionnaires.

Two issues remain before the data could be incorporated into this project. First, Edgar said he would approach ACARESC and request

formal permission to analyse the data. Secondly, the data are currently on the original questionnaires or on summary sheets prepared by ACARESC. They clearly need to be turned into machine-readable form. The survey instrument -- even at the end of the program -- is quite unstructured with farmers having to enter names of inputs and outputs, units etc. This makes it very difficult to code and enter into a computer. The summary sheets, however, have aggregated together the data in the weekly sheets, converted all goods into a common unit and added structure to the input and output lists. Edgar and I talked about the appropriate way to get the data into machine-readable form and it would seem that as a pragmatic strategy it would be sensible to have the summary sheets entered first.

The costs will be twofold. Firstly, a loss of some detail -- such as seasonal patterns of input (especially labor) usage and, more importantly, the loss of quantity information for inputs. According to Edgar, the latter problem is easy to resolve since these prices are available at a municipality level -- and, in fact, ACARESC has already computed many of the prices using these data. This apparently was one of the controls they used to check data quality. Secondly, as I undersand it, the 81/2 data have not been transferred to summary sheets; the initial phase of the project would have to exclude these data. Ultimately, it would be a good idea to have the entire PAT dataset put onto a computer. The data entered in the first stage would then be valuable as a control to check for errors in the complicated data entry of the second stage.

If the input price data are available, then it would seem worth transferring the data onto a computer as soon as possible. Edgar would clearly need resources for this task -- ideally a computer, programmer and some funds for data entry. In the short run, it may make sense to have one of the programmers who has worked with these sorts of surveys to collaborate with Edgar in setting up a data-entry system. The aim would be to make data-entry simple and to minimize the number of entry errors. One might, for example, use a database type product (such as DBASE3), set up a front-end which replicates the forms to be typed and have a series of realtime data checks performed as the data are entered. Since the data will only be available late in this project, it is imperative that the data be as clean as possible when the analysis stage starts. Even with fairly sophisticated data-entry procedures, the data will have to be pre-screened; several of the forms have margin comments and working which should be blocked out for data entry purposes.

(ii) Gestoa

The Gestoas Agricola program has been operating since 1984/5 covering the same 150-200 farmers on an annual basis. While the production data appear to be extraordinarily good, a consumption module was included only in 1985/6 and 1986/7 and covered 37 and 22 farmers respectively. The structure of the consumption survey is very similar to the PAT consumption module; it may be sensible to

include these farms in the analysis of the PAT data.

The production data, however, may well merit examination on their own as they appear to be very comprehensive and can be matched with a substantial amount of information on quantity and quality of services available. For example, the number, experience and qualfication of extension service agents; distances to facilities; existence of co-operatives can all be added to the dataset.

A series of input demand and output supply equations could be estimated, as could technology choice functions. In addition, the panel aspect of these data could be analysed to measure the impacts of changing environments, controlling for farmer heterogeneity. Again, Edgar's analysis of these data within a dynamic linear programming framework is likely to be very valuable indeed.

ACARESC is in the process of testing a data-entry and tabulation front-end written using DBASE3. Unfortunately all the data collected to date -- and being collected this year -- continue to remain on paper. There is, then, another major data entry effort required in order to work with these data. Edgar will also request permission for EMBRAPA to analyse these data. Presumably, in return, EMBRAPA could offer ACARESC the entire dataset in machine readable form so that they can provide the farmers with complete histories of production during feed-back sessions.

The sorts of studies suggested with the PAT and Gestoa are few and far between in the literature. The analysis of these data-- which appear to be of very high quality -- could make a substantial contribution to our understanding of farm decision making. It seems the marginal product of transferring the data onto computer may be very high indeed. I would strongly urge EMBRAPA to invest resources in this exercise.

3. Conclusions

It appears to me that there may be substantial gains from trade achieved through collaborative research. It might be worth encouraging more inter-action among EMBRAPA's own economists at different stations. Survey design, for example, could be substantially enhanced if several economists discussed proposed instruments prior to going into the field. Inevitably, drawing on other researchers' and programmers' experience will, I suspect, lead to improved instruments. Coding and data entry should also be improved; of course, once a data entry module has been developed in one station it should be shared among all researchers with the appropriate modifications. This suggests sharing of programmer resources -- which may be a good idea.

There is enormous heterogeneity in the quantity and quality of

computer services available at each center. It may be a good idea to attempt to rationalise at least those services available to economists. As I understand it, for example, in Brasilia there are no micro-computers attached to the mainframe; transfer of data from PC to mainframe and vice-versa is therefore rather difficult. Yet, most stations rely on PCs. It seems imperative to me that the facility to communicate with stations (through floppy disks) be available. In Goiania, there are several PCs but none are equipped with a serial board to enable them to be used as terminals on the mainframe. In Londrina, there is no statistical software. In Florianopolis, Edgar apparently has limited access to PCs and no programming support. Empirical research is not facilitated by these sorts of problems.

6. ACTIVITIES UNDERTAKEN BY THE CONSULTANT AND RESULTS

6.1 RESEARCH DONE UNDER DIRECT RESPONSIBILITY OF THE CONSULTANT

Research activities developed	Results Achieved

See attached

6.2 SUPPORT TO RESEARCH UNDERTAKEN BY OTHER EMBRAPA RESEARCHERS

Research activities developed	Results achieved

See attached

6.3 TRAINING ACTIVITIES DEVELOPED BY THE CONSULTANT

Date	Training subject matter	Type of event*	Number of beneficiaries	
			From EMBRAPA	From other institutions
18.8.88	Modelling technological innovation in the agricultural household model.	Seminar	30	
24.8.88 - 26.8.88	Survey data analysis using micro-computers	Short course	2	

6.5 ACTIVITIES IN SUPPORT OF RESEARCH STRATEGY AND PLANNING

Research subject matter	Research program to which subject matter is concerned

Discussions of data availability and modelling strategy with emphasis on:
- Adoption of technology by farmers
- Modelling farm production decisions
- Modelling farm production decisions — a dual approach
- Integrating production and consumption decisions in farm household models

6.6 ACTIVITIES IN SUPPORT OF OTHER CENTERS AND UNIVERSITIES IMPROVING THE RESEARCH CENTERS LINKS WITH ABROAD

Subject matter on which links were recommended	Persons, centers and universities recomended for contact

7. OTHER NATIONAL SYSTEM CENTERS, APART FROM DUTY STATION CENTER, ASSISTED BY THE CONSULTANT

Research center	Area of assistance provided by the consultant
Goiania Londrina Florianopolis IEA, Sao Paulo Sede, Brasilia	--- see attached

8. CONSULTANT'S SUGGESTIONS AND TECHNICAL OR INSTITUTIONAL RECOMMENDATIONS FOR THE IMPROVEMENT OF THE RESEARCH SERVICE

See attached

9. AGREEMENTS OR COMMITMENTS ESTABLISHED WITH EMBRAPA RESEARCHERS IN-SERVICE OF THE FUTURE DEVELOPMENT OF RESEARCH IN THE CONSULTANT'S FIELD OF SPECIALIZATION

To continue work with ENDEF in consultation with Elmar de Cruz and Mariza Barbosa. Continue working with datasets discussed in report.

10. CONSULTANT'S COMMENTS ON CIRCUMSTANCES WHICH AFFECTED THE CONSULTANCY WORK

A rationalization of computer software and hardware availability would probably have substantial positive productivity effects.

Date: 9 September 1988

Signature 

Programa II. Geração e Transferência de Tecnologia

O Programa de Geração e Transferência de Tecnologia é a resposta do IICA a dois aspectos fundamentais: (i) o reconhecimento, por parte dos países e da comunidade técnico-financeira internacional, da importância da tecnologia para o desenvolvimento produtivo do setor agropecuário; (ii) a convicção generalizada de que, para aproveitar plenamente o potencial da ciência e da tecnologia, é necessário que existam infra-estruturas institucionais capazes de desenvolver as respostas tecnológicas adequadas ás condições específicas de cada país, bem como um lineamento de políticas que promova e possibilite que tais infra-estruturas sejam incorporadas aos processos produtivos.

Nesse contexto, o Programa II visa a promover e apoiar as ações dos Estados membros destinadas a aprimorar a configuração de suas políticas tecnológicas, fortalecer a organização e administração de seus sistemas de geração e transferência de tecnologia e facilitar a transferência tecnológica internacional. Desse modo será possível fazer melhor aproveitamento de todos os recursos disponíveis e uma contribuição mais eficiente e efetiva para a solução dos problemas tecnológicos da produção agropecuária, num âmbito de igualdade na distribuição dos benefícios e de conservação dos recursos naturais.

INSTITUTO INTERAMERICANO DE COOPERAÇÃO PARA A AGRICULTURA

O Instituto Interamericano de Cooperação para a Agricultura (IICA) é o organismo especializado em agricultura do Sistema Interamericano. Suas origens datam de 7 outubro de 1942, quando o Conselho Diretor da União Pan-Americana aprovou a criação do Instituto Interamericano de Ciências Agrícolas.

Fundado como uma instituição de pesquisa agronômica e de ensino, de pós-graduação para os trópicos, o IICA, respondendo às mudanças e novas necessidades do Hemisfério, converteu-se progressivamente em um organismo de cooperação técnica e fortalecimento institucional no campo da agropecuária. Essas transformações foram reconhecidas oficialmente com a ratificação, em 8 de dezembro de 1980, de uma nova convenção, que estabeleceu como fins do IICA estimular, promover e apoiar os laços de cooperação entre seus 31 Estados membros para a obtenção do desenvolvimento agrícola e do bem-estar rural.

Com um mandato amplo e flexível e com uma estrutura que permite a participação direta dos Estados membros na Junta Interamericana de Agricultura e em seu Comitê Executivo, o IICA conta com ampla presença geográfica em todos os países membros para responder a suas necessidades de cooperação técnica.

As contribuições dos Estados membros e as relações que o IICA mantém com 12 Países Observadores, e com vários organismos internacionais, lhe permitem canalizar importantes recursos humanos e financeiros em prol do desenvolvimento agrícola do Hemisfério.

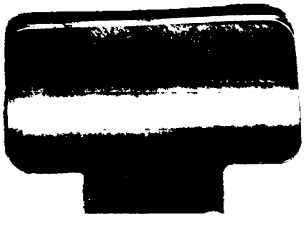
O Plano de Médio Prazo 1987-1991, documento normativo que assinala as prioridades do Instituto, enfatiza ações voltadas para a reativação do setor agropecuário como elemento central do crescimento econômico. Em vista disso, o Instituto atribui especial importância ao apoio e promoção de ações tendentes à modernização tecnológica do campo e ao fortalecimento dos processos de integração regional e sub-regional.

Para alcançar tais objetivos o IICA concentra suas atividades em cinco áreas fundamentais, a saber: Análise e Planejamento da Política Agrária; Geração e Transferência de Tecnologia; Organização e Administração para o Desenvolvimento Rural; Comercialização e Agroindústria, e Saúde Animal e Sanidade Vegetal.

Essas áreas de ação expressam, simultaneamente, as necessidades e prioridades determinadas pelos próprios Estados membros e o âmbito de trabalho em que o IICA concentra seus esforços e sua capacidade técnica, tanto sob o ponto de vista de seus recursos humanos e financeiros, como de sua relação com outros organismos internacionais.

Esta publicação foi reproduzida, em abril de 1989, numa tiragem de 80 exemplares.


FECHA DE DEVOLUCION
IICA-PM-
A4/BR-87-019
Autor
Título Ex-post evaluation system fo
PROCENSUL II
Fecha
Devolución
Nombre del solicitan



INSTITUTO INTERAMERICANO DE COOPERAÇÃO PARA A AGRICULTURA
SHIS QI 5, Conj. 9, Bl. D, Comercial Local, Caixa Postal 09-1070, Brasília, D.F., Brasil
Tel.: (061) 248-5477 - Telex 611959 INAG-BR - Correio Eletrônico 1536 - Fac-Símile: (061) 248-5807