

Thesis: Impact of Production Support on Farm Performance: a Case of Ghana's Agricultural Support Programme

Author: Ing. Sylvester Amoako Agyemang

Review by Mgr. Milan Ščasný, Ph.D.

Charles University, Environment Center

Charles University, Institute of Economic Studies Faculty of Social Sciences

Faculty of Business Administration, Faculty of Accounting and Finance, Prague

University of Warsaw, Faculty of Economic Sciences, Warsaw

Prague December 9, 2021

The thesis is prepared around three linked studies. The first one aims at farmers' decision to participate in production support program, second deals with the impact of this production support on adoption of sustainable agricultural practices by these farmers, and the last one examines whether there is an effect of the agricultural production support on farm performance. Each of these studies relies on a farmer-level data that were collected in the Northern Ghana by the student using own survey instrument.

After introducing the problem, a wide and comprehensive review of relevant literature is provided. This review documents very well author's knowledge about the literature, including the methods used and their applications in developing countries. At some places, its scope is on the cost of not going deep. For instance, Section 2.2.3 aims at various attitudinal (soft) measures (pages 12–15), while next section is devoted to psychological variables, let me correct it to *social-psychological* behavioural models and measures (pages 15–16). Naturally, these two describe the same, however, they are described in the thesis as they are not related one to the other. Moreover, these constructs, like attitudes, subjective norms, or social control, are all latent that requires specific modelling and models, however, the review is quiet about it, what is a pity, since proper measurement and modelling latent variables might improve the modelling presented in this thesis. The review – that is really wide– is divided in four subsections, each devoted to one study plus a review on production support programmes is carried out, however, as it is now, it seems to be a long list of problems and issues rather synthetic review tailored to the problem(s). One of the reasons to review a literature (additionally to learn about applied appropriate economic and/or econometric models) is to make an informed decision about the explanatory variables and other controls in the model and get a prior knowledge about expected direction in estimated associations (or effects). One would therefore expect such comprehensive review is better organised, e.g. summarising the key findings coming from the literature in a table, making useful link to the analysis that follows. Another reason to conduct a review is to identify research gap and suggest ways how to fill the gap. A question follows: based on the literature review and what you have been doing, **could you highlight clearly what is a research gap, how did you contribute to fill the gap, and what is a novelty of your research?**

Section 3 provides historical and institutional background, including The Planting for Food and Jobs (PFJ) programme—the must to be described in any empirical study. This part is clear and all important

information is provided (I had a problem to find a description of all of the six pillars, though). The most important point mentioned here is on *self-selection* to participate in PFJ program (i.e. the programme is voluntary...."with peculiar attention given to farmers who are willing to raise their productivity levels."

Section 5.1 describes the conceptual model, and I really appreciate this part, while previous section sketches the objectives and hypotheses (I think the order of these two sections shall be reverted, meaning that research objectives should follow the flows described in the conceptual model). I have several comments here:

- First, the conceptual model describes how the analysed measures and inter-linked. There are beneficiaries and non-beneficiaries (of PFJ program), it is then examined whether the beneficiaries adopted more SAPs, and, last, it is questioned whether those who participated in PFJ program also benefited in terms of performance on their farm. First issue that come into my mind is whether participation in the program and adopting the SAPs are really independent (mental) processes or, in other word, **whether there is endogeneity issue** (that is different issue than the self-selection; I mean those who adopted more SAP measures might more likely seek support and hence participate in the program). Please, comment this point.
- Second, it is intuitive and well documented in empirical literature that agents are heterogenous in their response (and behaviour). However, author's suggestion to investigate **heterogeneity wrt farm size** is falling from the sky and I was expecting to get more about this reasoning in the conceptual model.
- Third, it is suggested that author is building a structural model (for instance, at page 48 "i.e., Poisson regression"). In my opinion author is **not building nor using any structural model** that is really a pity let me add. My question is therefore to describe during student's defence **what type / kind of a structural model one might build and use** suitable for the concerned problem(s) in this thesis and the data collected (e.g., allowing to enter the latent variables, to model relationship in more complex way, and/or jointly).
- Forth, at several places author correctly argue he is not (and can't) analyse a casual effect. However, at another places his terminology is not precise, like saying "the effect of farmers' risk awareness" (at page 48), since what is doing here is analysing associations, likelihood or relationship, not (casual) effect. The terminology used might be improved when the thesis is prepared for a publication in a journal.

Section 5.2 describes the sampling and the data. First, let me congratulate to the author to get this data and I acknowledge it is not easy task to get them. I have three specific questions on the data:

- Please, describe **how the data were cleaned**, in particular, how you define the outlier and why it was needed to drop out almost 10 % of observations.
- The final dataset consists of 502 obs., **please, explain why you are using 469 and 423 obs. only in logit and Poisson model, respectively** (when descriptive statistics are provided for all 502 obs., see Table 6). I note that dropping respondents who behave systematically different (or not providing certain information that is important for the research) may bias the estimate.
- It is not clear described in the thesis, **for how many years the data on farm performance were gathered** (for one year, for two years of PFJ programme, or for more years covering also the period before PFJ program) that also have consequences on the analysis and econometric model used. Please, comment this point, in particular in relation to the third study.

Methodology (Section 5.4)

Logit and Poisson models were selected for the first two studies correctly (including testing for over-dispersion for the count data) and they are appropriate to the data and analysed research question(s).

I am only concerned that participation in PGJ program might be endogenous in the second study (on SAP's) and hence the estimate is potentially biased. **Shall be PGJ instrumented, and if it is so, what instrument(s) you would suggest to use?** I started to worry more about this issue when reading that PFJ beneficiaries are indeed different ("...had access to credit, participated in SAP training, ...", see page 75-76). **Please, comment in detail this point.**

I am a bit puzzled with the third model. When the analysis of the effect of PGJ program (the treatment) is concerned, I was expecting that one should at first match the treated (i.e. the beneficiaries) and the control (i.e. non-beneficiaries) before the treatment (using PSM, or even better *the coarsened exact matching, CEM*, that is a relatively newer casual inference technique) and then estimate the effect using panel FE model. As far as I understand the model presented in the thesis, instead of analysing the average treatment effect of treated (participation in PGJ), the study rather explores whether the intensity of PGJ support has an effect on the treated (without having a control!). I would like to ask the author **to explain these differences in modelling, and justify the particular approach used in this study. Why the effect is not analysed via comparing the outcome variables for the treated and not treated, after matching?** I agree there is a brief note along these my worries at page 59, but I did not get the presented argument, in particular why the problem (related to selection, to endogeneity, or to both) shall be solved "*...by balancing the treatment and control groups (farmers who received more production support and those who received lesser support*", page 59). **Can be justified an approach that builds a control group around a segment of beneficiaries who received lesser (not higher) support?**

Definition of variables (Section 5.5)

In the second study (the counts on SAPs), farmer's performance of eight different practices are measured. It is also argued that there is no correlation "between the adoptions of individual SAPs", page 67. I am a bit surprised about that, but still some of individual SAPs (not all) might be implemented likely together than the others. I would like to recall my comment **on structural modelling using latent variables** (to behave sustainable might be manifested by such (pre-selected) practices) and **suggest thinking about different type of modelling, relying on a measurement model for the latent constructs. Could you comment this?** Then, in some stream of research, factors rather count individual actions are used; have you thought about creating factors, using FA?

Results

I have here a couple of comments:

- **Is it correct approach to work with categorical variables, measured at 5-point Likert-scale (if 1= strongly disagree; 5=strongly agree), as they are continuous in the model?** Is it correct to interpret the result from the estimation that "This suggests that an additional point in attitudes (higher level of agreement that PFJ participation helps farmers to acquire more inputs at affordable price and to enhance productivity; scale 1 – 5) increase farmers' probability ..., page 82) or "...Thus, a one-point increase in corruption perception (i.e., high level of agreement....."? **What is appropriate modelling and interpretation for this type of (categorical) variables, if my argument is right?**

- Please, could you **comment why there is negative association between pests invasion and SAPs adoption**, while the association is positive for flood and soil erosion (true, you describe the (positive) associations at page 86, but not providing an explanation why the association for the three measures on awareness // experience differ. Btw, why the fourth measure (experiencing soil infertility) is missing in the model?
- On average, the beneficiaries received 551 GHC (see Table 8) Please comment your reason **to predict the dose-response functions up to 4000 GFC of the support** (is this range going beyond the interval what the beneficiaries received?). Why you report statistical inference of estimates (based on bootstrapping), see CI90 for SAP adoption and CI95 for farm performance?

List of publications

The list of publications included four titles. One article has published in *Eur J Devlp Res* that is ranked in Q2 with relatively high AIS (0.643). Second is published in *Agr Fin Rev* ranked 11/35 in *Agri Econ & Policy* with AIS =0.343. Third one is published in a local journal, *Central Eur Bus re* ranked 218/285 in *Business* with low AIS (0.143). The last is a book chapter published by the university.

It seems to me that the articles, except the first one, deal with topic that is different than the topic of presented thesis. Please, comment how your three articles published in the peer-reviewed journals are linked to your thesis. I encourage you, after considering the comments during the defence, to prepare manuscript and submit them in good peer-reviewed journals.

Minor comments:

- Page 56: the dependent variable in the empirical model is a binary 0/1, not the probability
- Page 57: I would not say that Poisson model is argued to be the most suitable count data estimator in the literature. For instance, in recreational demand studies, the data often suffer from over-dispersion, and hence NB outperforms the Poisson model.
- Page 74: can be $SD(PFJ)=0.48$ reported in Table 6 for the sample of beneficiaries if, I assume, every beneficiary by definition participated in the PFJ programme?
- Be consistent how you measured a measure and how you interpret it, for instance, soil erosion is measure based on answer on question “Have you been experiencing/experienced soil erosion on your cropland since 2016?” (page 70), while you conclude that “if they are aware of the existence of soil erosion...” (page 82)

In general, the thesis is well written and it provides new empirical findings on how public programs work. The author proofed his skills in empirical analysis. The thesis represents a solid piece of empirical work in the field of development and agriculture economics, based on unique data coming from own survey.

I consider the thesis to be suitable for final defence and after its successful defence the student, Ing. Sylvester Amoako Agyemang, to be awarded the title Doctor of Philosophy – PhD.