The effect of PROCAMPO on farms' technical efficiency: a stochastic frontier analysis

Saul Basurto-Hernandez
University of Birmingham UK
SXB1022@bham.ac.uk

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7

Abstract

This paper investigates the effects of the PROCAMPO subsidy on farm's technical efficiency in Mexico using the parametric stochastic frontier approach. The main findings suggest that: (i) the average technical efficiency of the 33,721 farms in the sample ranges between 43-46%, (ii) the negative effect of the PROCAMPO subsidy on farms' TE increases as technical inefficiency rises, (iii) PROCAMPO negatively affects farms' TE in 68.52-73.54% of farms in the sample (positive effects in the remaining farms), and (iv) age, years of schooling, the area-owned, use of hired labour, diversification and use of irrigation increase TE scores. These findings further support the view that agricultural subsidies negatively influence farms' technical efficiency, not only in Europe, but also in developing countries. Policy-makers should be aware about the possible effect of PROCAMPO payments on technical efficiency.

Keywords: stochastic frontier, technical efficiency, farms, subsidy

JEL: Q12, Q16, Q18

1. Introduction

Renegotiating or withdrawing from the North American Free Trade Agreement (NAFTA) would likely pose significant challenges to the agriculture sector of certain constituent countries. Furthermore, some possible outcomes from ongoing negotiations might well force policy-makers to re-evaluate the effectiveness of public policies on farms' performance thereby helping agriculturalists to adapt to changed circumstances. Since the implementation of NAFTA in 1994, the Government of Mexico has supported a direct cash transfer programme called PROCAMPO intended to shrink the difference between subsidies paid to domestic and foreign agriculturalists. It replaces the previous price-support policies, which ensured fixed prices to the farmer and has grown to become the Government programme with the largest number of recipients in the rural sector.

PROCAMPO consists of a single payment per hectare of cultivated area given to those farmers that own eligible lands.[1] The government defined the eligible land in 1994 and it is not modifiable. It comprises all agricultural fields where farmers cultivated any of the following crops between the 1992 and 1993 summer-spring agricultural seasons: cotton, rice, safflower, barley, beans, corn, sorghum, soy or wheat.[2] Since 1995, the government has removed restrictions and now agriculturalists can grow any (legal) crop. Eligibility for PROCAMPO payments has therefore become a characteristic of the land, transferable between property owner and tenant, but farmers cannot enrol new fields into the subsidisation programme. Nowadays, the main justification for PROCAMPO is to enhance productivity of Mexican farms (DOF, 1994).

[1] In 2014, small or self-consumption farmers (up to 5 and 0.2 hectares of rain-fed and irrigated lands respectively) receive $1,500 MXN/ha for the first 3 hectares of rain-fed land and $1,300 MXN/ha for the remaining fields. Medium-sized farms (5-20 and 0.2-5 hectares of rain-fed and irrigated lands respectively) receive $963 MXN/ha. Large farms (20 or more and 5 or more hectares of rain-fed and irrigated lands respectively) receive $963 MXN/ha up to 100 hectares of land (DOF, 2014). $1 USD=$13.29 MXN.

[2] Intercropping practices are also eligible but if it includes a perennial crop, the corresponding land was/is not eligible.

Over the period 1994-2015 the average payment per hectare ranged between $732-$1,615 MXN/ha (57-126 USD/ha).[3] The PROCAMPO programme covers 10.92 and 3.18 million hectares in the Spring-Summer and Autumn-Winter agricultural seasons respectively (49% and 14% of the total cultivated land in Mexico respectively). It benefits 2.41 million and 0.46 million farmers in both agricultural seasons, respectively. On average, the total budget of PROCAMPO is equivalent to 3.45% of the Gross Domestic Product (GDP) of the arable and livestock farming, forestry use, fishery and hunting sector. Moreover, it currently accounts for 16% of the Secretariat of Agriculture, Livestock and Rural Development, Fisheries and Food's (SAGARPA) budget (see Figure 6 in Appendix 1).

Previous studies have examined the effect of PROCAMPO on agriculturalists' income, migration and food security. Among others, Sadoulet et al. (2001) find that such transfers create a multiplier income-effect in the *ejidal* sector.[4] The income multiplier ranges between 1.5 and 2.6. Gonzalez-Konig and Wodon (2005) and Scott-Andretta and Cuecuecha (2010) find that PROCAMPO discourages migration from Mexico to the United States and increases the use of labour in the production of corn and beans.

Regarding food security, Garcia-Salazar et al. (2011) argue that since corn receives the largest amount of subsidy payments, this programme reduces corn imports by 40.5%. Furthermore, Ruiz-Arranz et al. (2002) find empirical evidence against the conventional wisdom that men just drink away PROCAMPO subsidies and argue that these cash transfers enhance food security through investments in domestic production. Although the existing literature has examined the effects of PROCAMPO on different areas, to the best of our knowledge there is no study investigating the effect of PROCAMPO payments on the Technical Efficiency (TE) of the agriculture sector in Mexico.

[3] Measured in constants prices 2013=100

[4] Areas of communal land used for farming activities where members individually exploit designated parcels.

Stochastic Frontier Analysis (SFA) relaxes the implausible assumption that all farms are fully efficient and allows for inefficiencies in the production process. It defines the production frontier as the maximum attainable output that can be produced using existing technology and inputs (Minvel and Latruffe, 2017). Any output-input combination lying behind the frontier indicates the existence of inefficiencies. To measure the extent of Technical Inefficiency (TI), the SFA uses two approaches. The output-oriented (OO) approach explores whether a particular farm can produce a higher level of output using the same amount of inputs. On the other hand, the input-oriented (IO) approach explores whether a farm can produce the same output using fewer inputs. Following previous studies, we use the OO approach, which is the standard method, to examine the effect of PROCAPO on farms' TE.

The primary goal of subsidisation programmes, such as PROCAMPO, is to influence farmers' income, boost productivity or to prevent beneficiaries from choosing undesired practices e.g. those that are environmentally damaging. However, subsidy payments might have a positive, neutral or potentially negative effect on farms' TE. For example, whilst some cash transfers might lead to technology modernisation, a process which could increase TE (e.g. Zhu and Oude Lansink, 2010) some recipients might use cash transfers merely to augment their income providing them with less incentive to produce efficiently (e.g. Martin and Page, 1983). Among others, Serra et al. (2008), Kumbhakar and Lien (2010), and Zhu and Oude Lansink (2010) argue that any conclusion concerning the effect of subsidies on farms' TE must be drawn from empirical evidence rather than from theorising.

To investigate the effect of PROCAMPO on farms' TE, we use representative cross-sectional data drawn from 33,721 crop farms in Mexico. In so doing this research contributes to the literature by: (i) providing empirical evidence on the link between agricultural subsidies and TE in a large developing country where there is no prior evidence concerning any such relationship and (ii) computing observation-specific and percentile-specific estimates for the subsidised-farms' TE relationship using Wang's (2002) formula and Recentered Influence Function (RIF) regressions respectively. The observation-specific and percentile-specific estimates allow us to investigate the differential impact of PROCAMPO on farms' TE.

The remainder of this paper is structured as follows. Section 2 presents an overview of the existing literature investigating the link between agricultural subsidies and farms' TE. Section 3 describes the SFA method and the database. Section 4 presents the results and discuss a set of policy implications arising out of them. Section 5 concludes with some suggestions for further research.

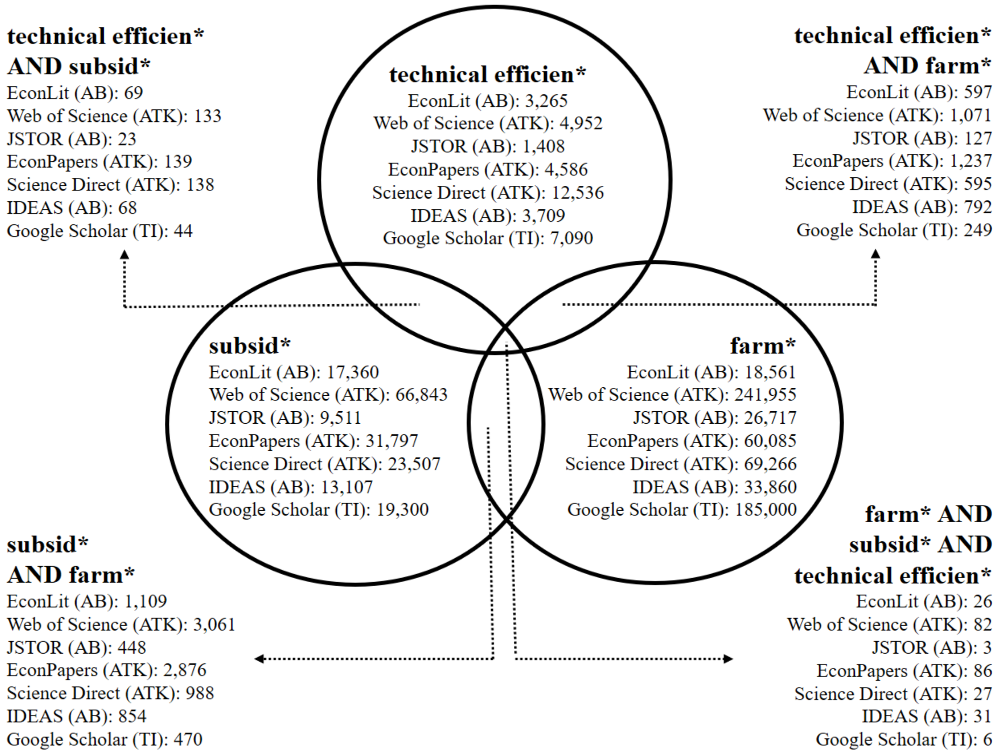
2. Literature review

For presentation purposes, we organise this section as follows. Subsection 2.1 describes the literature survey. It identifies the set of relevant materials for the literature review. Section 2.2 briefly describes two methodological approaches that account for technical inefficiencies in farming activities. This subsection discusses the main advantages of the SFA over the Data Envelope Analysis (DEA) method. Subsection 2.3 presents an overview of previous studies analysing the effect of agricultural subsidies on farms' TE. To guide the implementation of the SFA method, we compare model specifications encountered in the empirical literature, which allows us to place the novelty of this research into the existing literature.

2.1. Literature survey

A systematic literature survey helps us to identify those studies analysing the effect of subsidy payments on farms' TE. To survey the existing literature, we follow three steps. First, we select a set of search terms closely related to our topic of interest: *subsidies*, *cash transfers*, *farm*, *technical efficiency*, and *agriculture*. Second, we refine our literature search by using different combinations and word endings of these keywords in the EconLit, Web of Science, JSTOR, EconPapers, Science Direct, IDEAS, and Google Scholar databases. These materials include published papers, working papers, books, chapters and technical reports. Third, we exclude irrelevant publications by looking at their abstracts and add other relevant materials cited in the initial set of documents. Figure 1 shows the literature survey from the abovementioned databases.[5]

Figure 1. Literature survey



Source: EconLit, Web of Science, JSTOR, EconPapers, Science Direct, IDEAS, and Google Scholar databases.
Note: AB: anywhere in the abstract; TI: anywhere in the title; ATK: anywhere in the abstract, title or key words.

The initial set of materials comprises 173 different documents that result from the combination of seven outcomes at the bottom right of figure 1.[6] We exclude 95 irrelevant entries because subsidy payments are not part of the analysis and/or simply because the authors do not estimate a second stage regression for technical efficiency.[7] To complete the set of materials, we also

[5] The search tools limit us to use the same criteria in all of them therefore, we take into account those studies in which the search terms appear either in the abstract (AB), abstract, title or keywords (ATK) or title (TI).
[6] We use those studies where the three search terms, subsidies, farm and technical efficiency, appear in the abstract, title or keywords. We also use agricultur* and cash transfer* as alternative search terms for farm* and subsid* respectively but, the criteria in figure 1 outperform other alternative criteria.
[7] Some authors do not estimate a TI equation either because they assume that farms are fully efficient or because the main purpose of the research is to compute average technical efficiency scores and compare such scores between subsamples.

use two meta-analyses that examine the link between public subsidies and TE in all types of farms (Minviel and Latruffe, 2013, 2014, 2017) and organic farms (Lakner and Breustedt, 2017). As a result, we add 31 materials to the original set of references. Thus, we analyse previous findings encountered in 109 published papers, working papers, technical reports, chapters and books.

2.2. Methodological approaches

To identify the effect of public subsidies on TE, previous studies apply either the non-parametric DEA or the parametric SFA. Both methods compute strictly positive TE scores, allowing researchers to evaluate the performance of particular farms. The production frontier is the maximum attainable output produced using some inputs and existing technology. Thus, all points behind the frontier are suboptimal. The OO approach computes the size of TI by measuring the distance between the current production level and the production frontier. On the other hand, the IO approach uses the distance between current level of inputs used to produce the corresponding output and the (lower) level of inputs required to produce exactly the same output to measure the size of TIs (Kumbhakar et al. 2015).

Based on the work of Farrell (1957), Charnes et al. (1978) introduce the DEA approach and define it as a mathematical programming model that identifies economic relations such as production functions and efficient production possibilities using real data. Although there are various DEA models in the linear programming literature, Table 8 in Appendix 1 contains an overview of the Charnes et al. DEA model. The IO (OO) setting minimises (maximises) the difference between (the aggregate of) efficiency scores θ (φ) and the adjusted non-Archimedean element (ε) subject to three constraints.

The first constraint states that the sum of input x_{ij} over all farms times the corresponding parameter λ_j plus the slack variable s_i^- must be equal to the efficiency score θ times the observed input value of the corresponding farm. Second, the sum of output y_{rj} over all farms times the corresponding parameter λ_j minus the slack variable s_i^+ must be equal to the observed output value of the corresponding farm (the observed value of the corresponding farm times the efficiency score (φ)). Third, all λ_j parameters are strictly non-negative. Cooper et al. (2011) state that a farm is fully efficient if and only if $\theta^* = 1$ ($\varphi^* = 1$) and all slack variables $s_i^{-*} = s_r^{+*} = 0$. The farm is weakly efficient if and only if $\theta^* = 1$ ($\varphi^* = 1$) and $s_i^{-*} \neq 0$ and/or $s_r^{+*} \neq 0$ for some input or output. After computing the farm-specific efficiency scores, this approach examines the determinants of inefficiency in a second stage, which is a separate regression, e.g. truncated regressions or censured Tobit models.

Aigner et al. (1977) and Meeusen and Van Den Broeck (1977) introduced the parametric SFA approach into the economic literature. Among others, Kumbhakar and Lovell (2003), Coelli et al. (2005) and Greene (2008) define the SFA as a composite econometric method that accommodates technical inefficiencies and random shocks in the production of commodities (see Table 9 in Appendix 1 for an overview). These models fit a production frontier using either the Cobb-Douglas (CD), the generalised, the transcendental or the translog (TL) specifications. Using the parameter estimates of the frontier, the SFA then computes observation-specific TE scores.

The SFA approach splits the error term from the production function into a random term (v), which accounts for unobserved heterogeneity across farms and unanticipated events (white noise), and a non-negative error term (u), which accounts for TI. Therefore, a farm is fully efficient if and only if $u = 0$. The SFA approach further hypothesises the non-negative error term to be a function of variables linked to inefficiency. Thus, this method estimates a separate equation in order to identify the main determinants of TI (Kumbhakar et al, 2015). Recently, empirical studies have used a single-step maximum likelihood (ML) estimator to simultaneously obtain parameter estimates for both the frontier and the inefficiency equation since this estimator outperforms the two-step procedure (Wang and Schmidt, 2002).

Latruffe et al. (2008) and Justyna (2015) analyse the effect of agricultural subsidies on farms' TE applying both the DEA and SFA approaches. The former study reaches similar conclusions from both methods: the ratio of operational subsidies to total revenue negatively influences TE. Conversely, the latter investigation reaches opposing results; the SFA suggests that total subsidies positively affect TE but the DEA identifies harmful effects on TE of subsidy payments. Although these articles both suffer from data limitations, their findings show that the selection of method matters. Pechrova (2013) argues that changes in input levels in an inefficient farm do not alter TE scores of other farms in the DEA. However, it may modify TE scores in the SFA since this change may affect the random error term. If that happens, estimates from the DEA and SFA may differ and lead to different conclusions. Furthermore, the inclusion of new observations in the sample may shift the frontier in the DEA. In the SFA, TE scores will always be different since increasing the sample size has an inevitable effect on the random and non-negative error terms.

To summarize, the main advantages of DEA over SFA are: (i) this method does not impose any assumption on the functional form of the frontier and (ii) it is able to accommodate multiple inputs and outputs in the analysis (Bojnec and Latruffe, 2009 and Minviel and Latruffe, 2017). Regarding the SFA approach its advantages are: (i) deviations from the frontier are not only attributable to TI since it accommodates random shocks and (ii) the single-step ML method is more consistent and efficient than the two-step DEA procedure (Bojnec and Latruffe, 2009; Bakucs et al. 2010; Wang and Schmidt, 2002).

The SFA has, it seems, become the workhorse in the literature investigating the effect of subsidies on TE. According to Minviel and Latruffe (2017), 76% of studies use the SFA approach while 20% use the non-parametric DEA model.[8] Some authors argue that DEA estimates are too sensitive to outliers. Then, if we remove those outliers, very efficient or very inefficient farms, from the sample DEA's TE scores might be biased. Furthermore, DEA's TE scores are downward-biased because the exclusion of random shocks (Bakucs et al., 2010; Bojnec and Latruffe, 2009; and Mamardashvili and Schmid, 2013). In what follows the literature review confines itself to those SFA studies which account for random shocks in the frontier and include a subsidy variable in the TI equation.[9]

2.3. An overview of empirical studies

2.3.1. The frontier and ATE scores

To guide the specification of our empirical model, Table 10 in Appendix 1 summarises the set of variables used to fit the frontier function in the existing literature. The dependent variable in the frontier function is either the value of output in currency units, the quantity of output in tonnes/litres or total sales in currency units of the corresponding agricultural commodities. The former indicator is preferred over quantities because farmers tend to diversify their production efforts and it facilitates the aggregation of distinct products. The value of output in currency units is also preferred over total sales because farmers might store some portion of the produce.

Most studies include land, capital, labour and intermediate inputs as determinants of the frontier function. To account for land in the production function, previous studies use the total number of hectares, or units of land, utilised to produce the corresponding output. Some studies also include aridity indices and soil characteristics to differentiate the quality of land (Dinar et al., 2007). Accounting for capital is rather less straightforward. In theory, one should include the cost of capital utilised to produce the corresponding output. Most empirical studies use the self-reported value of fixed assets or the annual depreciation of fixed assets. Studies analysing

[8] The remaining set of materials relies on correlation analyses or on comparisons between average technical efficiency scores from different subsamples (subsidised versus not subsidised farmers) calculated with either DEA, SFA, or both.

[9] Refer to Minviel and Latruffe (2017) for a literature review that includes both DEA and SFA empirical studies.

arable activities use either the total value or annual depreciation of manufactured capital. Regarding non-arable farming, we notice that existing literature uses the total (self-reported) value of biological capital, e.g. the value of milking or breeding cows. Alternatively, biological capital is accounted for using the size of herd. The selection of the capital variable depends upon data availability, which sometimes prevents researchers from correctly account for capital endowments in the frontier function.

Previous studies use the wage bill, number of workers or working hours per annum to account for labour in the frontier. Although the former measure captures different levels of workers' skills, it does not consider unpaid family labour. The second measure may misrepresent labour since farmers hire labour sporadically. Working hours have the advantage to include and aggregate all sources of labour, including family labour, into a single variable but this measure does not distinguish between different qualities of labour. To overcome this issue, some studies introduce education indices to differentiate the quality of labour, e.g. high skilled versus low skilled workers (Dinar et al., 2007). Regarding intermediate inputs, the existing literature uses total expenses on purchased inputs, quantities of different inputs or disaggregated expenses on fertilisers, seeds, crop protection, feed, veterinary fees, energy, etc.

Some studies introduce other variables as determinants of total output such as access to agricultural extension services that may boost the productivity of inputs, altitude and indicators of policy reforms, e.g. CAP reform (Hadley, 2006; Latruffe et al., 2011). Rather than including subsidy payments in the TI model, Mc Cloud and Kumbhakar (2008) use subsidies as factors of production since Mc Cloud and Kumbhakar argue that these payments facilitate the use of inputs and, consequently, farmers obtain higher levels of output. In panel data studies, a time trend is also an argument of the frontier accounting for technological progress.

Using the set of parameter estimates of the frontier, previous studies compute average technical efficiency (ATE) scores. Figure 2 displays the distribution of ATE scores in the SFA studies that examine the effect of agricultural subsidies on farms' TE.[10] Most of the estimations analyse TE in the production of crops (39%), milk (25%) and crops and livestock (14%). Other studies estimate a SFA model for individual commodities such as beef cattle (e.g. Manevska et al., 2013), pigs (e.g. Rasmussen, 2010), rice and corn (Tian and Wan, 2000), wheat (e.g. Tleubayev et al., 2017), olives (e.g. Zhu et al., 2011), cotton (Karagiannis and Sarris, 2002), fruits (Karagiannis and Sarris, 2002) and greenhouse horticulture (Karagiannis and Sarris, 2002).

Figure 2. Average Technical Efficiency in the existing literature (SFA studies)

Source: literature review

This strand of literature encounters a range of ATE scores between 45% and 96% in regional and country-level studies.[11] The existence of such inefficiencies in the abovementioned farms confirms the appropriateness of the SFA rather than the standard production function, which assumes that farmers are fully efficient. Comparisons between ATE scores from different

[10] From the 55 studies analysing the effect of subsidies on farms' TE, 43 papers report the ATE for the corresponding samples or subsamples. Since some of the articles compute ATE scores for different sectors and countries, this graph shows the results from 88 different estimations. In panel data studies, we use the ATE of the panel of farms rather than annual ATE scores.

[11] TE scores vary within individual studies. The ATE scores represent the mean of observation-specific TE scores.

countries or regions might not be appropriate because these scores come from different frontier functions (and samples). Latruffe et al., (2017) analyse technical efficiency of dairy farms in nine European countries. Latruffe et al find that the dairy sector in Europe seems to be more efficient than other farm types.

2.3.2. Technical inefficiency

Empirical studies use farmers' characteristics, managerial practices, farms' physical features and external factors to explain TIs. The *age of the farmer* is widely used as an indicator of experience in farming activities therefore older farmers tend to be more efficient (Coelli and Battese, 1996). The literature suggests that more *years of schooling* enhances farmers' abilities to use available resources and existing technologies more efficiently. Such an effect propels farmers closer to the production frontier (Sotnikov, 1998; Dinar et al., 2007). Regarding the set of managerial practices, the share of *family labour* to total labour has a positive effect on TE if family members are better skilled than hired labour or are sufficiently involved in farming activities (Zhu and Lansink, 2010). Conversely, Karagiannis and Sarris (2005) argue that a large share of *hired labour* to total labour incentivises farmers to be more efficient. This happens because farmers look for higher revenues in order to clear higher labour costs. Moreover, farmers can discipline hired workers, which is not always possible with family labour.

The studies in this literature review encounter both a positive and a negative association between *owned land* and TE. The share of owned land to total land negatively influences TE scores since agriculturalists do not pay land rents and, consequently, do not have to look for higher revenues in order to clear such costs (Rezitis et al., 2003). On the other hand, if the farmer owns his/her fields, he/she has more incentive to invest in modern technologies and in soil improvements that reduce the waste of resources in the long run (Zhu and Lansink, 2010).

Total debt may influence TE in both directions. It may force farmers to produce closer to the frontier in order to face such liabilities or may induce farmers to make inefficient decisions due to the financial stress (Foster and Rauser, 1991).

According to the existing literature, the degree of *diversification* may augment or reduce farms' technical efficiency. Farmers tend to diversify their production efforts because they own plots of land in different locations with different soil qualities (Niroula and Thapa, 2005; Tan et al., 2006). Moreover, farmers diversify in order to cope with production risks such as plagues, water shortages or natural disasters (Latruffe et al., 2011). Another advantage of diversifying production efforts is that certain combinations of crops or alternating crops from one season to another improve soil fertility. On the other hand, agriculturalists should concentrate their efforts in the production of a particular commodity and gain enough experience to produce it more efficiently.

To identify the effect of diversification (specialisation) on technical efficiency, previous studies use any of the following variables: the share of the main output to total output, the Herfindahl index[12] or another composite index. The literature review shows that both hypotheses hold in empirical studies. For example, Manjunatha et al., (2013) and Manevska-Tasevska et al., (2013)[13] encounter a positive relationship between specialisation and technical inefficiency. Among others, Dinar et al., (2007), Bojnec and Latruffe (2009) and Karagiannis and Sarris (2005) identify a negative association between specialisation and technical inefficiency. Other authors, such as Karagiannis and Sarris (2002) and Zhu and Lansink (2010)[14] encounter mixed results.[15]

[12] Refer to the methodological section for further details.
[13] Share of revenue from the main commodity to total revenue.
[14] Share of crop revenues to total revenue.
[15] Manjunatha et al., (2013), Dinar et al., (2007), Bojnec and Latruffe (2009) and Karagiannis and Sarris (2005) and Karagiannis and Sarris (2002) use the Herfindahl index.

Among others, Rezitis et al. (2003) argue that allocation of *time to off-farm activities* at the expense of farming may lead to lower levels of TE. In contrast, Bojnec and Ferto (2011) encounter a positive effect of off-farm activities on TE. These authors attribute such effects to the availability of additional funds (off-farm income) to invest in technologies that are more efficient. *Market-oriented* farms tend to be more efficient than other farms since the interaction with other competitors enables them to acquire knowledge and relevant information. However, subsistence agriculturalists might be more efficient than market-oriented farms because of their ability to manage scarce resources (Bojnec and Latruffe, 2009).

Regarding the characteristics of farms, *irrigation* enters in the TI equation as a risk-reducing factor. It has a negative effect on TI (Karagiannis and Sarris, 2002). To control for external factors, previous studies introduce regional dummy variables, indices or dummy variables for soil types, dummy variables for LFAs, road density, distance to the next farm and dummy variables accounting for structural (policy) changes and environmental restrictions in the TI model (see Table 11 in Appendix 1 for the full set of inefficiency explanatory variables). As stated in the introduction, *subsidy payments* might have a positive, neutral or negative effect on farms' TI. In this regard, the next subsection describes such a relationship with more details.

2.3.3. Subsidies and technical inefficiency

Table 1 shows the distribution of 243 empirical findings in 55 studies examining the effect of subsidisation on farms' TE.[16] The lack of information about units of measurement of the subsidy variable prevents us comparing the size of such effects. Instead, we examine the direction of the effect. For purposes of exposition, Table 1 displays the effect of subsidisation on TE using six different classifications. Overall, most of the estimations encounter a

[16] Some studies present the SFA results for the entire sector, for different subsamples such as farm types, particular areas or countries, ranges of elevation, or quantile regressions.

significant and negative subsidy-efficiency effect (48%). Regarding the type of subsidy, Minviel and Latruffe (2017) identify subsidies that aim to increase investment or production. Production subsidies include input subsidies, output subsidies also known as coupled subsidies[17], decoupled subsidies[18], environmental subsidies and subsidies provided to farms in less favoured areas (LFAs). See Table 1 for the distribution of types and their corresponding effects on TE.[19]

To control for the intensity of the subsidisation of a single farm, SFA models in Table 1 use the total value of subsidies, the share of subsidies to total revenue, the share of a particular subsidy to total subsidies, payments per unit of land or head, a dummy variable or the share of subsidised land to total land. Although the existing literature has not reached a consensus about the standard measure of subsidisation in the SFA model, Minviel and Latruffe (2017) argue that the total value of subsidies per farm may distort parameter estimates due to size effects. Having this potential issue in mind, 55.14% estimation results in Table 1 use the total value of subsidies in the SFA model. To avoid size effects, Minviel and Latruffe suggest that one should use subsidy rates rather than total subsidy payments. In some cases, data availability does not allow researchers to use a continuous variable in the TI equation, instead, they use a dichotomous variable to indicate whether the farmer receives the subsidy or not.

Grouping together the set of parameter estimates by sector, we observe a clear pattern regarding the effect of subsidisation on different sectors. Most of the studies analyse the production of crops (42%) and milk (29%), where subsidy payments clearly reduce TE. When researchers include both arable and non-arable farming activities in the same analysis (14% of all estimations), we observe a negative effect of subsidies on TE. Recently, organic farms have

[17] Subsidies linked to the level of output.
[18] Lump-sum payments.
[19] For a further discussion about the transition from coupled to decoupled subsidies in Europe (Common Agricultural Policy) refer to Anania and Pupo D'Andrea (2015).

become popular due to changes in consumers' preferences and the promotion of environmentally friendly food production (Sauer et al., 2002). Since this type of farms must adhere to more stringent environmental regulations, they tend to be less efficient than other farms (Kumbhakar et al., 2009) and according to Table 1, subsidy payments exacerbate inefficiencies in organic farms.

Table 1. Effect of subsidies on technical efficiency from previous studies

	Share (%) of the total estimation results		
	Negative effect (significant)	Null effect (non-significant)	Positive effect (significant)
All estimations (243 results)	**47.74**	34.98	17.28
Type of subsidy			
Total subsidies (coupled and decoupled)	**35.12**	23.14	10.74
Input subsidies	0.83	0.41	**1.24**
Agri-environmental subsidies	**7.44**	4.13	3.72
LFA subsidies	0.41	**6.20**	0.41
Investment subsidies	**3.72**	0.41	1.24
Price subsidies	0.00	**0.83**	0.00
Subsidy variable			
Value of subsidies (local currency)	20.16	**24.69**	10.29
Subsidies rate (subsidies to revenue)	**17.70**	2.88	1.23
Subsidies rate (subsidies to total subsidies)	**2.47**	1.23	0.41
Subsidies rate (subsidies per land units)	**3.70**	2.47	1.23
Subsidies rate (subsidies per animal)	0.00	0.41	**1.23**
Subsidies dummy (1=the farm receives a subsidy)	**2.88**	2.47	1.23
Proportion of land (subsidised land to total land)	0.82	0.82	**1.65**
Sector			
Crops and livestock	**9.47**	3.70	1.23
Crops only	**20.58**	16.46	4.53
Livestock only	4.12	**5.76**	1.23
Organic	**2.47**	0.82	0.82
Dairy	**11.11**	8.23	9.47
Data			
Cross-section	1.65	1.23	**2.06**
Panel or pooled data	**46.09**	33.74	15.23
Place*			
Europe	**43.21**	30.45	13.17
America	2.06	**2.88**	1.23
Asia	2.47	1.65	**2.88**
Endogeneity			
Addressing endogeneity issues	8.23	**13.58**	2.47
All explanatory variables are exogenous	**39.51**	21.40	14.81

**America*: Alberta (Canada), Kansas (USA), Saskatchewan (Canada), State of Wisconsin (USA), and United States. *Europe*: Austria, Belgium, Crete (Greece), Liberecky (Czech Republic), Czech Republic, Denmark, England, Wales, Finland, France, Schleswig-Holstein (Germany), Germany, Greece, Hungary, Ireland, Italy, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, The Netherlands, and United Kingdom.*Asia*: Bihar and Eastern Uttar Pradesh (India), Wuqi, Dingbian, and Huachi (China), China, Akmola region (Kazakhstan) and Russia.

The majority of estimations in the literature review use panel or pooled databases from European countries. The main target of such studies is to evaluate the performance of farms under the Common Agricultural Policy (CAP). Using single farm payments, single area payments, agri-environmental subsidies or payments to LFAs in Europe, these empirical analyses find a negative or a non-significant association in 43% and 30% of total estimations, respectively. Apart from Europe, this literature survey identifies SFA estimations in Canada (Giannakas et al., 2001; Samarajeewa et al., 2012), USA (Serra et al., 2008; Zaeske, 2012), India (Dung et al., 2011) and China (Tian and Wan, 2000; Zhao et al., 2015; Ito, 2015). By simply looking at figures in Table 1, the relationship between subsidies and TE in zones other than Europe seems to remain ambiguous. Moreover, other regions than Europe are not widely covered by this literature, especially developing countries in America, Africa and Asia (Minviel and Latruffe, 2017).

The SFA model suffers from endogeneity issues if there exists a correlation between inputs and the random error term in the frontier, a correlation between inefficiency effects and the random error term or a correlation between the use of inputs and technical inefficiencies.[20] Latruffe et al. (2017) address the first source of endogeneity by using a 4-step estimation procedure.[21] This issue arises if agriculturalists adjust intermediate inputs (e.g. fertilisers, irrigation or pesticides) as a response to stochastic events (e.g. weather shocks or plagues). If the SFA model does not account for such events, these are part of the error term. Quiroga et al. (2017) argue that coupled subsidies are endogenous since these payments depend on the level of output and farmers can influence this type of subsidisation. Quiroga et al. (2017) overcome

[20] This happens if less efficient farms use large quantities of inputs, which suggests a positive correlation. To the best of our knowledge, this source of endogeneity has not been addressed in the existing literature yet.

[21] First, it regresses the endogenous input on the exogenous variable vector using OLS. Second, it computes the non-linear least squares estimator to obtain the full set of parameters in the SF and use the estimated coefficients in the technical inefficiency equation to compute the instrument. Third, it computes the non-linear two-stage least squares to obtain the full set of parameters in the SF using the estimated instrument (step 2) and use the new estimates of the technical inefficiency equation to compute a better instrument compared with the one in step 2. Fourth, it replicates the previous step and uses the estimated instrument in step 3.

this issue by using a two-stage estimation.[22] Although Latruffe et al and Quiroga et al address the abovementioned sources of endogeneity using four-step and two-step sequential estimations, such results are always less efficient than the single-step estimation results. Unfortunately, the complexity of the model prevented Latruffe et al to fit a single-step model.

3. Methods and materials

3.1. Theory

The economic literature defines the production function as the process of transforming inputs into output(s) and its mathematical representation is as follows:

$$F(\boldsymbol{x}, \boldsymbol{y}) = 0 \tag{1}$$

Where $\boldsymbol{x}$ and $\boldsymbol{y}$ are J and M dimensional non-negative vectors of inputs and outputs respectively. For a single output, and for simplicity, we can rewrite expression (1) as:

$$y = f(\boldsymbol{x}) = f\left(x_1, \dots, x_J\right) \tag{2}$$

Where $f(.)$ is the maximum attainable output for a given set of inputs. Chambers (1988, p. 9) states that a well-defined function should satisfy certain regularity conditions. First, the production function is finite, non-negative, real-valued and single-valued for all non-negative and finite inputs. Second, the absence of inputs leads to no output. Third, additional inputs will never produce less output (monotonicity). Fourth, the production function is continuous and twice differentiable at any point. Fifth, the input set is convex and therefore, the production function is quasi-concave.

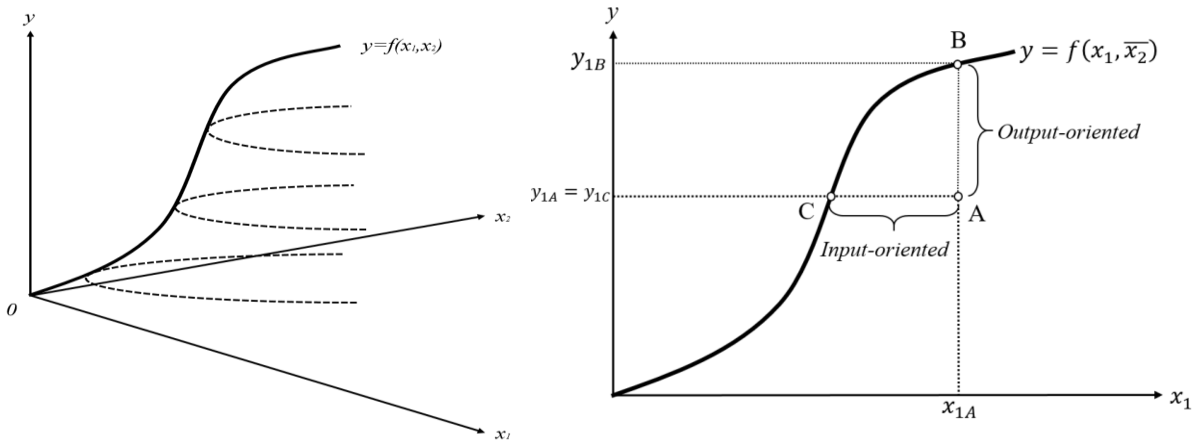
Figure 3a displays the feasible production set using two inputs, x_1 and x_2. The surface denoted by $y = f(x_1, x_2)$ is the maximum achievable output and it bounds the feasible production set.

[22] First, the ratio of coupled subsidies to total crop production is regressed on a vector of farm characteristics and indicators of policy reforms using the Fixed Effects (FE) or the Random Effects (RE) estimators. Second, a SFA model is estimated in which predicted values of coupled subsidies enter as an additional input in the production function. Actual coupled subsidies and particular decoupled payments are part of the technical inefficiency equation.

Fixing any of the two inputs to a certain level (e.g. $\overline{x_2}$), we obtain the total product curve for the remaining input, which captures the relationship between the corresponding input and the total output (e.g. $y = f(x_1)$). Thus, holding other inputs constant, the slope of the total product curve $\partial y/\partial x_1$, is the marginal product of x_1. To be consistent with theoretical underpinnings, $\partial y/\partial x_1 \geq 0$ and $\partial^2 y/\partial x_1^2 < 0$.[23] Both conditions together guarantee that increasing any input has a non-negative effect on total output.

Standard production theory assumes that farms operate along the frontier. Thus, only random noise prevents farms remaining on the production frontier. Nevertheless, the production efficiency literature relaxes this restriction. It allows farmers to operate on or below the frontier due to technical inefficiencies. Figure 3b displays the total product of x_1 and illustrates two different measures of technical inefficiency: Input-Oriented (IO) and Output-Oriented (OO) approaches. Point A is below the frontier and results from using x_{1A} units of input x_1. This research uses the OO approach.

Figure 3. Single output production function



3a. Single output production function (two inputs)

3b. Total product curve (x_1)

Source: adapted from Kumbhakar et al. (2015)

[23] Law of diminishing returns.

The OO approach indicates that a higher level of output y_{1B} is achievable using the same level of input x_1. Therefore, A is an inefficient production level and the size of technical inefficiency (TI) is equal to $(y_{1B} - y_{1A})/y_{1B}$, and consequently, technical efficiency (TE) is y_{1A}/y_{1B}. Accounting for such inefficiencies, we can rewrite the production function in (2) as follows:

$$y = f(\boldsymbol{x}) * \exp(-u) \tag{3}$$

Where u is non-negative and measures TI. For small values of u, $\exp(-u) \cong 1 - u$. Thus, TE $= \exp(-u) = 1 - u = 1 - TI$.

Using equation (3), we can also measure the effect of increasing inputs on the level of output, or returns to scale (RTS). According to the widely known economic literature, a production function is homogeneous if the following condition holds (Kumbhakar et al. 2015):

$$\lambda^{\gamma} y = f\left(\lambda x_1, \dots, \lambda x_J\right) \tag{4}$$

Here, condition (4) implies that if all inputs rise by the same proportion, λ, total output increases by λ^{γ}. Then, this production function is homogeneous of degree γ. For $\gamma > 1$, $\gamma < 1$, or $\gamma = 1$, we observe increasing, decreasing or constant RTS respectively. For homogeneous functions, RTS is the sum of input elasticities:

$$RTS = \sum_{j=1}^{J} \varepsilon_j(\boldsymbol{x}), \text{ where } \varepsilon_j(\boldsymbol{x}) = \frac{\partial \ln f(\cdot)}{\partial \ln x_j}. \tag{5}$$

Accounting for TI in the SFA does not alter formulation (5) since this term appears additively after taking logs of equation (3).

3.2. Method

The estimation of the SFA model includes parameter estimates of the frontier and the technical inefficiency functions. To obtain such parameters, we apply the single-step maximum likelihood (ML) method.[24] The SFA literature typically uses the CD and/or the TL production functions to identify the frontier in equation 3. The CD function with TIs and in its logarithm form is as follows:

$$y_i = \beta_0 + \sum_{j=1}^{J} \beta_j x_{ij} + v_i - u_i \tag{6}$$

Where, v_i is the random noise in the frontier, u_i is the TI term, and $\beta_0 = \ln a$. For strict concavity (quasi-concavity), it requires $0 < \beta_j < 1 \ \forall\ j = 1, \dots, J$, $0 < \sum_{j=1}^{J} \beta_j < 1$ and $a > 0$ ($\beta_j > 0 \ \forall\ j = 1, \dots, J$ and $a > 0$). The CD function is homogeneous of degree $\sum_{j=1}^{J} \beta_j$ and the corresponding elasticities of total output with respect to individual inputs are equal to $\varepsilon_j = \frac{\partial \ln y}{\partial \ln x_j} = \beta_j$. Therefore, $RTS = \sum_{j=1}^{J} \varepsilon_j = \sum_{j=1}^{J} \beta_j$. On the other hand, the TL production function is as follows:

$$y_i = \beta_0 + \sum_{j=1}^{J} \beta_j x_{ij} + \frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{jk}\, x_j x_k + v_i - u_i \tag{7}$$

Where $\beta_{jk} = \beta_{kj}$. Unlike the CD function, this specification is not necessarily homogeneous, unless $\sum_{k=1}^{K} \beta_{jk} = 0 \ \ \forall\ j$, and does not assume a constant elasticity of substitution (equals unity in the case of the CD function). The change of total output given by a change in any of the inputs depends on the use of other inputs since $\varepsilon_j = \frac{\partial \ln y}{\partial \ln x_j} = \beta_j + \frac{1}{2} \sum_{k=1}^{K} \beta_{jk}\, x_k$. Hence, RTS are equal to $\sum_{j=1}^{J} \varepsilon_j = \sum_{j=1}^{J} \left(\beta_j + \frac{1}{2} \sum_{k=1}^{K} \beta_{jk}\, x_k\right)$.

Notice that the CD function is a special case of the TL specification. The latter reduces to the former if $\beta_{jk} = 0 \ \ \forall\ jk$. To empirically test for the appropriateness of the two functional forms, the existing literature uses a likelihood ratio test of the form: $LR = -2(L_{CD} - L_{TL})$, where the CD model is nested in the TL model. Here, L_{CD} and L_{TL} are the log-likelihood values of the

[24] Refer to Wang and Schmidt (2002) for a further discussion about the superiority of the single-step estimator over the traditional two-step estimation procedure.

CD and TL models respectively. The LR-statistic follows a χ^2 distribution with $df_{TL} - df_{CD}$ degrees of freedom, that is, the difference between the degrees of freedom of the corresponding models (Greene, 2012, p. 526-527).

To identify the two elements of the composite error term in the SFA model, Aigner et al. (1977) and Meeusen and van den Broeck (1977) impose parametric distributions on both terms. The SFA assumes that v_i is an i.i.d. random term with zero mean and constant variance $\left(N(0, \sigma_v^2)\right)$. It accounts for unobserved heterogeneity across farms, stochastic events involved in production activities and errors in the functional form of the frontier. Moreover, it assumes independency between v_i and u_i.[25] Regarding the non-negative error term, empirical studies adopt either a half-normal ($u_i \sim i.i.d.N^+(0, \sigma_u^2)$), a truncated-normal ($u_i \sim i.i.d.N^+(\mu, \sigma_u^2)$), or an exponential ($f(u_i) = \frac{1}{\eta} * \exp\left(-\frac{u_i}{\eta}\right)$) distribution for the TI term.

The one-parameter half-normal and exponential distributions cluster the majority of observations close to full-efficiency, that is, the mode of the distribution of TI is zero. This seems to be a restrictive assumption as one may observe high inefficiencies in farming activities. The truncated-normal distribution relaxes such restriction by allowing the mode of u_i to be nonzero. Furthermore, if $\mu = 0$ the truncated-normal is identical to the half-normal distribution. For these reason, we assume a (more flexible) truncated-normal distribution of TI in the SFA model.

To empirically disentangle the composite error and compute the estimator of u_i, the SFA model uses the Jondrow (JLMS) formula (Jondrow et al. 1982, p. 235) along with the previous assumptions:

[25] As pointed out in Kumbhakar et al. (2015, p. 55), since v_i captures exogenous shocks, it is unlikely that it might be related to production inefficiencies. However, this random term may capture risks in the production process and farmers' risk-attitudes may be captured by the inefficiency term. See Smith (2008) for further details about such dependency.

$$\hat{E}[u|\delta] = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi\left(\frac{\delta\lambda}{\sigma}\right)}{1-\Phi\left(\frac{\delta\lambda}{\sigma}\right)} - \left(\frac{\delta\lambda}{\sigma}\right)\right] \tag{8}$$

Where, $\sigma = \sqrt{\sigma_v^2 + \sigma_u^2}$, $\lambda = \frac{\sigma_u}{\sigma_v}$, ϕ and Φ stand for the standard normal and cumulative density functions, respectively, and $\delta = v - u$. In the JLMS formula, $e_i = y_i - \boldsymbol{\hat{\beta}}'\boldsymbol{x_i}$ is the estimator of δ_i. This estimator allows us to obtain observation-specific TE and TI scores.

Among others, Kumbhakar, Ghosh, and McGuckin (1991), Reifschneider and Stevenson (1991), Huang and Liu (1994), and Battese and Coelli (1995) argue that the assumption of a truncated-normal distribution ($\mu \neq 0$) enables us to parameterise the expected value of the non-negative error term. Thus, μ is a linear function of a vector of exogenous variables, $\boldsymbol{z}$, that determines technical inefficiencies:

$$\mu_i = f(\boldsymbol{z_i}) = \gamma_0 + \sum_{m=1}^{M} \gamma_m z_{mi} \tag{9}$$

Where, γ_m are the corresponding parameters. Since both moments of u_i are observation-specific, we can account for heteroscedasticity in the TI term and parameterise σ_u^2 as in Caudill and Ford (1993), Caudill, Ford, and Gropper (1995) and Hadri (1999).

Following Caudill and Ford, Caudill, Ford, and Gropper and Hadri, Wang (2002) proposes a model in which both μ_i and σ_u^2 are linear functions of the same vector of exogenous variables $\boldsymbol{z}$. Wang adds the production uncertainty (σ_i^2) equation to the traditional model in Kumbhakar, Ghosh and McGuckin (1991), Huang and Liu (1994) and Battese and Coelli (1995)[26], which is as follows:

$$\sigma_i^2 = f(\boldsymbol{z_i}) = \exp(\vartheta_0 + \sum_{m=1}^{M} \vartheta_m z_{mi}) \tag{10}$$

This model relaxes the assumption that TI increases (decreases) monotonically with the corresponding inefficiency effect. In such a case, the relationship between a variable in the

[26] Equations (6) or (7) and the parameterisation of μ_i in equation (9)

inefficiency equation and technical inefficiency may alternate signs within the sample. This extension of the SFA model adds more complexity to the analysis because the single-step method estimates the parameters of the frontier, technical efficiency (μ_i) and production uncertainty (σ_u^2) equations simultaneously.

Using Wang's (2002) model, the marginal effect of the m-th inefficiency effect on μ_i is as follows[27]:

$$\frac{\partial \mu_i}{\partial z_m} = \gamma_m \left[1 - \Lambda_i \left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)}\right] - \left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)}\right]^2\right] + \vartheta_m \frac{\sigma_i}{2}\left[(1 + \Lambda_i^2)\left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)}\right] + \Lambda_i \left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)}\right]^2\right] \quad (11)$$

Where $\Lambda_i = \mu_i/\sigma_{u,i}$ and ϑ_m is the m-th corresponding coefficient in the production uncertainty equation. These marginal effects are observation-specific and their signs reveal the direction of the effect of elements in $\boldsymbol{z}$ on TI.

Linking observation-specific TE scores and marginal effects allows us to see if there are differential effects of PROCAMPO on TE in the sample. Furthermore, we use RIF-regressions to examine differential effects of PROCAMPO on farms' TE. This method is similar to a standard OLS regression but the dependent variable ($u = -\ln(TE)$) is replaced by its RIF function. For a quantile analysis, Firpo et al. (2009) defines the RIF function as follows:

$$RIF(u; q_\tau) = q_\tau + \frac{\tau - \mathbb{1}(\mathbf{u} \leq q_\tau)}{f_u(q_\tau)} \quad (12)$$

Where q_τ is the τ-th quantile of the unconditional distribution of u (maximum value of u in the corresponding quantile), τ is the quantile, $f_u(q_\tau)$ is the probability density function of u evaluated at q_τ, and $\mathbb{1}(\mathbf{u} \leq q_\tau)$ is an indicator variable equals one if the outcome value (u) is less than or equal to q_τ and equal to zero otherwise.[28]

[27] If σ_u^2 is not parameterised, the second term in the right-hand side of Wang's (2002) formula vanishes in equation (11).
[28] For a further discussion about the advantages of unconditional over conditional quantile regressions approaches refer to Borah and Basu (2013). Overall, the conditional quantile regression (alternative method) identifies the effect of an explanatory variable on a specific quantile of the outcome (dependent) variable. To do that, the conditional quantile regression assess such an effect using specific values of the remaining covariates. Borah and Basu (2013) argue and provide empirical evidence that

3.3. Data description

To estimate the SFA model, we use information from the 2014 National Agricultural Survey (NAS). The National Institute of Statistics and Geography in Mexico (INEGI) releases the NAS. It contains data on 66,483 farms and is a representative sample of the 34 major agricultural commodities. The lack of information on perennial crops and livestock, e.g. value of milking cows, age of perennial trees/plants or age of breeding pigs/cows, prevents us to compute the cost of capital for these farms. Moreover, the production cycle of some perennial crops and livestock activities typically last more than one year, e.g. the production of avocado or fattening cattle, and then some farms may not report the corresponding annual value of output in the NAS.[29] Additionally, PROCAMPO does not cover perennial crops and livestock activities. Therefore, the dependent variable in the production frontier is the value of output(s) from arable activities that last at most one agricultural year. To aggregate all agricultural commodities produced within the corresponding farm into a single category, we use self-reported farm gate prices.[30] Since we use the value of the produce and not the marketed output, the dependent variable does not suffer from the storage effect.

In line with the literature review in section 2, the SFA model includes measures of capital, land, labour and intermediate inputs purchased from outside the farm in the OO frontier equation. The total ownership cost of capital per farm controls for different capital endowments. The NAS contains detailed information on all types of machinery and equipment, which allows us to compute the corresponding ownership costs. According to Edwards (2015), the ownership cost is equal to capital recovery cost (total depreciation times the corresponding capital

the results from the conditional quantile regression are not always interpretable in a policy or population context. These quantile effects are usually valid for the corresponding quantile and not for the whole sample or population. Conversely, unconditional quantile regressions such as the RIF-function provide generalizable results since this method computes quantile-specific marginal effects using the (entire) distributions of other covariates in the model.

[29] For instance, a farmer cultivating avocados may report zero output in 2014 because avocado trees are not in the productive phase.

[30] If the farmer does not report farm gate prices, we use averages of farm gate prices at municipality-level instead.

recovery factor plus the salvage value times the real interest rate) plus taxes, insurance and housing costs (0.05 times the purchase price of the corresponding equipment plus the salvage value).[31] Some farms substitute tractors with a yoke of oxen. In this regard, the NAS only collects data on whether the farmer uses oxen or not. Since 20% of farms in the sample use this type of capital, we account for this using a dummy variable in the frontier equation. Unfortunately, data on buildings and facilities is not available in the NAS. Total land utilised to produce the composite output is also included in the SFA equation. To control for the amount of labour, we use the total number of working hours spent on farming activities per farm in the 2014 agricultural year. It includes working hours from full-time workers, temporary workers, '*jornaleros*'[32] and family members. Regarding intermediate inputs, we aggregate together all annual expenses into a single indicator (see Appendix 1 for further details).

To explain technical inefficiencies in farming activities, we use the standard set of explanatory variables in the literature. Characteristics of the farmer include their age and years of schooling. Regarding managerial practices, we use the share of owned land to total land, the share of hired labour to total labour, the Herfindahl (diversification) index[33], the share of irrigated area to total area and a dummy variable for farms selling agricultural commodities abroad directly, especially in the US market. According to section 2.3.2, total debt and off-farm income may determine the size of TIs. However, the NAS does not collect information on off-farm activities, remittances or total debt. In this regard, we acknowledge that not all forms of financial capital are accounted for in the model since there is a delay between incurring expenditures and receiving revenue from harvest even in the production of annual crops. To investigate the effect

[31] The capital recovery factors and salvage values are available on https://www.extension.iastate.edu/agdm/crops/html/a3-29.html. For purchase prices of equipment and machinery refer to http://www.sagarpa.gob.mx/agricultura/Precios/Paginas/PreciosdeMaquinariaAgricola.aspx

[32] Employees hired for sporadic activities (per day) such harvesting activities.

[33] The Herfindahl index is equal to $HI_i = \sum_{j=1}^{J} \left(A_{ji}/TA_i\right)^2$, where A_{ji} is the total area allocated to crop j and TA_i total land in farm i.

of PROCAMPO on farms' TI, we use a dummy variable. The NAS comprises self-reported data on whether the corresponding farm receives PROCAMPO or not.

After removing entries with missing data, impossible values and farms with perennial crops or livestock activities, the database contains 33,721 valid observations. We exclude 5,070 farms that do not report socio-demographic characteristics of the farmer, age and years of schooling. Also excluded from the database are the 5,371 farms that do not report or report impossible information on output(s), working hours, intermediate inputs and proportions of irrigated areas (e.g. none working hours in the agricultural year, zero expenses on intermediate inputs or proportions of irrigated land to total land greater than 100%). We also remove 22,321 farms from the sample, which allocate their production efforts to perennial crops or livestock activities. Thus, the final sample comprises data on 33,721 farms that derive 100% of their revenue from annual (or seasonal) crops. Table 2 displays definitions, summary statistics and expected signs of the corresponding variable in the frontier and technical inefficiency models.

Table 2 shows that 46% of farms in the sample received the subsidy payment in the 2014 agricultural year. In average, we observe that beneficiaries of PROCAMPO obtain higher revenues but also, utilise larger amounts of capital, land and intermediate inputs that non-recipients. Conversely, non-beneficiaries tend to use labour more intensively than farmers receiving PROCAMPO.

Agriculturalists are on average 57 years old and have 6 years of academic studies, which is equivalent to a primary school education. Overall, most of the farmers are the owners of the sampled fields (82% of the total farmland). The production of agricultural commodities mainly relies on family labour (76% of total working hours). There exists a farmers' specialisation towards particular crops such as corn, beans, sorghum and wheat (Herfindahl index equals 0.91).

Table 2. Definitions, descriptive statistics and expected signs

Variable	Description	Units	Mean		SD		Min.		Max.		Sign
			Stochastic Frontier equation								
			P	WP	P	WP	P	WP	P	WP	
Output	Value of the produce (all agricultural commodities)	$*1,000	449.47	372.28	1,335.31	1,435.63	0.15	0.15	23,900.00	23,800.00	NA
Capital	Total ownership cost of capital	$*1,000	92.16	40.35	212.19	144.02	0.00	0.00	3,867.87	2,799.85	+
Land	Total utilised agricultural land to produce	has	36.43	23.92	117.47	126.68	0.02	0.01	6,055.41	8,221.00	+
Labour	Total labour including full-time, temporary, 'jornaleros' and family workers	hrs*1,000	5.54	6.34	12.74	23.26	0.05	0.05	575.33	1,288.28	+
Inputs	Annual expenses on intermediate inputs	$*1,000	283.36	201.32	1,616.29	3,239.09	0.00	0.00	174,000.00	427,000.00	+
Oxen	Farms using a yoke of oxen	0,1	0.19	0.21	-	-	0.00	0.00	1.00	1.00	
			Technical inefficiency equation								
Age	Age of the farmer	years	58.81	55.24	13.74	14.47	16.00	16.00	100.00	100.00	±
Schooling	Farmer's education	years	6.01	6.22	4.83	4.84	0.00	0.00	24.00	26.00	-
Owned	Ratio of owned land to total agricultural land	%	83.48	80.98	32.87	37.19	0.00	0.00	100.00	100.00	-
Hired	Ratio of hired labour to total labour	%	24.58	23.41	32.29	32.98	0.00	0.00	100.00	100.00	-
Herfindahl	Herfindahl index (the closer to one, the closer to full specialisation)	0-1	0.89	0.93	0.19	0.16	0.17	0.21	1.00	1.00	±
Irrigated	Ratio of irrigated land to total agricultural land	%	32.23	27.99	43.67	42.50	0.00	0.00	100.00	100.00	-
Abroad	Farm directly selling some of the produce abroad (1=yes and 0=no)	0,1	0.004	0.004	0.06	0.06	0.00	0.00	1.00	1.00	-
Procampo	Farm receives PROCAMPO (1=yes and 0=no)	0,1	0.46		-		0.00		1.00		±

P: the farmer receives PROCAMPO (15,420 farms), NP: the farmer does not receive PROCAMPO (18,301 farms)
Source: National Agricultural Survey (2014) and SAGARPA (2014)

Almost one third of the sampled fields have an irrigation system (30% of total farmland). In addition, few farms sell their produce abroad directly (0.4% of farms). Although Mexico exports large quantities of agricultural commodities, farmers usually sell their output(s) to intermediaries (42% of the total number of arable-farms in Mexico), food processors (9%) and other buyers, who finally sell these products abroad, especially in the US market.[34] The following section examines the effect such variables on farms' TE.

4. Results

To present the set of findings, we organise this section as follows. First, we analyse the parameter estimates of the frontier model, the ATE scores and the corresponding input elasticities. Second, we discuss the implications of parameter estimates of the TI equation. We perform a set of RIF-regressions to compute marginal effects of the subsidy variable on TE for each percentile of the distribution and show the distribution of the observation-specific marginal effects of PROCAMPO on farms' TE scores. Third, we develop a set of robustness checks to verify the consistency of our findings.

4.1. Production frontier

The SFA model uses equations (6-7), (9) and (10) to examine the PROCAMPO-TE link. Table 3 shows the parameter estimates of the OO production frontier. To identify the functional form that better fits out data, we estimate both the CD and the TL models with and without technical inefficiencies. To minimise biases resulting from omitted variables, we use regional fixed effects to control for heterogeneous climate, economic policies, traditions and other regional factors in the frontier models[35]. SFA models (1) and (4) in Table 3 assume that there are no

[34] http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/agropecuarias/ena/ena2014/doc/tabulados.html

[35] Using the SAGARPA's regionalisation, we create four dummy variables. Region 1 (Centre): Ciudad de Mexico, Hidalgo, Estado de Mexico, Morelos, Puebla, y Tlaxcala. Region 2 (Centre-West): Aguascalientes, Colima, Guanajuato, Jalisco, Michoacan de Ocampo, Nayarit, Queretaro, San Luis Potosi, y Zacatecas. Region 3 (North): Baja California, Baja California Sur, Chihuahua, Coahuila de Zaragoza, Durango, Nuevo Leon, Sinaloa, Sonora y Tamaulipas. Region 4 (South-East): Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, y Yucatan. The inclusion of state fixed effects adds more complexity to the model and sometimes prevent it to converge.

technical inefficiencies in production activities, that is, all parameters in the mean and variance TI equations are equal zero. So that, all deviations from the frontier arise from random shocks. SFA models (2) and (5) parameterise the mean TI equation but assume that all parameters in the variance TI equation are zero. Allowing for non-monotonic effects, SFA models (3) and (6) parameterise both moments of the TI error term as in Wang (2002).

Before analysing the set of findings, we test for the existence of technical inefficiencies in the agriculture sector (hypothesis 1) and for the appropriateness of the Wang's model (hypothesis 2). Hypothesis 1 states that farms are fully efficient, and consequently, parameters in the (mean and variance) TI equation are simultaneously zero. Regarding hypothesis 2, it states that the parameterisation of the variance of TI is not appropriate, then, all coefficients are zero. Table 12 in the Appendix indicates that all slopes in the (mean) TI equations are different from zero, that is, we can reject hypothesis 1. This result holds for both the TL and CD functions. Moreover, table 12 shows evidence in favour of the appropriateness of Wang's model rather than the standard approach. The likelihood ratio test indicates that we can reject hypothesis 2. This test indicates that the parameterisation of production uncertainty matters. Therefore, SFA models (3) and (6) should be preferred over models (1-2) and (4-5).

Table 3. Parameter estimates of the production frontier

Variables	***Frontier***					
		Cobb-Douglas			***Translog***	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Capital	0.0245***	0.0178***	0.0172***	0.1417***	0.1164***	0.1204***
	(0.0017)	(0.0016)	(0.0016)	(0.0179)	(0.0169)	(0.0169)
Capital square				0.0002	0.0015*	0.0012
				(0.0009)	(0.0008)	(0.0008)
Land	0.3523***	0.4136***	0.4156***	0.4129***	0.4229***	0.4253***
	(0.0077)	(0.0077)	(0.0077)	(0.0633)	(0.0606)	(0.0609)
Land square				-0.0935***	-0.0941***	-0.0936***
				(0.0041)	(0.0041)	(0.0041)
Labour	0.1018***	0.0939***	0.0918***	0.1104*	0.3108***	0.3008***
	(0.0074)	(0.0071)	(0.0072)	(0.0630)	(0.0597)	(0.0606)
Labour square				0.0255***	0.0120***	0.0128***
				(0.0040)	(0.0038)	(0.0039)
Inputs	0.6569***	0.5536***	0.5465***	-0.0424	-0.0018	-0.0043
	(0.0052)	(0.0057)	(0.0058)	(0.0482)	(0.0463)	(0.0466)
Inputs square				0.0657***	0.0552***	0.0549***
				(0.0020)	(0.0020)	(0.0021)

*Capital*Land*				0.0242***	0.0225***	0.0228***
				(0.0016)	(0.0016)	(0.0016)
*Capital*Labour*				0.0073***	0.0061***	0.0060***
				(0.0018)	(0.0016)	(0.0016)
*Capital*Inputs*				-0.0223***	-0.0203***	-0.0205***
				(0.0013)	(0.0013)	(0.0013)
*Land*Labour*				0.0753***	0.0665***	0.0661***
				(0.0075)	(0.0071)	(0.0072)
*Land*Inputs*				-0.0284***	-0.0151***	-0.0151***
				(0.0046)	(0.0046)	(0.0046)
*Labour*Inputs*				-0.0610***	-0.0570***	-0.0569***
				(0.0054)	(0.0052)	(0.0052)
Oxen	-0.4586***	-0.3931***	-0.3864***	-0.4200***	-0.3652***	-0.3622***
	(0.0171)	(0.0173)	(0.0174)	(0.0167)	(0.0169)	(0.0170)
Constant	2.9033***	4.1287***	4.2659***	5.9658***	5.6361***	5.7669***
	(0.0707)	(0.0816)	(0.0850)	(0.3690)	(0.3522)	(0.3572)
Observations	33,721	33,721	33,721	33,721	33,721	33,721
Log-likelihood	-53495	-52288	-52226	-52537	-51533	-51489
Fixed effects (regions)	Yes	Yes	Yes	Yes	Yes	Yes
Inefficiency effects (mean)	No	Yes	Yes	No	Yes	Yes
Inefficiency effects (variance)	No	No	Yes	No	No	Yes
Subsidy variable (Dummy)	No	Yes	Yes	No	Yes	Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Using the parameter estimates in models (3) and (6) and the JLMS formula in Jondrow et al. (1982)[36], we encounter that ATE scores are 43% and 46% respectively (see Table 4). Farm-specific TE scores vary between 0.42% and 90.23% in these models. Other things equal and using sample means, Table 4 shows the elasticities of the corresponding inputs. These results show that farms in the sample exhibit increasing RTS since we cannot reject the null hypothesis of sum of elasticities equal 1.08 in model (3) and equal 1.14 in model (6) at the 1% significance level. Furthermore, we can reject the null hypothesis of constant RTS at the 1% significance level.

Annual expenses on intermediate inputs (seeds, fertilisers, herbicides, etc.) are the main determinants of total output. A 1% rise in intermediate inputs increases total output by approximately 0.55-0.56%.[37] Such finding is in line with other studies (see for example

[36] Technical efficiency via: $TE = \exp(-E(u|e))$.
[37] One may argue that the aggregation of intermediate inputs in a single category assumes that the production process is separable. It implies that, for example, the marginal rate of technical substitution (MRTS) between seeds and fertilisers is independent of the number of tractors, working hours, and land. This assumption seems slightly unrealistic; however, we do not address this issue in this analysis and the reader may be aware of the implications of such assumption.

Latruffe et al. 2017). These findings suggest that the allocation of one additional hectare of land[38] to (annual or seasonal) crops leads to a 1.40-1.56% rise in total output (at means). Among others, Zhu and Lansink (2010) and Giannakas et al. (2001) encounter a similar land-output elasticity in crop farms in Sweden (0.43) and wheat farms in Saskatchewan, Canada (0.44), respectively.

Table 4. ATE, elasticities in the frontier, and returns to scale

	Model (3)	**Model (6)**
	Average Technical Efficiency	
Mean	42.70%	45.80%
SD	20.22%	19.78%
Range	[0.42%-89.66%]	[0.47%-90.23%]
Variable	**Estimated elasticities (at means)**	
Capital	0.0172***	0.0165***
	(0.006)	(0.0039)
Land	0.4156 **	0.4636 **
	(0.0077)	(0.0087)
Labour	0.0918***	0.0889***
	(0.0072)	(0.0077)
Inputs	0.5465***	0.5654***
	(0.0058)	(0.0063)
	Returns to scale	
Null hypothesis: CRTS (Chi-2)	73.25***	170.15***
Probability (Chi-2)	[0.0000]	[0.0000]
Null hypothesis: IRTS (Chi-2)	1.14	0.31
Probability (Chi-2)	[0.2857]	[0.5805]

CRTS: Constant Returns to Scale.
IRTS: Increasing Returns to Scale (sum of elasticities equal 1.08 CD and 1.14 TL)
*** p<0.01, ** p<0.05, * p<0.1

Regarding labour, farmers need to spend or hire 650-672 additional working hours[39] on farming activities to increase total output by 1%. Hadley (2006) computed similar labour-output elasticities in England and Wales. Surprisingly, the capital-output elasticity is slightly small. However, Brumer and Loy (2000) identified an output elasticity of capital of 0.049 in Northern Germany. We acknowledge that the capital variable suffers from measurement errors,

[38] One hectare represents 3.37% of the sample average (29.64 has).
[39] It represents 11% of the current sample mean.

e.g. it does not include the cost of buildings and oxen, and that may influence the size of the corresponding elasticity.

4.2. Technical inefficiencies

The TI equation uses the JLMS estimator of u_i in equations (9) and (10) as dependent variable. Table 5 shows the parameter estimates of inefficiency effects. Regardless of the functional form of the frontier, all coefficients are in line with our initial expectations. Since *age* of the farmer proxies experience in farming activities, the negative sign on the associated coefficient indicates that older farmers are more efficient than young agriculturalists. This finding is in line with previous estimations (e.g. Coelli and Batesse, 1996). More *years of schooling* may improve farmers' abilities to acquire knowledge related to farming activities, especially to avoid waste of resources. The corresponding coefficient on education suggests that additional years of schooling reduce technical inefficiencies. This result further support the view of education contributes to the efficient allocation of resources and the optimal use of existing technology (Sotnikov, 1998; Dinar et al., 2007).

Parameter estimates of the TI equation reveal that the ratio of *owned area* to total farmland shrinks the gap between the current output and the frontier. In this case study, borrowing, renting or using land under the *'a medias'* scheme[40] leads to technical inefficiencies. Rezitis et al. (2003) argue that renting land increases TE but we do not encounter evidence in favour of such conjecture. Such finding might be an indication that tenants or farmers borrowing land through the *'a medias'* scheme exhaust the properties (fertility) of land and therefore, such fields are less productive. *Hiring labour* incentivises farmers to operate closer to the frontier. This effect contradicts the hypothesis that family labour is better skilled or more involved in

[40] Agreement between landowners and farmers in which both parts split production costs and total revenue (or losses) in halves.

farming activities than hired labour (Zhu and Lansink, 2010). Thus, the pressure for clearing labour costs forces farmers to be more efficient.

Specialisation alienates farmers from the frontier. The associated coefficient to the Herfindahl index indicates that higher proportions of land allocated to the production of a single crop increases TIs. Such a finding suggests that benefits from diversification exceed benefits from specialisation. The standard assumption that specialisation boosts efficiency does not hold (Latruffe et al., 2011). In line with the initial expectations, as the ratio of *irrigated area* to total farmland increases, TI goes down. Since the share of irrigated land is an indicator of land with unreliable rainfall, this effect captures to what extent famers can cope with shortages of water by replacing rainfall with irrigation (Karagiannis and Sarris, 2002).

The associated coefficients to the *abroad* variable are not statistically different from zero. It is been argued in the literature that farms selling some of or the produce abroad tend to use resources more efficiently due to high competition in such markets. Surprisingly, we do not encounter evidence supporting this argument. There are few farms in Mexico selling their products abroad directly. Most producers sell raw products to intermediaries and food processors, who finally export such agricultural commodities to the US and other markets. The NAS survey does not collect information about the value chain of crops. Therefore, measurement errors of this variable may lead to insignificant results.

The SFA model identifies a negative *subsidy*-TE link in crop farms in Mexico. The negative effect is consistent under various specifications of the SFA model. The subsidisation programme indirectly pushes farmers to operate further away from the maximum attainable output. Such result suggests that PROCAMPO discourages farmers' to use available resources more efficiently in order to obtain higher revenues. PROCAMPO subsidy payments might compensate low-income farmers, and then such farmers put less effort on farming activities.

Furthermore, some beneficiaries might not use inputs, purchased with the subsidy payment, optimally, or might never use such inputs. This behaviour leads to higher inefficiencies.

Table 5. Parameter estimates of the inefficiency effects model

Variables			***Technical inefficiency (mean)***			
		Cobb-Douglas			***Translog***	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Age		0.0004	-0.0049*		0.0001	-0.0078***
		(0.0016)	(0.0026)		(0.0018)	(0.0029)
Schooling		-0.0353***	-0.0618***		-0.0451***	-0.0734***
		(0.0057)	(0.0110)		(0.0069)	(0.0119)
Owned area		-0.0171	-0.1060***		-0.0303**	-0.1171***
		(0.0132)	(0.0185)		(0.0150)	(0.0211)
Hired labour		-0.2921***	-0.2214***		-0.3042***	-0.2851***
		(0.0212)	(0.0274)		(0.0264)	(0.0387)
Herfindahl		2.3092***	0.6779**		2.2954***	1.0188**
		(0.2874)	(0.3234)		(0.3360)	(0.4107)
Irrigated area		-0.5871***	-0.3750***		-0.6422***	-0.4545***
		(0.0455)	(0.0414)		(0.0597)	(0.0550)
Abroad		-0.1936	0.3365		-0.0188	0.5321
		(0.6136)	(0.6980)		(0.6957)	(0.7926)
Procampo (dummy)		0.3587***	0.7289***		0.4435***	0.7109***
		(0.0460)	(0.0890)		(0.0566)	(0.0906)
Constant		-0.7258**	1.1612***		-0.9732***	0.9965***
		(0.3028)	(0.3068)		(0.3676)	(0.3546)
Variables			***Technical inefficiency (variance)***			
		Cobb-Douglas			***Translog***	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Age			0.0037*			0.0051***
			(0.0020)			(0.0019)
Schooling			0.0204***			0.0205***
			(0.0067)			(0.0062)
Owned area			0.0962***			0.0836***
			(0.0161)			(0.0153)
Hired labour			-0.0166			0.0213
			(0.0170)			(0.0181)
Herfindahl			0.8711***			0.4884*
			(0.2586)			(0.2518)
Irrigated area			-0.0487***			-0.0268
			(0.0186)			(0.0192)
Abroad			-0.3834			-0.3316
			(0.4683)			(0.4275)
Procampo (dummy)			-0.3975***			-0.2732***
			(0.0589)			(0.0544)
Usigma_constant	-0.8162***	0.6073***	-0.6143**	-0.9010***	0.6333***	-0.4798**
	(0.0347)	(0.0697)	(0.2485)	(0.0348)	(0.0860)	(0.2381)
Vsigma_constant	-0.0249*	-0.2348***	-0.2549***	-0.0712***	-0.2259***	-0.2547***
	(0.0139)	(0.0179)	(0.0190)	(0.0137)	(0.0172)	(0.0182)
Sigma_u	0.6649***	1.3548***		0.6373***	1.3725***	
	(0.0115)	(0.0472)		(0.0111)	(0.0590)	
E(sigma_u)			1.2182			1.2633
Sigma_v	0.9876***	0.8892***	0.8803***	0.9650***	0.8932***	0.8804***
	(0.0069)	(0.0080)	(0.0084)	(0.0066)	(0.0077)	(0.0080)
Lambda	0.6732***	1.5236***		0.6604***	1.5366***	
	(0.0168)	(0.0474)		(0.0161)	(0.0590)	
Observations	33,721	33,721	33,721	33,721	33,721	33,721
Log-likelihood	-53495	-52288	-52226	-52537	-51533	-51489

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Interestingly, parameter estimates of the variance TI equation in Table 5 indicate that PROCAMPO reduces production uncertainty, that is, the variance of TI (Wang, 2002). Somehow, these cash transfers ensure a portion of farmers' revenues. We also encounter empirical evidence in favour of the standard assumption that specialisation increases production uncertainty (risk). The associated coefficient to the Herfindahl index in the variance TI equation suggests that as diversification of crops increases (lower values of the Herfindahl index) the variance of TI goes down. As expected, irrigation also reduces uncertainty in the production of crops.

Before we analyse observation-specific marginal effects from the single-step estimation, let us examine the marginal effects along the entire distribution of TI using RIF-regressions. Using the JLMS formula in equation (8) and parameter estimates in Table 3, we compute the values of the predicted values of the non-negative error term $\widehat{u_i}$. Figures 4a and 4b display the results of the RIF-regressions for each percentile of the TI distribution[41] in the CD and TL models respectively. The horizontal axis indicates the percentile of the predicted non-negative error term ($\widehat{u_i} = -\ln TE_i$). Thus, the closer to zero the more efficient the farm is. The vertical axis measures the marginal effect (coefficient associated to PROCAMPO in the RIF-regression). To adjust standard errors, we use the bootstrapping method with 100 repetitions for each percentile-specific coefficient. Using the adjusted standard errors, dashed lines are the lower and upper limits of the 95% confidence intervals of the percentile-specific marginal effects.

Figures 4a and 4b suggest that the negative effect of PROCAMPO on farms' TE is not the same for all farms. Interestingly, we encounter differential effects, that is, the size of the negative association between PROCAMPO and TI increases with TI. The marginal effect of the subsidy payments is larger for those farms that operate further away from the frontier. For

[41] The set of RIF-regressions uses the same model specification for the mean TI equation as in Table 5.

example, Figure 4b shows that at the 10-th percentile, PROCAMPO increases TI by 3.70%. In contrast, at the 90-th percentile, PROCAMPO rises TI by 15.15%. Both figures show that the subsidy-TI link is not monotonic, which also justifies the use of Wang's model.

Figure 4. Percentile-specific marginal effects of PROCAMPO on technical inefficiency

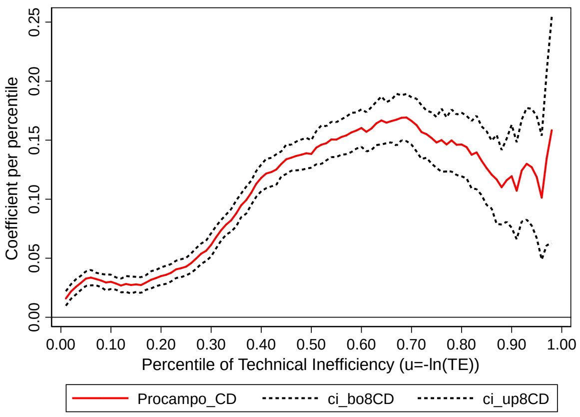


Figure 4a. Model (3)

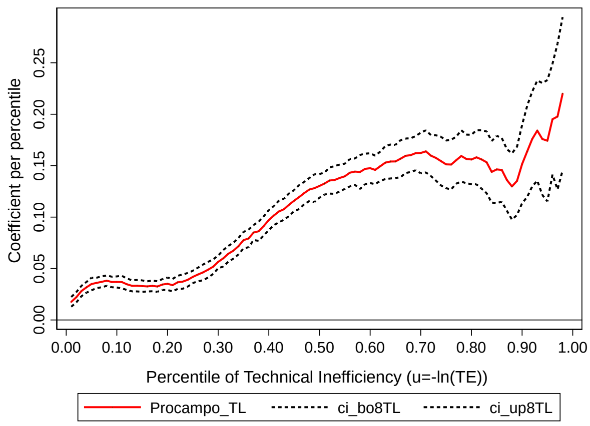


Figure 4b. Model (6)

Turning now to the observation-specific effect of PROCAMPO on farms' TE, we use equation (11), to compute such effects. Table 6 displays the distribution of marginal effects of all variables in the mean TI equation. On average, and other things equal, one additional year of experience reduces TI by 0.02%. Taking the mean of the marginal effect, a 38 years old farmer is approximately 0.38% less efficient than a 57 years old agriculturalist.[42] Thus, experience on farming activities slightly reduce inefficiencies in the production process.

Regarding years of schooling, one more year of education improves the manner in which farmers use available resources and makes them 1.44-1.45% more efficient. Before 1993, primary school education was compulsory and free in Mexico.[43] These studies require 6 years of schooling (sample mean equals 6.13 years) thus most farmers in the sample were subject to such regulation. Since 1993, both primary and secondary education are compulsory and free.

[42] Sample mean equals to 58 years.
[43] Most of the farmers born in 1956 and completed their studies before the 1993 reform.

Therefore, one may expect that three additional years of (secondary) education of younger farmers will rise TE in the subsequent years by approximately 4.32-4.35%.

The mean of the observation-specific marginal effects of land ownership has opposite signs in the CD and TL models. The former (latter) model suggests that land ownership slightly increases (reduces) TI. Therefore, the results for the relationship between area-owned and TI are not conclusive. In any case, increasing the share of area-owned to total area by 10% leads to 0.02% and -0.004% changes in TI scores in average, respectively. Thus, buying or renting land does not considerably improve the use of available resources.

Holding other things fixed if farmers increase the proportion of hired labour to total labour by 10%, TI diminishes between 0.80% and 0.90%. Using the mean of the corresponding variables, hiring an additional full-time worker or 253 *jornales* (2,024 working hours per annum) increases the proportion of hired labour to total labour from 24% to 58%. Consequently, such an adjustment leads to a reduction of 3% in TI scores. Further research is required to distinguish the size of the corresponding marginal effects between full-time workers, temporary workers and *jornaleros*.

Results in Table 6 suggest that a 1% increase in the degree of specialisation makes farmers 0.55-0.64% more inefficient. Currently, 75.54% of the 33,721 farms in the sample allocate all their land to a single commodity. This high degree of specialisation might be the reason for such an effect. To contextualise the size of the (average) marginal effect in Table 6, if an average farmer equally allocates his land to the production of two different crops (Herfindahl equals 0.50), we expect that this farmer would be 25-29% more efficient than a fully specialised farmer (Herfindahl index equals 1). Therefore, benefits from coping with production risks and selecting the most suitable crop for heterogeneous qualities of land (diversification) exceed benefits from specialisation, e.g. more experience on the production of a particular crop. Such finding coincides with previous findings in the existing literature (Manjunatha et al., 2013; Manevska-Tasevska et al., 2013).

Table 6 also shows that installing an irrigation system in a rain-fed field, which is a 100% increase in the percentage of irrigated area, reduces the waste of resources and makes farmers 16% more efficient.[44] For some plants, water stimulates fertilisers and nutrients uptake therefore, the availability of irrigation is crucial since it guarantees the farmer can cope with water shortages form unreliable rainfall. Policy-makers should use this finding to design policies that facilitate the acquisition of irrigation equipment and infrastructure. Selling agricultural commodities abroad is not significant.

Table 6. Marginal effects of variables in the technical inefficiency model

	Model (3)			Model (6)		
Variable	**Mean**	**Min.**	**Max.**	**Mean**	**Min.**	**Max.**
Age	-0.0002	-0.0046	0.0013	-0.0002	-0.0073	0.0019
Schooling	-0.0144	-0.0592	0.0028	-0.0145	-0.0704	0.0033
Owned area	0.0022	-0.0994	0.0388	-0.0004	-0.1103	0.0328
Hired labour	-0.0910	-0.2156	-0.0216	-0.0827	-0.2756	-0.0120
Herfindahl	0.6388	0.2258	0.8226	0.5490	0.1849	1.0071
Irrigated area	-0.1632	-0.3660	-0.0407	-0.1591	-0.4418	-0.0368
Abroad	-0.0411	-0.1726	0.3125	0.0234	-0.1185	0.5030
Procampo (dummy)	**0.1011**	**-0.1048**	**0.6932**	**0.1075**	**-0.0659**	**0.6787**

Source: own elaboration based on inefficiency effects models and Wang (2002)

In average, PROCAMPO increases TI by 10-11% (see Table 6). To examine the distribution of observation specific-marginal effects, Figures 5a and 5b shows the relationship between TE scores (horizontal axis) and the observation-specific marginal effect of PROCAMPO on TI (vertical axis). The vertical (horizontal) dashed line is the ATE in Table 4 (zero line). From

[44] This effect does not arise due to a misspecification of the frontier function since the cost of irrigation in part of the intermediate inputs variable.

Table 6, we find that the subsidy-TI link alternate signs within the sample. This confirms the existence of differential effects of subsidy payments on TI encountered in the RIF-regressions. By simply looking at Figures 5a and 5b, the positive effect of PROCAMPO on TI decreases as TE scores increases. Figures 4a and 4b show similar trends. However, using a more flexible specification of the SFA model, via Wang's model, we encounter that PROCAMPO reduces TI in some farms (farms below the zero line in Figures 5a and 5b). According to models (3) and (6), 31.48% and 26.46% of the 33,721 farms in the sample are below the zero line in Figures 5a and 5b respectively, that is, these farms use PROCAMPO to improve TE.

Figure 5. Marginal effects of PROCAMPO on technical inefficiency

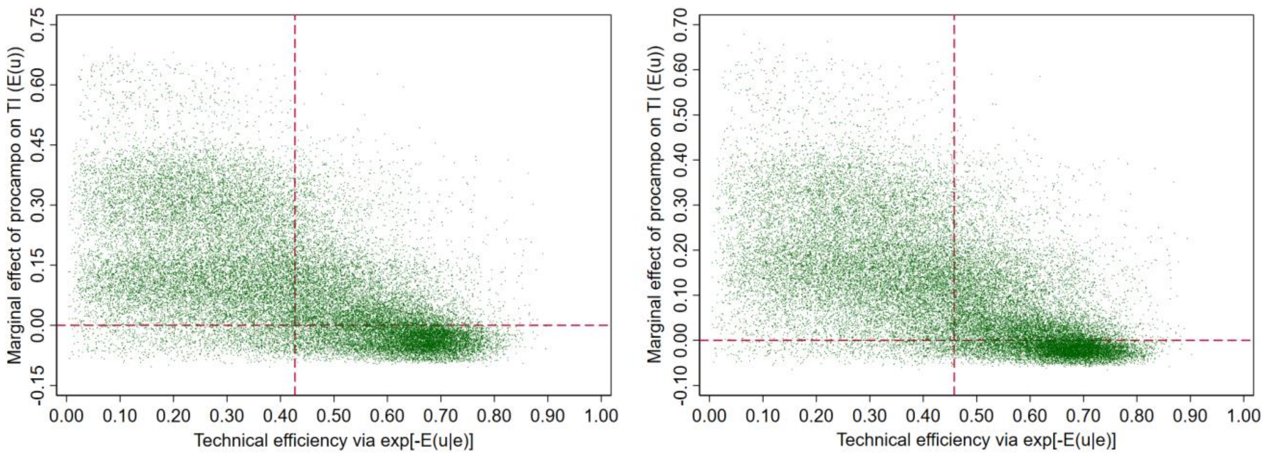


Figure 5a. Model (3) Figure 5b. Model (6)

Models 3 and 6 suggest that PROCAMPO reduces (increases) TI by 3.91% (16.55%) and 2.25% (15.43%) in those farms below (above) the zero line, respectively. We notice that farms above the zero line use less capital, e.g. machinery and equipment, than those below the zero line ($37,974 versus $136,471 of average ownership cost of capital). Furthermore, 23.31% of farms with a positive PROCAMPO-TI relationship use oxen more frequently than farms with a negative subsidy-TI link (10.80%). Farms below the zero line are larger than other farms: 51

hectares versus 22 hectares of utilised land, 11,049 versus 4,145 working hours per annum and $642,069 versus $93,746 of annual expenses on intermediate inputs.[45]

Regarding inefficiency effects, we do not observe significant differences in the age of the farmer and area-owned. However, farmers, who use PROCAMPO to increase TE, tend to be more educated than their counterparts (9.54 versus 4.90 years of schooling). Moreover, such farms hire more labour than other farms (51.27% versus 14.11% of total labour). In this regard, farms with a positive PROCAMPO-TI link tend to use more family labour than other farms, which the farmer cannot discipline easily. The degree of specialisation is slightly larger for farms above the zero line (Herfindahl index equals 0.92 versus 0.89). Farms with a positive PROCAMPO-TE association irrigate 82% of total farmland while other farms only irrigate 11% of the total area. Thus, policy-makers should reverse the negative association between PROCAMPO and TE by: (i) helping farmers to mechanise the production of crops, (ii) providing farmers with extension services (or education), (iii) facilitating the procedures to hire labour, (iv) incentivising crop diversification practices and (v) helping farmers to install irrigation facilities.

4.3. Robustness checks

To verify the consistency of previous results, we conduct a set of robustness checks. Table 7 shows the size of the sample or subsample used in the SFA model, the distribution of ATEs, the coefficient associated to the PROCAMPO variable in the mean TI equation and the distribution of observation-specific marginal effects of PROCAMPO on TI from the CD and TL models respectively. To be consistent with previous results, we use the same functional forms as in Tables 3 and 5.

[45] All these figures are the corresponding subsample means.

We estimate SFA models for both annual and perennial crops. Overall, we do not observe significant differences by including farms with perennial crops in the analysis. We also estimate SFA models for different farm types. The existing literature uses the proportion of revenue attachable to a particular activity to classify farms. Most of the empirical analyses use a threshold of 2/3 of total revenue. Using this criterion, we estimate SFA models for all, beef cattle, arable, mixed and pigs farms. We encounter strongly significant effects of PROCAMPO on farms' TI in the entire sample and the subsample of arable farms. However, for non-arable activities such an effect is not significant. Some of these farms might receive the subsidy since at most 1/3 of total revenues comes from arable activities. Under such circumstances, the subsidy might not be enough to influence TI at the farm-level. This may also apply to mixed farms, which derive less than 2/3 of revenue from arable activities, and therefore the quantity of land they can enrol in PROCAMPO might not be sufficiently large to influence TI.

To examine whether the subsidy-TI link varies among farm size, we split the sample of annual crops into small, medium and large-sized farms. The former type comprises those farms with less than or 5 hectares of land. Medium-sized farms utilise more than 5 hectares and less than 20 hectares of land. Large farms use 20 or more than 20 hectares of land. Table 7 suggests that the parameter associated to the PROCAMPO variable in the mean TI equation is always positive. Nonetheless, observation-specific marginal effects vary among farm sizes. For instance, PROCAMPO reduces TI in some small farms. Such finding suggests that scarcity of resources forces small farms to use inputs more efficiently (see for example Helfand and Levine (2004)). The PROCAMPO-TI relationship is not statistically significant for irrigated farms. However, we encounter a significant subsidy-TI relationship in rain-fed farms. In this regard, rain-fed farms receiving PROCAMPO are in average 12%-13% less efficient than rain-fed farms without the subsidy.

Table 7. Robustness checks

Model/sample	Obs.	Model (3) ATE Mean	Model (3) ATE Min.	Model (3) ATE Max.	Model (3) PROCAMPO (TI equation)	Model (3) Marginal effect Mean	Model (3) Marginal effect Min.	Model (3) Marginal effect Max.	Model (6) ATE Mean	Model (6) ATE Min.	Model (6) ATE Max.	Model (6) PROCAMPO (TI equation)	Model (6) Marginal effect Mean	Model (6) Marginal effect Min.	Model (6) Marginal effect Max.
Annual and perennial crops															
Annual crops	33,721	42.70%	0.43%	89.66%	0.7289***	0.10	-0.10	0.69	45.80%	0.47%	90.23%	0.7109***	0.11	-0.07	0.68
Annual and perennial	36,719	41.52%	0.42%	89.67%	0.6760***	0.10	-0.14	0.65	44.71%	0.50%	90.11%	0.6675***	0.10	-0.08	0.65
Farm types															
All farms	56,529	47.79%	0.70%	89.84%	0.6349***	0.04	-0.18	0.58	50.50%	0.56%	90.18%	0.7208***	0.05	-0.08	0.60
Beef cattle	5,792	16.72%	8.15%	44.82%	0.0436	0.04	0.04	0.04	9.41%	4.03%	28.14%	0.0244	0.13	0.13	0.13
Arable	46,300	46.15%	0.57%	89.39%	0.7403***	0.07	-0.19	0.71	48.55%	0.57%	90.01%	0.6817***	0.08	-0.09	0.65
Mixed	3,908	64.26%	1.08%	91.01%	1.7535	0.00	-0.06	1.36	64.60%	0.83%	91.28%	1.0424	0.01	-0.01	0.91
Pigs	529	66.49%	5.09%	97.98%	1.7521	0.04	-0.33	1.74	37.96%	2.30%	78.26%	-0.3697	-0.10	-0.37	0.15
Farm size															
Small (<=5 has)	13,994	27.73%	3.65%	81.92%	0.3177***	-0.82	-11.89	0.32	27.27%	3.47%	81.81%	0.2905***	-0.75	-10.80	0.29
Medium (>5 & <20 has)	10,764	51.43%	0.53%	89.47%	28.7912	4.83	-2.50	17.32	49.96%	0.35%	89.77%	0.8540***	0.10	-0.03	0.73
Large (=>20 has)	8,963	48.20%	0.04%	92.28%	0.3202	0.02	-0.02	0.31	48.67%	0.04%	91.54%	0.4422*	0.03	-0.03	0.40
Irrigated and rain-fed farms															
Irrigated	12,439	56.17%	0.10%	92.11%	0.6291	-0.03	-0.07	0.62	55.35%	0.08%	92.79%	0.3243	-0.02	-0.04	0.32
Rain-fed	21,282	52.64%	0.83%	88.73%	0.7712***	0.12	0.03	0.52	55.18%	0.74%	88.61%	2.1352**	0.13	0.02	1.00

Annual crops: 100% of total revenue comes from annual crops. Annual crops and perennial crops: 100% of total revenue comes from annual and perennial crops.

All farms: all farms with complete information in the NAS. Beef cattle, arable and pigs farm types: at least 2/3 of total revenue comes from the corresponding activity. Mixed: any of the activities account for 2/3 of total revenue.

Rain-fed farms: farms with share of irrigated area to total area equals zero. Irrigated farms: farms with share of irrigated area to total area greater than zero.

PROCAMPO (TI equation): this is the coefficient of PROCAMPO in the (mean) Technical Inefficiency equation for the corresponding sample.

*** p<0.01, ** p<0.05, * p<0.1

ATE scores vary among farm types, size and the availability of irrigation. On the one hand, mixed farms tend to be more efficient than other farm types. Such result support the previous finding about the positive effect of diversification on TE. Farms producing beef cattle have the lowest ATE scores. There are two potential explanations for such finding. First, the production of beef cattle mainly relies on extensive practices in Mexico therefore larger quantities of land appear in the frontier function. Second, the time at which revenues from this type of farm become apparent does not necessarily coincide with the 2014 agricultural year. Thus, total output reported in the NAS might be below the actual output. Moreover, we use the same measure of capital as in the main SFA estimation, which is not a correct measure of capital endowments for beef cattle farms. Interestingly, we encounter that medium-sized farms tend to be more efficient than small and large farms, which suggests that the optimal size of crop farms is within the 5-20 hectares range. Further investigation is needed to confirm such result. As expected, irrigated farms are slightly more efficient than rain-fed farms.

5. Conclusions

Using the stochastic frontier approach and cross-sectional data on 33,721 crop farms, this research investigates the effect of PROCAMPO subsidy payments on farms' technical efficiency in Mexico. This study contributes to the existing literature by providing empirical evidence on the link between agricultural subsidies and TE in a large developing country where there is no prior evidence concerning any such relationship and computing observation-specific and percentile-specific marginal effects of subsidy payments on TI using Wang's formula and RIF-regressions respectively.

This investigation uses a dummy variable indicating whether the farmer receives the subsidy or not in order to identify the subsidy-TI link. The main findings suggest that: i) the average technical efficiency in the 33,721 crop farms is between 43% and 46%; ii) the negative effect of PROCAMPO on farms' TE increases as technical inefficiency rises; iii) according to the CD and TL models, PROCAMPO negatively influences farms' TE in 68.52% and 73.54% of farms in the sample respectively (positive effects in the remaining farms); and iv) age, years of schooling, area-owned, hired labour, diversification and irrigation increase TE scores.

The estimation of farms' TE scores and the examination of inefficiency effects become relevant since policy-makers should re-evaluate the effectiveness of public policies on farms' performance. Looking at the characteristics of those farms with a positive subsidy-TE relationship, policy-makers should contribute to the mechanisation of production of crops, provide farmers with extension services, facilitate the procedure to hire labour, incentivise crops diversification and help farmers to install irrigation facilities. Furthermore, policy-makers should re-formulate the allocation criteria of the subsidisation programme, which may not be suitable for the current context, e.g. subsidies linked to the level of TE.

To interpret the set of findings in this research the reader should be aware of the following caveats. First, TE (TI) scores of crop farms may be upward (downward) biased because the ownership cost of buildings is not accounted for in the frontier model. Second, some farms might put more land into cultivation in order to enrol their lands into the subsidisation programme. These fields appear in the frontier equation and might cause some biases in the non-negative error term. Further steps of this research should account for endogeneity issues since some of the inputs in the frontier equation might be correlated with stochastic events, which are usually part of the error term.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, *6*(1), 21-37.

Amara, N., Traore, N., Landry, R., & Remain, R. (1999). Technical efficiency and farmers' attitudes toward technological innovation: The case of the potato farmers in Quebec. *Canadian Journal of Agricultural Economics/Revue canadienned'agroeconomie*, *47*(1), 31-43.

Anania, G., Pupo, M. R., & Andrea, M. R. (2015). The 2013 Reform of the Common Agricultural Policy. *The Political Economy of the 2014-2020 Common Agricultural Policy: An Imperfect Storm*, 33-86.

Bojnec, Š., & Latruffe, L. (2009). Determinants of technical efficiency of Slovenian farms. *Post-Communist Economies*, *21*(1), 117-124.

Bojnec, S., & Ferto, I. (2011, April). Impact of off-farm income on farm efficiency in Slovenia. In *Congress of the European Association of Agricultural Economists. Zurich, Switzerland*.

Carroll, J., Greene, S., O'Donoghue, C., Newman, C. F., & Thorne, F. S. (2009, April). Productivity and the determinants of efficiency in Irish agriculture (1996-2006). In *83rd Annual Conference, March 30-April 1, 2009, Dublin, Ireland* (No. 50941). Agricultural Economics Society.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The review of economics and statistics*, 28-45.

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review*, *18*(1), 139-165.

Coelli, T. J., & Battese, G. E. (1996). Identification of factors which influence the technical inefficiency of Indian farmers. *Australian journal of agricultural and resource economics*, *40*(2), 103-128.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., &Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Springer Science & Business Media.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Data envelopment analysis: History, models, and interpretations. In Handbook on data envelopment analysis (pp. 1-39). Springer US.

DOF (1994). Decreto que regula el Programa de Apoyos Directos al Campo denominado PROCAMPO. Diario Oficial de la Federación. México, DF. 25 de Julio de 1994.

Dinar, A., Karagiannis, G., & Tzouvelekas, V. (2007). Evaluating the impact of agricultural extension on farms' performance in Crete: a nonneutral stochastic frontier approach. *Agricultural Economics*, *36*(2), 135-146.

Dung, K. T., Sumalde, Z. M., Pede, V. O., McKinley, J. D., Garcia, Y. T., & Bello, A. L. (2011). Technical Efficiency of Resource-Conserving Technologies in Rice-Wheat Systems: The Case of Bihar and Eastern Uttar Pradesh in India. *Agricultural Economics Research Review*, *24*(2).

Edwards, W. M. (2015). *Estimating farm machinery costs*. Cooperative Extension Service, Iowa State University.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253-290.

Foster, W. E., &Rausser, G. C. (1991). Farmer behavior under risk of failure. *American Journal of Agricultural Economics*, *73*(2), 276-288.

García-Salazar, J. A., Skaggs, R., & Crawford, T. L. (2011). PROCAMPO, the Mexican corn market, and Mexican food security. Food Security, 3(3), 383.

Giannakas, K., Schoney, R., &Tzouvelekas, V. (2001). Technical efficiency, technological change and output growth of wheat farms in Saskatchewan. *Canadian Journal of Agricultural Economics/Revue canadienned'agroeconomie*, *49*(2), 135-152.

González-Konig, G., &Wodon, Q. (2005). Do cash transfers to farmers reduce migration? Procampo in Mexico. No. EC200501, Department of Economics and Finance Working Papers, Universidad de Guanajuato. http://econpapers. repec. org/paper/guawpaper/ec200501. htm.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*, *1*, 92-250.

Hadley, D. (2006). Patterns in technical efficiency and technical change at the farm-level in England and Wales, 1982–2002. *Journal of Agricultural Economics*, *57*(1), 81-100.

Hallam, D., & Machado, F. (1996). Efficiency analysis with panel data: A study of Portuguese dairy farms. *European review of agricultural economics*, *23*(1), 79-93.

Halter, A. N., Carter, H. O., & Hocking, J. G. (1957). A Note on the Transcendental Production Function cx 1 a 1 e b 1 x 1 x 2 a 2 e b 2 x 2. *Journal of farm Economics*, 966-974.

Helfand, S. M., & Levine, E. S. (2004). Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural economics*, *31*(2-3), 241-249.

Iraizoz, B., Bardaji, I., &Rapun, M. (2005). The Spanish beef sector in the 1990s: impact of the BSE crisis on efficiency and profitability. *Applied Economics*, *37*(4), 473-484.

Ito, J. (2015). Diversification of agricultural production in China: Economic rationality of Crop choice under the producer subsidy program. *The Japanese Journal of Rural Economics*, *17*, 1-17.

Justyna, Góral. (2015). Subsidies and Technical Efficiency of Large-Scale Farms in Poland. Agrarian Perspectives XXIV. – Global Agribusiness and Rural Economy International Scientific Conference, At Prague.

Karagiannis, G., & Sarris, A. (2002). Direct subsidies and technical efficiency in Greek agriculture. *Zaragoza (Spain)*, *28*, 31.

Karagiannis, G., & Sarris, A. (2005). Measuring and explaining scale efficiency with the parametric approach: the case of Greek tobacco growers. *Agricultural Economics*, *33*(s3), 441-451.

Karagiannis, G., &Tzouvelekas, V. (2005). Explaining output growth with a heteroscedastic non-neutral production frontier: the case of sheep farms in Greece. *European Review of Agricultural Economics*, *32*(1), 51-74.

Kumbhakar, S. C., & Lien, G. (2010). Impact of subsidies on farm productivity and efficiency. In *The Economic Impact of Public Support to Agriculture* (pp. 109-124). Springer New York.

Kumbhakar, S. C., & Lovell, C. K. (2003). *Stochastic frontier analysis*. Cambridge university press.

Kumbhakar, S. C., Tsionas, E. G., &Sipiläinen, T. (2009). Joint estimation of technology choice and technical efficiency: an application to organic and conventional dairy farming. *Journal of Productivity Analysis*, *31*(3), 151-161.

Kumbhakar, S. C., Wang, H. J., & Horncastle, A. P. (2015). *A practitioner's guide to stochastic frontier analysis using Stata*. Cambridge University Press.

Lakner, S., & Breustedt, G. (2017). Efficiency Analysis of Organic Farming Systems–A Review of Concepts, Topics, Results and Conclusions. *German Journal of Agricultural Economics*, *66*(2), 85-108.

Lambarraa, F., & Kallas, Z. (2009). Subsidies and technical efficiency: An application of stochastic frontier and Random-effect Tobit models to LFA Spanish olive farms. In *The 113th EAAE seminar: A resilient European food industry and food chain in a challenging world*.

Latruffe, L., Bakucs, L., Bojnec, S., Ferto, I., Fogarasi, J., Gavrilescu, C., ...&Toma, C. (2008). Impact of public subsidies on farms' technical efficiency in New Member States before and after EU accession. *ratio*, *3,* 3-24.

Latruffe, L., Bravo-Ureta, B., Moreira, V., Desjeux, Y., &Dupraz, P. (2011, August). Productivity and subsidies in European Union countries: An analysis for dairy farms using input distance frontiers. In *EAAE 2011 International Congress*.

Manevska-Tasevska, G., Rabinowicz, E., & Surry, R. Y. (2013). Policy impact on farm level efficiency in Sweden: 1998–2008. *Policy Impact on Farm Level Efficiency in Sweden: 1998–2008*.

Manjunatha, A. V., Anik, A. R., Speelman, S., & Nuppenau, E. A. (2013). Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land Use Policy*, *31*, 397-405.

Martin, J. P., & Page, J. M. (1983). The impact of subsidies on X-efficiency in LDC industry: theory and an empirical test. *The Review of Economics and Statistics*, 608-617.

McCloud, N., &Kumbhakar, S. C. (2008). Do subsidies drive productivity? A cross-country analysis of Nordic dairy farms. In *Bayesian Econometrics* (pp. 245-274). Emerald Group Publishing Limited.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

Minviel, J. J., & Latruffe, L. (2013). Explaining the impact of public subsidies on farm technical efficiency: Analytical challenges and Bayesian meta-analysis. *2013; 7. Journées de recherchesen sciences sociales (JRSS), Angers, FRA, 2013-12-12-2013-12-13.*

Minviel, J. J., & Latruffe, L. (2014). Meta-regression analysis of the impact of agricultural subsidies on farm technical efficiency. In *presentation at the EAAE 2014 Congress 'Agri-Food and Rural Innovations for Healthier Societies', August* (Vol. 26).

Minviel, J. J., & Latruffe, L. (2017). Effect of public subsidies on farm technical efficiency: a meta-analysis of empirical results. *Applied Economics*, *49*(2), 213-226.

Niroula, G. S., & Thapa, G. B. (2005). Impacts and causes of land fragmentation, and lessons learned from land consolidation in South Asia. *Land use policy*, *22*(4), 358-372.

Pechrova, Marie. (2013). Efficiency of biodynamic farms. Czech University of Life Science Prague.

Rasmussen, S. (2010). Scale efficiency in Danish agriculture: an input distance–function approach. *European Review of Agricultural Economics*, *37*(3), 335-367.

Rezitis, A. N., Tsiboukas, K., &Tsoukalas, S. (2003). Investigation of factors influencing the technical efficiency of agricultural producers participating in farm credit programs: The case of Greece. *Journal of Agricultural and Applied Economics*, *35*(3), 529-541.

Ruiz-Arranz, M., Davis, B., Stampini, M., Winters, P., &Handa, S. (2002). More calories or more diversity? An econometric evaluation of the impact of the PROGRESA and PROCAMPO transfer programmes on food security in rural Mexico (No. 02-09).

Sadoulet, E., De Janvry, A., & Davis, B. (2001). Cash transfer programs with income multipliers: PROCAMPO in Mexico. World development, 29(6), 1043-1056.

SAGARPA (2016). ACUERDO por el que se dan a conocer las Reglas de Operación del Programa de Fomento a la Agricultura de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación para el ejercicio 2017. DiarioOficial de la Federación. México, DF. 31 de Diciembre de 2016.

Samarajeewa, S., Hailu, G., Jeffrey, S. R., &Bredahl, M. (2012). Analysis of production efficiency of beef cow/calf farms in Alberta. *Applied Economics*, *44*(3), 313-322.

Sauer, J., Park, T., &Graversen, J. (2007). Organic farming in Denmark-productivity, technical change and market exit. *Schriften der GesellschaftfürWirtschafts-und Sozialwissenschaften des Landbaus eV Band 43*, 379.

Scott-Andretta, J., &Cuecuecha, A. (2010). The Effect of Agricultural Subsidies on Migration and Agricultural Employment (No. DTE 474).

Serra, T., Zilberman, D., & Gil, J. M. (2008). Farms’ technical inefficiencies in the presence of government programs. *Australian Journal of Agricultural and Resource Economics*, *52*(1), 57-76.

Sotnikov, S. (1998). Evaluating the effects of price and trade liberalisation on the technical efficiency of agricultural production in a transition economy: The case of Russia. *European Review of Agricultural Economics*, *25*(3), 412-431.

Tan, S., Heerink, N., & Qu, F. (2006). Land fragmentation and its driving forces in China. *Land use policy*, *23*(3), 272-285.

Tian, W., & Wan, G. H. (2000). Technical efficiency and its determinants in China's grain production. *Journal of productivity analysis*, *13*(2), 159-174.

Tleubayev, A., Bobojonov, I., Götz, L., Hockmann, H., &Glauben, T. (2017). *Determinants of productivity and efficiency of wheat production in Kazakhstan: A stochastic frontier approach* (No. 160). Discussion Paper, Leibniz Institute of Agricultural Development in Transition Economies.

Zaeske, A. L. (2012). Aggregate Technical Efficiency and Water Use in US Agriculture.

Zellner, A., &Revankar, N. S. (1969). Generalized production functions. *The Review of Economic Studies*, *36*(2), 241-250.

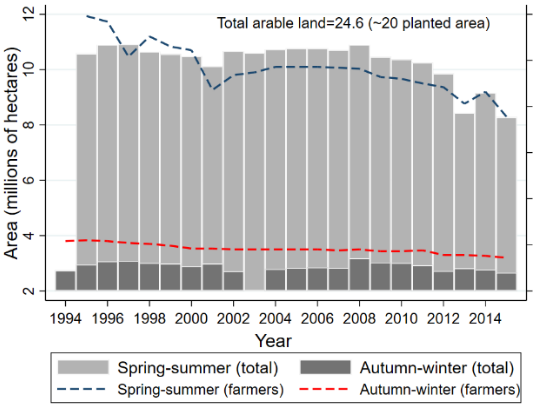
Zhao, M., Yin, R., Yao, L., & Xu, T. (2015). Assessing the impact of China's sloping land conversion program on household production efficiency under spatial heterogeneity and output diversification. *China Agricultural Economic Review*, *7*(2), 221-239.

Zhu, X., & Lansink, A. O. (2010). Impact of CAP subsidies on technical efficiency of crop farms in Germany, the Netherlands and Sweden. *Journal of Agricultural Economics*, *61*(3), 545-564.

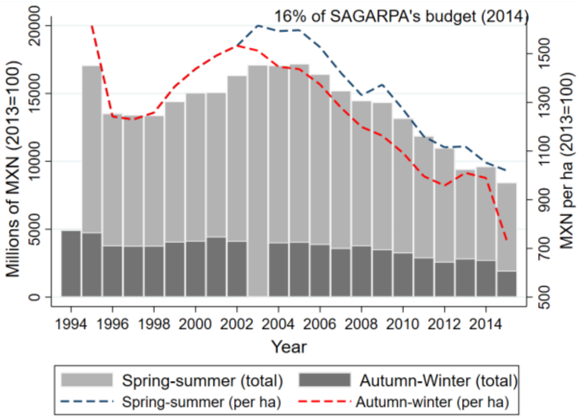
Zhu, X., Karagiannis, G., & Oude Lansink, A. (2011). The Impact of Direct Income Transfers of CAP on Greek Olive Farms' Performance: Using a Non-Monotonic Inefficiency Effects Model. *Journal of Agricultural Economics*, *62*(3), 630-638.

Appendix 1

Figure 6. Distribution of PROCAMPO



8a. Subsidised area and beneficiaries



8b. Total cash transfer

Source: own elaboration based on INEGI (2014) and SAGARPA (2017)

Table 8. CCR Data Envelopment Analysis models

Input-oriented	
Envelopment model	**Multiplier model**
$\min \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)$	$\max z = \sum_{r=1}^{s} \mu_r y_{ro}$
subject to	subject to
$\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = \theta x_{io} \quad i = 1, \ldots, m;$	$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0$
$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{ro} \quad r = 1, \ldots, s;$	$\sum_{i=1}^{m} v_i x_{io} = 1$
$\lambda_j \geq 0 \quad j = 1, \ldots, n;$	$\mu_r, v_i \geq \varepsilon > 0$
Output-oriented	
Envelopment model	**Multiplier model**
$\max \varphi + \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)$	$\min q = \sum_{i=1}^{m} v_i x_{io}$
subject to	subject to
$\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = x_{io} \quad i = 1, \ldots, m;$	$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} \mu_r y_{rj} \geq 0$
$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = \varphi y_{ro} \quad r = 1, \ldots, s;$	$\sum_{r=1}^{s} \mu_r y_{ro} = 1$
$\lambda_j \geq 0 \quad j = 1, \ldots, n;$	$\mu_r, v_i \geq \varepsilon > 0$

θ: efficiency score (ratio) φ: efficiency score (ratio) ε: non-Archimedean element smaller than any positive real number m: total number of i-th inputs s: total number of r-th outputs n: total number of j-th farms to be evaluated s_i^- , s_r^+ : slack variables to transform inequalities into equalities x_{ij}: amount of input i used by farm j	x_{io}: observed input value of the farm to be evaluated y_{rj}: level of output r produced by farm j y_{ro}: observed output value of the farm to be evaluated λ_j: set of parameters z: efficiency score (ratio) q: efficiency score (ratio) μ_r: returns to scale parameter (multipliers) v_i: returns to scale parameter (multipliers)

Source: adapted from Cooper et al. (2011)

Table 9. An overview of the SFA models

Input-oriented	Output-oriented
$y = f(\boldsymbol{x} * \exp(-\eta)), \quad \eta \geq 0$ or $y = f(x * e^{-\eta})$ where for small values of η: $\text{TE} = \exp(-\eta) = 1 - \eta = 1 - TI$	$y = f(\boldsymbol{x}) * \exp(-u), \quad u \geq 0$ or $y = f(x) * e^{-u}$ where for small values of u: $TE = \exp(-u) = 1 - u = 1 - TI$

Functional forms of production functions

Cobb-Douglas production function (Cobb and Douglas (1928))

$$\ln y = \beta_0 + \sum_{j=1}^{J} \beta_j \ln x_j$$

Generalised production function (Zellner and Revankar (1969))

$$\ln y + \theta y = \beta_0 + \sum_{j=1}^{J} \beta_j \ln x_j$$

Transcendental production function (Halter (1957))

$$\ln y = \beta_0 + \sum_{j=1}^{J} \beta_j \ln x_j + \sum_{j=1}^{J} \alpha_j \,.\, x_j.$$

Translog production function (Christensen et al. (1973))

$$\ln y = \beta_0 + \sum_{j=1}^{J} \beta_j \ln x_j + \frac{1}{2}\sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{jk} \ln x_j \ln x_k, \qquad \beta_{jk} = \beta_{kj}$$

y: level of output $f(.)$: production function (frontier) $\boldsymbol{x}$:non-negative input vector η: measurement of input-oriented technical inefficiency	u: measurement of output-oriented technical inefficiency TE: technical efficiency scores TI: technical inefficiency J or K: total number of inputs

Source: adapted from Kumbhakar et al. (2015)

Table 10. Set of variables in the existing literature (Frontier function)

Variable	Description	Units
	Stochastic Production Function	
y_i	Total value of output deflated by the corresponding price index (or quantities)	\$, litres, kg
$land_i$	Total utilised agricultural land	ha
$labour_i$	Total labour including hired and family workers	hours, \$
$capital_i$	Total value of stock of capital or total depreciation value (machinery, buildings, equipment, breeding herd, etc.) or total horsepower of agricultural machinery and total electric motors	\$, HP
$inputs_i$	Other expenses on purchased inputs (intermediate expenses) or quantities e.g. fertilisers, or disaggregated expenses on fertilisers, seeds, crop protection, feed, veterinary fees, energy, etc.	\$, kg/ha
	Stochastic Production Function (other variables)	
ext_i	Use of extension services (public and private)	visits
$arid_i$	Aridity index, ratio of annual temperature to total volume of rainfall	ºC/mm
alt_i	Altitude	masl
$soil_i$	Soil quality (dummy variables)	dummy
$change_i$	Dummy variables for policy reforms	dummy
lfa_i	LFA payments (dummy variables)	dummy

Table 11. Set of variables in the existing literature (Technical inefficiency function)

Variable	Description	Units	Expected sign
	Farmers' characteristics		
edu_i	Farmer's years of education	years	+
age_i	Age of the farmer or farm's manager	years	+
$exper_i$	Number of years as a farmer	years	+
	Farm characteristics (managerial practices and physical characteristics)		
$rented_i$	Ratio of rented/owned land to total agricultural land	%	-+
$hire_i$	Ratio of hired/family labour to total labour or dummy variable for farm hiring labour	%, dummy	-+
$debt_i$	Ratio of total debt to total assets (cows) or total debt	%, \$	-+
$spec_i$	Specialisation, share of the main output in total output, or the Herfindahl index, or multiple cropping index	%	-+
$exter_i$	Share of output sold in the external market	%	-+
$market_i$	Share of marketed output in total output (or self-consumption)	%	-+
$offinc_i$	Off-farm income or off-farm job	\$, dummy	-+
$ltol_i$	Land to labour or capital to labour ratio	ha-\$/worker	
sub_i	Share of total, coupled, or decoupled subsidies in total farm income; or total value of subsidies (per unit of land); or dummy variables indicating whether a farmer receives a subsidy or not	\$, \$/ha, dummy	
ext_i	Use of extension services or participation in management workshops (years)	dummy, years	
$insem_i$	Use of artificial insemination	dummy	
$syst_i$	Different production/farming systems (dummy variables)	dummy	
$treat_i$	Cows under bovine somatotropin treatment	%	
$stall_i$	Use of free stall housing	dummy	
$feed_i$	Ratio of purchased feedstuffs to the number of cows	%	

$save_i$	Family savings	$	
$inves_i$	Investment per cow or total investment	$	
$inten_i$	Intensive farming operations or hectares per livestock unit	dummy, ha/units	
$seed_i$	Type of seeds (modern variety or not)	dummy	
$crop_i$	Share of cropped land	%	
$insur_i$	Share of crop insurance income to total farm income as proxy to weather conditions (or disaster payments)	%, $	
$inputs_i$	Expenses of different inputs per acre (seed, fertiliser, pesticides, veterinary fees, and machinery)	$/acre	
$plan_i$	Improvement plan is carried out in the farm	dummy	
pes_i	Farm environmental payments as a proportion of total income or total amount of environmental payments	%, $	
$enter_i$	Entrepreneurial orientation index	index	
$organ_i$	Organic farms	dummy, %	
$legal_i$	Legal status of the farm	dummy	
mec_i	Number of hours of mechanical operations	hours	
$mobile_i$	Number of telephones per 100 people	units	
$manage_i$	Workers per manager	persons	
$size_i$	Farm or herd size (farmland or number of animals)	ESU	-+
$irriga_i$	Type of irrigation (water pump or not) or irrigation (irrigated or not) or share of irrigated land	dummy, %	+
	External factors		
$price_i$	Price of the relevant output	$/litre,$/kg	
$region_i$	Regional dummies	dummy	
$water_i$	Dummy variables for water rights regimes	dummy	
$road_i$	Road density in the corresponding region	km/km2	
$soil_i$	Index of soil quality	%,EMZ/ha	
lfa_i	Less Favoured Areas (dummy variables)	dummy	
alt_i	Altitude (proxy of geoclimatic heterogeneities)	dummy	
$quota_i$	Milk quota	kg/year	
tax_i	Simplified sales tax	dummy	
$vote_i$	Green voters	persons	
$dist_i$	Distance to the next dairy	km	
$change_i$	Dummy variables for structural changes such as EU accession, policy reforms, etc.	dummy	
env_i	Environmental restrictions	dummy	

Dependent variable:

a) Total output

$$y_i = \sum_{s=1}^{5} y_{s,i}$$

where, y_i is the total output per farm and $s =$ {crops, crops (protected), beef cattle, milk, pigs}. It includes the value of annual crops produced in the 2014 agricultural year[46] and the value of perennial crops[47], beef cattle, milk, and pigs sold in the same period.

Frontier variables:

a) Total ownership cost of capital

b) Total area

c) Working hours

$$working\ hours_i = wh_{hg6,i} + wh_{hl6,i} + wh_{jor,i} + wh_{fam,i}$$

where, wh_{hg6}, wh_{hl6}, wh_{jor}, and wh_{fam} are the total number of working hours spend on farming activities from workers hired for 6 months or more, workers hired for less than six months, 'jornaleros', and family members respectively[48].

d) Intermediate inputs

$$intermediate\ inputs = \sum_{l=1}^{L} expenses_l$$

where, $expenses_l$ are annual expenses and $l =$ {preparation of land/substrate, sowing/planting, fertilisers, plagues control, irrigation, harvesting activities, balanced feed, medicines, vaccines, surgeries, veterinary fees, rent payments (machinery, equipment, and facilities), technical support (extension services), gasoline, diesel, oils, additives, electricity, freight charges, irrigation rights, other expenses, output for self-consumption (seeds and livestock feed)}.

Technical inefficiency variables:

a) Age of the farmer

b) Schooling (years of study)

c) Owned area

$$owned\ area_i = (owned\ area_i / total\ area_i) * 100$$

[46] Barley, maize, oat, rice, sorghum, wheat, beans, chillies, cotton, potatoes, soy, green tomato, melon, onion, red tomato, squash, and watermelon.

[47] Cacao, coffee, apples, avocado, bananas, grapes, lemon, mango, oranges, alfalfa, and sugar cane

[48] There were 253 working days from 1 October 2013 to 30 September 2014. Assuming a working day of 8 hours, the total number of working hours per annum is 2,024. Thus, $wh_{hg6,i}$ is the number of workers hired for or more than 6 months times 2,024. $wh_{hl6,i}$ is the number of workers hired for less than 6 months times 1,012. $wh_{jor,i}$ is the number of 'jornaleros' times the average number of working hours per day and 'jornal' times the total number of working days of the corresponding 'jornalero'. Moreover, $wh_{fam,i}$ is the sum over family members of the average number of hours that each member spends on farming activities per day times 253.

d) Hired labour

$$hired\ labour_i = \left((wh_{hg6,i} + wh_{hl6,i} + wh_{jor,i}\right)/working\ hours_i) * 100$$

e) Specialisation: Herfindahl index

$$HI_i = \sum_{j=1}^{J} \left(A_{ji}/TA_i\right)^2$$

f) Irrigated area

$$irrigated\ area_i = (irrigated\ area_i/total\ area_i) * 100$$

g) External market (dummy variable for farms selling abroad)

h) Procampo (dummy variable)

$$Procampo\_d_i = \begin{cases} 1 & receives\ Procampo \\ 0 & otherwise \end{cases}$$

Table 12. Model specification tests

	Hypothesis 1		**LR-Test**	**d.f.**	**Prob>χ^2**
Model (1) nested in model (2)	H1: $\gamma_0 = 0$ and $\gamma_m = 0$	$\forall\ m$	2,414***	9	0.00
Model (1) nested in model (3)	H1: $\gamma_0 = 0$ and $\gamma_m = 0$	$\forall\ m$	2,537***	17	0.00
Model (4) nested in model (5)	H1: $\gamma_0 = 0$ and $\gamma_m = 0$	$\forall\ m$	2,007***	9	0.00
Model (5) nested in model (6)	H1: $\gamma_0 = 0$ and $\gamma_m = 0$	$\forall\ m$	2,095***	17	0.00
	Hypothesis 2				
Model (2) nested in model (3)	H1: $\vartheta_0 = 0$ and $\vartheta_m = 0$	$\forall\ m$	123***	8	0.00
Model (5) nested in model (6)	H1: $\vartheta_0 = 0$ and $\vartheta_m = 0$	$\forall\ m$	88***	8	0.00