Technology Adoption And Pest Control Strategies Among UK Cereal Farmers: Evidence from Parametric and Nonparametric Count Data Models

Abhijit Sharma and Alastair Bailey and Iain Fraser

Bradford University School of Management, University of Kent, University of Kent

18. June 2009

Technology Adoption And Pest Control Strategies Among UK Cereal Farmers: Evidence from Parametric and Nonparametric Count Data Models*

Abhijit Sharma[†], Alastair Bailey[‡] and Iain Fraser[§]

Thursday 18[th] June, 2009

Abstract

This paper examines technology adoption and integrated pest management strategies employed by UK farmers, using both parametric and nonparametric methods. We employ a unique survey data set collected from UK cereal farmers to assess the determinants of technology adoption in relation to pest management. Our preferred model specification is nonparametric which makes use of the recently developed methods of Li and Racine (2007) and Racine and Li (2004). These methods allow us to combine categorical and continuous data and thereby avoid sample splitting and resulting efficiency losses. Our analysis reveals that total area farmed is positively related to the number of technologies adopted, whereas age is negatively related. We also find evidence of significant statistical differences for number of adoptions by region across the UK.

JEL Codes: O14, Q16
Keywords: technology, adoption, cereal farming, UK, nonparametric

*We wish to acknowledge financial support received from the UK Research Councils through the Rural Economy and Land Use Grant (RES-224-25-0093). We would also like to thank survey participants in our pilot survey, as well as Rothamsted Research for granting us access to these farmers and to all respondents of the main survey itself. We acknowledge the help of the HGCA for allowing us the use of their mailing database to undertake the main survey. We would also like to thank Marco Bertaglia, Elodie Douarin, Mark Freeman, Alex Mohr and Chengang Wang for helpful comments received. The usual disclaimers apply.

†Corresponding author: Abhijit Sharma, Bradford University School of Management, Emm Lane, Bradford BD9 4JL, UK. Ph: +44-1274-234781; Email: A.Sharma12@brad.ac.uk.

‡Department of Economics, University of Kent, UK.

§Department of Economics, University of Kent, UK.

1 Introduction

In this paper we consider the issue of technology adoption by UK cereal farmers in relation to pest management (PM). In particular, we are interested in developing a better understanding of the number of PM technologies adopted by farmers. This is a departure from most conventional studies that examine the adoption of a single technology. However, there are many reasons to assume that farmers employ a portfolio mix approach to PM. To begin with, there is growing agronomic literature examining technology adoption in terms of farm level Integrated Pest Management (IPM) (e.g., Krupinsky et al, 2002). The adoption of IPM requires the use of several related activities within a given farming system (Fernandez-Cornejo, 1996). It is also the case that farmers are more likely to adopt a mix of strategies to deal with pests as a response to ever increasing agricultural policy demands (e.g. Voluntary Initiative in the UK) to reduce pesticide use. Increasing external pressures to reduce pesticide use as well as alternative PM technologies are being actively supported by various government policies such as Environmental Stewardship (Chalak et al., 2009).

In terms of a portfolio approach to PM Holland and Oakley (2007) note that insecticides are frequently employed to reduce expected crop damage, but that they do interact with alternative PM strategies such as the control effect of having beneficial insects within an ecosystem. Even if pesticides remain an important part of the PM tool kit, a reduction in pesticide use could see the wider use of complementary alternative technologies such as biological control and other novel approaches. Complementarity between techniques may include simple additive relations (site and time specific) or temporal additive relations (early to late season active).

In an effort to better understand this issue of portfolio mix, we surveyed cereal farmers in the UK by asking them about the number of PM strategies employed. Following Soule (2001) and Lohr and Park (2002) we interpret the number of technologies adopted as a measure of intensity and diversity of adoption. The motivation for this research stems from the observation that farmers frequently adopt a portfolio of PM strategies, but little is known about this portfolio choice problem in UK cereal farming.

The data we use to examine the PM portfolio choice problem leads us to employ count data models in our analysis. To undertake our analysis we employ a combination of parametric and nonparametric methods. The parametric models estimated are the typical models employed in literature. The nonparametric count data model we employ is based on a recently developed estimation technique introduced by Racine and Li (2004). The appeal of this nonparametric method is that it allows the researcher to efficiently combine categorical and continuous data that has, until now, been a limitation of nonparametric approaches. Thus, by being able to combine data we are able to achieve significant efficiency gains (Hayfield and Racine, 2008; Li and Racine, 2007).

Overall, our analysis provides useful insights into evidence on technology adoption and the role played by farm specific characteristics in the adoption process. Furthermore, we show that our chosen nonparametric model comfortably outperforms conventional parametric models such as the Poisson and Negative Binomial, as well as an OLS model specification. The structure of this paper is as follows. In Section 2 we begin by briefly reviewing the adoption literature. In particular, we examine the agricultural economics literature and the issue of multiple technology adoptions. We then briefly describe the methods we employed to examine our data (Section 3). In Section 4 we describe our data and presents results. Finally, in Section 5, we conclude.

2 Literature Review

There exists a vast literature on the topic of technology adoption in all areas of economics including agriculture (e.g., Grilliches, 1957, Feder et al., 1985, Abadi Ghadim and Pannell, 1999, and Abdulai and Huffman, 2005). This diverse literature has been concerned with all facets of technology adoption. In our research, we examine the literature specifically linked to the number of different technologies adopted by farmers. Our research is driven by the underlying assumptions that the adoption of one specific technology does not necessarily prevent the use of another. It is also likely that the of number of technologies adopted may not be independent in that there may well be a degree of path dependency (Cowen and Gunby, 1996). In other words, the choice of technologies adopted more recently by farmers to deal with pests may well be partly dependent on earlier technology choices. In this context we are far more likely to observe a portfolio of technologies being employed by farmers.

Many existing studies model technology adoption using a dichotomous variable (adopt or not), where determinants of this choice are assessed econometrically (Fernandez-Cornejo et al., 1994). However, in a number of cases, it is not appropriate to model technology adoption as a simple dichotomous choice, as it is the combination of technologies employed that matters. Such situations would include IPM, soil conservation and precision agriculture (Lohr and Park, 2002). That is to say situations where a large set of techniques, with potential complementarities, are available to farmers. It is in this context that technology adoption is more appropriately modelled as a portfolio selection problem.

In principle, portfolio selection can be modelled using a multinomial Logit or Probit specification, where the dependent variable is a categorical variable taking a different value according to the portfolio selected. Such an approach would, however, ignore the issue of complexity or intensity of adoption that is of prime concern when one tries to understand the adoption of potentially complementary techniques. Also multinomial Logit and Probit estimation can be prohibitively computationally demanding when the number of portfolios considered becomes large (Isgin et al., 2008) and would also require a very large sample of respondents.

An alternative approach employed in this paper is to use count data models. Focusing specifically on the issue of technology choices, the reliance on count data is often viewed as limiting, as the complementarities and substitutions among the technologies considered are not modelled explicitly. However, count data models do offer some useful advantages. Count data models pay little attention on the exact combination of techniques chosen within a portfolio and focus instead on the more interesting notion of portfolio complexity or adoption intensity. They enable a more informative investigation of decision making compared to dichotomous choice data models (often involving binary dummies). They also avoid making strong assumptions about relationships between technologies being investigated, as no arbitrary aggregation or lumping of techniques is assumed.

Despite the vast literature on technology adoption there is very little literature that has set out to examine, let alone understand, issue of portfolio choice (size) in technology adoption. Those papers in the literature of most relevance to this research include, Lohr and Park (2002), Isgin et al. (2008), and Kim et al. (2005). An important aspect of all of these studies is that they employ count data models to examine this aspect of technology adoption. The models reported in the literature to date employ either Poisson or the Negative Binomial and could be considered restrictive.

In terms of existing research on this topic the work of Lohr and Park (2002) is important with respect to our analysis. Lohr and Park outlined a model to describe the effects of farm level and regional variables on alternative insect management technology choice within the context of organic farming. In their paper, the number of adoptions is used as a measure of intensity, or portfolio complexity, of adoptions. Lohr and Park consider 11 identified alternative management techniques. In terms of econometric specification, they reject the Poisson model in favour of the Negative Binomial. They found that full time farming does not influence the number of adoptions whereas years of experience and level of education are positively related.

Lohr and Park (2002) find farm size to be negatively related which is as they note unexpected as the conventional hypothesis is that bigger farms are more likely to adopt more technologies. This hypothesis is related to the literature on IPM which suggests the need for more intensive management and associated skills, and these are likely to be found on larger farms implementing IPM. However, they argue that it is less farm size that matters but rather the fixed costs of the technology which can vary by technology type, which is in turn specific to the sample of farms they employ. Lohr and Park (2002) also modelled regional effects which turn out to be statistically significant. They suggest that this result reflects variation in cropping systems, soil & climatic differences as well as differences in U.S. State level policy activities associated with organic farming.

More recently Isgin et al. (2008) examined the number (they also refer to it as intensity) of precision farming technologies adopted by farmers using count data methods. They consider how many technologies are adopted by employing survey data from a group of farmers in Ohio (sample size of n=491). The count data regression models they employ are Poisson and Negative Binomial count data models. Isgin et al. collected a data set of 18 technologies. Like Lohr and Park (2002) they assume that precision farming such as IPM, is more likely to be characterised by the adoption of a number of technologies.

Both Lohr and Park (2002) and Isgin et al. (2008) discuss the main assumptions underpinning a count of technologies as a means to proxy intensity of adoption. The first assumption they consider is that adoption of any one technology does not preclude the adoption of any other. To a degree this reduces the importance of the path dependence argument. Thus, adoption decisions are a result of heterogeneity of the farming situation. Isgin et al. provide empirical support for this argument by referring to sample correlation coefficient estimates. They found that only 12% of their sample adopt more than 50% of the available technologies. Thus, they argue that there is evidence of diffused adoption of the available technologies. [1]

The second main assumption is that there is no limit to the number of technologies adopted except as it relates to profitability. So farms will adopt up to the point at which marginal benefit of the next adoption is equal to the marginal cost of the adoption. This means that choosing more technologies is not necessarily better: it is only better if it is needed and appropriately employed such that marginal portfolio cost does not exceed marginal portfolio benefit. Generally speaking, one would expect that adopting more technologies is better from a portfolio perspective, but it is true only if the farmer knows how to use efficiently all the techniques; if there is an increased risk of pest invasion; and by a variety of pests. In this context, the count data model is used to elucidate

[1] Note that in our sample 22% of farms adopt more than 50% of the technologies considered. However, if we go to the 70% level of adoptions, we have less than 1% of farms undertaking such adoptions. Thus, the extent of sequential adoption decisions does not appear to be too high and thus should not introduce severe selectivity bias.

which farmers are in need of a more complex portfolio (or a more intense adoption of technologies) according to their characteristics and the characteristics of their farms. The model does not assume that adopting more technologies is better, instead it simply investigates the determinants of more complex portfolio choice.

Finally, the existing literature does provide a useful overview of expectations regarding model parameters. Thus, based on the available literature for the farm size variable, Isgin et al. (2008) expect a positive estimate and they find positive evidence. Equally, they expect education to be positively related to the number of technologies adopted whereas age or other variables measuring experience do suggest a lower number of adoptions. This argument is based on the premise that there is a reduced time period over which a new technology will be rewarded. Also farmers with greater experience with existing technologies farmers may be willing to continue their reliance on existing methods and as such there may be a status quo bias. This can also be thought of as a consequence of risk aversion.

In terms of farmer type it is assumed that full time status is positively related to adoption as it requires greater managerial effort. This could act as a constraint on the adoption of the full set of available practices. In contrast, off-farm work simply reduces time available to implement these technologies. Finally, turning to regional impacts, Isgin et al. (2008) indicate that there is little or no theoretical justification to guide us on the impact of such variables on the intensity of technology adoption.

3 Econometric Methods

The existing count data literature on technology adoption typically employs parametric specifications such as the Poisson model and the Negative Binomial. The number of technologies adopted is assumed to be the dependent variable and a set of common characteristics the set of explanatory variables. As these models typically have have no strong theoretical basis, there is little guidance on appropriate functional form or set of explanatory variables to employ. As a result a more meaningful approach to employ in terms of model estimation is to use a nonparametric specification.

In terms of a nonparametric specification we can begin expressing the number of technologies adopted as t_i, using an equation like (1)

$$t_i = f(x_i, z_i) \tag{1}$$

The explanatory variables are grouped into two vectors. A vector z comprises discrete dummy variables and categorical variables. The vector x captures continuous or non-dichotomous variables.

Among the most important problems identified in estimating a relationship of this type are (i) the lack of guidance to choice of functional form for the function $f(\cdot)$, and (ii) mechanisms used to deal with estimation in the presence of categorical explanatory variables z_i. The simplest solution to (i) involves OLS or maximum likelihood estimation of a linear version:

$$t_i = x_i\beta + z_i\delta + u_i \tag{2}$$

where β and δ, while u_i is an error term assumed to have a mean of zero and a finite variance. In the econometric literature Racine and Li (2004) explain how estimation of (1) using nonparametric estimation techniques which do not impose a specific functional form upon some elements of $f(\cdot)$ increases the explanatory power of the regression equation.

The attractiveness of nonparametric methods is that they allow the researcher to determine the appropriate functional form for the data being examined. However, until recently, nonparametric estimation methods have not been able to accommodate categorical data, so semi-parametric methods would be required to estimate the following version of equation (1):

$$t_i = g(x_i) + z_i\delta + u_i \tag{3}$$

In this form the continuous variables are arguments of an unknown function $g(\cdot)$ which is estimated using nonparametric methods, but the discrete variables (i.e., dummy, discrete and categorical) z are accommodated as a linear part of the regression equation. However, using the generalised kernel estimation methods introduced by Racine and Li (2004), we are able to incorporate all of our discrete data as well as the continuous data in a fully nonparametric regression framework. For example, this implies that if there are two continuous variables x and y and a binary categorical variable m for a total of k observations, then if there are p observations for one value of the categorical variable m, we would have to create two subsets. One subset would contain $k - p$ observations and the other subset would contain all the remaining (p) observations. Traditionally, smooth nonparametric regression models would have to be constructed for both subsets separately and further analysis carried out on each individual subset. By employing the methods described by Hayfield and Racine (2008) efficiency losses associated with a reduced sample size are avoided.

The appeal of the Racine and Li (2004) approach to our technology adoption problem is that this type of model cannot be derived precisely from economic theory. By its very nature this issue is ultimately an empirical issue. The key feature of nonparametric methods is that they employ local averaging by use of kernel methods. This implies that they compute a consistent estimate of a conditional mean by locally averaging dependent variable values that are defined as 'close' to the independent regressors. Specifically, we can define our nonparametric regression model as follows:

$$t_i = f(x_i, z_i) + u_i. \tag{4}$$

where x_i and z_i are the vectors of continuous and discrete variables employed in the analysis. We impose no a priori functional form on $f(\cdot)$ allowing it to be flexible. Finally, we assume that u_i has mean zero and variance $var(u_i|x_i) = \sigma^2(x_i)$.

To estimate $f(\cdot)$ we employ the kernel methods developed by Racine and Li (2004) which we describe as follows. We begin by deriving the kernel function for our continuous variables. Assume that $k(\cdot)$ is the univariate kernel function such that the product kernel can be expressed as

$$K((x_i - x)/h) = \prod_{s=1}^{S} k((x_{si} - x_s)/h_s) \tag{5}$$

where x_{si} and x_s are the sth components of x_i and x, and h_s is the smoothing parameter for all the continuous variables. To implement our kernel function we employ a standard normal kernel function

$$k(v) = e^{-v^2/2}/\sqrt{2\pi} \tag{6}$$

Next we summarise the kernel function estimation for the discrete variable z introduced by Racine and Li (2004). The kernel function is defined as

$$l(z_{si}, z_s, \lambda_s) = 1 \text{ if } z_{si} = z_s, \text{ and } \lambda_s \text{ otherwise.} \tag{7}$$

where z_{si} and z_s are sth element ($s = 1$ to S).

We define the smoothing parameter λ_s which is assumed to lie in the unit interval [0,1]. When $\lambda_s = 0$ then the kernel function becomes an indicator function and when $\lambda_s = 1$ the kernel function is a constant for all values of the discrete variables.

Following Racine and Li (2004) we can then write the product kernel for our discrete variables as

$$L(z_i, z, \lambda) = \prod_{s=1}^{S} l(z_{si}, z_s, \lambda_s). \tag{8}$$

Given equations 5 and 8 we can employ a Nadaraya-Watson type kernel to estimate $f(x_i, z_i)$

$$\hat{f}(x, z) = \frac{\sum_{i=1}^{n} p_i K(\frac{x_i - x}{h}) L(z_i, z, \lambda)}{\sum_{i=1}^{n} K(\frac{x_i - x}{h}) L(z_i, z, \lambda)}. \tag{9}$$

An important aspect of the estimation process is the choice we make regarding the smoothing parameters (h, λ). In this paper we select the appropriate bandwidths for our discrete and continuous variables by employing the Expected Kullback-Leibler Cross Validation method. Examples of other studies that have used this approach in the literature include Min et al. (2004) and Henderson and Millimet (2005). By employing cross validation for both discrete and continuous variables, if any variable is found to be independent of nt_i then it is completely smoothed out. This is an important feature of the estimation process that makes this form of nonparametric estimation more efficient than other forms of nonparametric estimation (Racine and Li, 2004).

Finally, given the above nonparametric model we will present our regression results using partial regression plots. These results allow us to present the multivariate regression function via a series of bivariate plots. This is the same approach that has been adopted by Maasoumi et al. (2007).

4 Survey Data and Model

4.1 Survey Data

The data employed in this paper has been obtained as a result of a postal questionnaire sent to UK cereal farmers. The survey targeted farmers involved in cereal production as PM is of prime concern for both the quantity and quality of the harvested cereals. Additionally, it can be reasonably assumed that the cereal farmers all have to decide on how best to prevent pest invasion by selecting a number of practices within a similar set of available options.

Our survey instrument was developed by undertaking a pilot study distributed to 150 farmers. From this we modified the survey instrument. To gain access to a large sample of cereal producers we obtained access to a mailing list owned by the Home Grown Cereals Authority (HGCA) (www.hgca.com). We distributed our survey instrument to 7,500 randomly selected members of the mailing list from a total of 30,000. Overall we received 645 returns of which 574 are useable. Although the response is somewhat on the low side it is in keeping with previous survey returns (e.g., Agricultural Development and Advisory Service (ADAS), 2002).

In terms of some key descriptive data the average size of farm is 295 hectares with some 177 hectares being owned. In terms of the crops being grown wheat was the most common followed by barley. For both crops the reported yields from the survey are in keeping with those reported by Department of Environment, Food and Rural Affairs (DEFRA) (2007). DEFRA reported the mean yield for wheat is 8 tonnes per hectare and 5.9 tonnes per hectare for barley. The survey respondents reported a mean yield for milling wheat of 8.6 tonnes per hectare, for feed wheat of 8.8 tonnes per hectare, for malting barley of 6.7 tonnes per hectare and feed barley of 8.4 tonnes per hectare. Thus, the sample figures are once again comparable with the population statistics for the UK as a whole.

The vast majority of respondents are conventional farmers at 93%. The remaining 7% are organic and these are a reasonable reflection of the proportion of organic farmers in the UK. The average age of respondents is 55 years whereas DEFRA (2007) report the average age of farmers to be 58 years. We also found that the average number years farming is 33 years. In terms of the highest qualification held by farmers the survey found that 85% had achieved some level of college qualification with 21% attaining a university qualification. Again the survey responses are in keeping with those available in the literature. Thus, we are confident that our data are reasonably representative of the UK arable farming sector.

The key piece of information that we wished to collect from the survey relates to the number of pest management technologies adopted. This is because the dependent variable in our regression will be a count of the various technologies adopted by each farmer. A list of the 18 main technologies employed in the UK was constructed as part of the survey design process. This list, along with the rate of adoption by our sample of farmers, is shown in Table 1.

The results in Table 1 indicate a division between pest management practices that are widely adopted and those that are far less prevalent. As might be expected the use of practices such as crop rotation, treatment of seeds and the use of resistant varieties is relatively high. However, one particularly interesting result is the number of farmers using improvements in field margins which can probably be attributed to the widespread adoption of farm management practices as a result of joining an agri-environmental schemes such as Environmental Stewardship. Here, nearly 55% of the sample were members of the Entry Level Scheme, 35% are in the Countryside Stewardship Scheme and approximately 5% are members of the Higher Level Scheme. At the same time other activities that are encouraged by the use of financial incentives such as bettle banks and the growing of flower strips are much lower. This result is likely to have been driven by effectively low payment rates for these manipulations. Finally, some of the more recently developed technologies such as using pheromones, trap crops and beneficial insects, not surprisingly, display very low levels of adoption.

4.2 Count Data Model

Given the characteristics of the data we have collected in our survey we decided to employ a count data regression model. As already noted our dependent variable is the number of pest management technologies adopted by a farmer (which is a simple sum of technologies adopted, reported in Table 1).

Table 1: **Pest Management Technology Adoption**

Technology	**Percentage Adopted**
Rotating crops specifically to prevent pest problems	78
Treating seeds/seedlings (chemical, heat, microbial)	71
Using field margins	70
Planting disease- or insect-resistant varieties	67
Adjusting time of planting to avoid pests	65
Hand rogueing	64
Using different varieties in different fields	63
Spot spraying/ spraying field edges	60
Rotating pesticide classes to avoid resistance	59
Cultivation or using rotary hoe for weeds	39
Using flower strips to encourage beneficial insects	33
Using pheromones for monitoring insects	32
Using beetle banks	28
Using mixed varieties in each field	18
Using pheromones for controlling insects	14
Introducing predators/parasites of insect pests	11
Using a trap crop	9
Other	2

In terms of explanatory variables we employed a large set drawn from our survey instrument. These are shown along with the units of measurement in Table 2.

The set of variables presented in Table 2 have been selected based on existing research in the literature (for instance, Lohr and Park, 2002). In addition, we have constructed a spatial location indicator based on postcodes and reflecting the main agricultural areas in our sample. The inclusion of this variable is to ensure that we capture regional effects that might be associated with climate as well as localised network effects in terms of adoption choice.

For the variables that are categorical (ordered or unordered) such as Education, Profitability, Farm Type, Production System and Regional Location these have to be employed as conventional dummy variables in the parametric models. So, for example, Education enters all of our parametric models as a set of dummy variables for each of the levels attained. In the case of Education, Schooling is omitted from the parametric specifications so as to avoid the dummy variable trap. In addition, for the parametric models estimated, Northern Ireland is the excluded region.

Finally, it is worth reiterating that, because of the power of flexibility of the nonparametric methods developed by Racine and Li (2004) we can use all our dummy and categorical variables without the need to construct a whole set of numerical binary dummy variables. As a result our nonparametric results are for the full set of explanatory variables as presented in Table 2.

Table 2: **Explanatory Variables**

Variable	**Units**
Total Farm Area	Hectares
Years Farming	Years
Employment Status	1 if full time, 0 otherwise
Education	Schooling, Secondary, Vocational, Diploma, University
Farm Profitability	Low, Medium or High
Farm Classified as Arable	1 if Arable, 0 otherwise
Farm Classified as Mixed	1 if Mixed, 0 otherwise
Farm Classified as Other	1 if Other, 0 otherwise
Conventional Production System	1 Conventional, 0 otherwise
Organic Production System	1 Organic, 0 otherwise
Regional Location	**1 if in Region, 0 otherwise**
Region 1	Scotland
Region 2	Wales
Region 3	Northern Ireland
Region 4	East Midlands
Region 5	West Midlands
Region 6	North East
Region 7	North West
Region 8	South West
Region 9	South East

5 Empirical results and discussion

5.1 Empirical results: Parametric

In this section we focus on the results generated by the OLS, Poisson and Negative Binomial specifications. These are presented in Table 3.[2]

Table 3: **Parametric Regression Results**

	OLS		**Poisson**		**Negative Binomial**	
	Estimate	p-value	Estimate	p-value	Estimate	p-value
Intercept	2.975	**0.03**	1.162	**0.000**	1.155	**0.000**
Total Area	0.001	0.134	0.001	**0.096**	0.001	0.164
Years Farming	-0.046	**0.002**	-0.006	**0.000**	-0.006	**0.001**
Full Time	1.17	**0.004**	0.15	**0.001**	0.151	**0.007**
Secondary	-0.524	0.163	-0.072	**0.081**	-0.073	0.158
Vocational	0.175	0.619	0.019	0.599	0.019	0.691
Diploma	-0.064	0.795	-0.011	0.691	-0.078	0.819
University	-0.096	0.767	-0.016	0.642	-0.022	0.624
Medium Profit	0.858	**0.075**	0.128	**0.019**	0.128	0.059
High Profit	1.424	**0.014**	0.189	**0.003**	0.188	**0.017**
Arable	4.143	**0.000**	0.794	**0.000**	0.804	**0.000**
Mixed Farm	3.746	**0.000**	0.751	**0.000**	0.762	**0.000**
Conventional	0.537	0.423	0.068	0.344	0.065	0.47
Scotland	-1.25	0.179	-0.183	**0.089**	-0.174	0.194
North East	0.751	0.439	0.088	0.421	0.105	0.441
North West	1.277	0.32	0.148	0.275	0.156	0.372
East Midlands	1.246	0.141	0.146	0.165	0.161	0.223
West Midlands	0.516	0.612	0.061	0.595	0.074	0.607
Wales	-1.887	0.1	-0.344	**0.019**	-0.355	**0.042**
South East	1.401	0.147	0.163	0.133	0.173	0.205
South West	-0.245	0.793	-0.035	0.752	-0.026	0.847
R^2	21.48					
Pseudo R^2			8.57		5.14	

Dependent - **numadopt**: Number of adoptions.

The first thing to observe about the results presented in Table 3 is that there is a reasonable degree of uniformity regarding the sign of the parameter estimates as well as those which are statistically significant across specifications. For example, we observe that being a Full Time farmer is positively related to the number of technologies adopted, whereas Years Farming is negatively related. We also can observe that, for all specifications, the degree of farm profitability is positively related to the number of technologies adopted. A similar finding is reported for the type of farm.

We find less consistent results with respect to Total Area in that only the Poisson model yields a statistically significant estimate. However, the sign and magnitude of

[2]It should be borne in mind that coefficients for OLS estimations are interpreted differently from those for Poisson and Negative Binomial models.

the estimate is positive and similar for all models. Of the other variables included in the various models all Education dummies, as well as the type of Farming System (Conventional or Organic) and the vast majority of the Region dummies are statistically insignificant.

Finally, comparing relative model fit we observe our OLS specification has the highest R^2. However, for both the Poisson and Negative Binomial models the R^2 is a pseudo R^2. It is well known that goodness-of-fit measures such as a pseudo-R^2 have to be interpreted with extreme caution. Values indicating 'high' levels of fit should be cross-checked with other diagnostics and not over-interpreted. However, values of a pseudo-R^2 indicating poor fit can be relied upon to indicate poor fit. It can be seen that both the Poisson and Negative Binomial models have low pseudo-R^2 values and we read these as such.

5.2 Empirical results: Nonparametric

We now present our nonparametric results.[3] Our model has been estimated using the Kullback-Leibler cross-calidation local-linear bandwidth selection method (Racine and Li, 2004). This data driven method will automatically remove irrelevant variables from the final regression model. To see which variables are irrelevant we present the cross validation estimates for our nonparametric regression in Table 4.

Table 4: **Nonparametric Regression Results**

	Cross validation Values
Total Area	$6\text{x}10^6$
Full Time	0.107
Years Farming	$1\text{x}10^6$
Education	1
Farm Profitability	1
Farm Type	0.018
Farming System	1
Region	0.162
R^2	29.69

The results in Table 4 provide the following information. For all of the discrete variables a bandwidth estimate of one indicates that a variable has been smoothed out of the model. When so indicated, that variable can be considered as insignificant. Thus, Education, Farm Profitability and Farming System are all irrelevant regressors. However, the remaining Regressors, including both continuous variables, are significant and help to explain the number of adoptions.

In terms of overall model performance we found that the resulting R^2 for this specification is 29.69. This is significantly higher than those for our other models. Therefore, in terms of model fit we argue that our nonparametric specification outperforms the other parametric models estimated.

In keeping with existing Racine and Li (2004) applications in the literature we also present our results as a set of partial regression plots. That is, we plot each independent

[3] All estimations have been done using `R` 2.9.0 and the package `np`. We also estimated nonparametric models based on the Aitchison and Aitken kernel as well as binary specification for categorical variables. All results are broadly consistent with those reported here.

variable against the dependent variable of number of technologies employed assuming that all other independent variables take their mean value. These plots show both a mean estimate and an associated confidence interval. Figure 1 shows the plots of our set of partial regression parameter estimates.

For the continuous variables the plots (see Figure 1) show both a point estimate and an associated confidence interval. Thus, for Total Farm Area we can see that as farm size increases, an increasing number of technologies are adopted, although there is also increasing variation as farm size increases. So, although the algorithm has not smoothed this variable out of the specification, statistical support for this relation is lacking. When we consider Years Farming we can see that there is a declining trend and the associated confidence interval is narrow indicating that this is a statistically significant variable. Both of these findings are in agreement with those reported for our parametric specifications.

We next turn to our various categorical variables. First, we observe that being in farming Full Time is significant where Full Time farming is coded as one and not Full Time is zero. We also observe that Farm Type is statistically significant where the labels indicate Arable (1), Other (2) and Mixed (3). It comes as little surprise to find those farms that are classified as being solely Arable employ more PM technologies. It is interesting to note that Mixed Farms employ fewer PM technologies than Arable farms. Mixed farms could be considered as relatively more complex operations than pure arable farms and this may limit the attention farmers can place of pest management as a separate activity. It may also reflect the fact that a lower quality of the crop produced may be tolerated because much of the output may well be intended for animal feed. Finally, the other category describes farms that are predominately livestock orientated who only grow these crops for their use as animal feed. As such the marginal benefit of employing a larger number of PM technologies as a means to prevent pest and disease problems will likely be lower relative to the use of the crop.

In keeping with our parametric results neither Education or Production System are statistically significant. In addition, unlike the parametric specification we do not find the level of Farm Profitability to be statistically significant. The fact that Farm Profitability is not related to the number of adoptions potentially indicates that the choice of how many technologies to adopt is driven by farming system and resulting farm profitability is more to do with how the technologies are employed as opposed to whether they are used or not.

However, we now find that Farming Region is significant (see Figures 1 and 2, especially Figure 2). Figure 2 shows the pattern of adoptions across nine regions of the UK graphically where adoption ins Scotland (1), Northern Ireland (2), North East (3), North West (4), East Midlands (5), West Midlands (6), Wales (7), South East (8) and South West (9) are shown. Thus, from the plot on farming regions we can see that farms in the North East, East Midlands and South East have adopted the greatest number of technologies. On the other hand, farmers in Scotland, Wales and the South West have adopted the smallest number of technologies. This regional break down is probably to be expected given the mix of farm types by location in the UK, climatic differences varying pest and disease pressures and the network effects of the proportion of host crops in the landscape. Since we see more intense arable farming in the east and more mixed and livestock farming in the west, the regions with lower counts could be more marginal areas for arable production with a lower pest pressure, more varied landscapes and cropping systems etc. An alternative explanation could be that in these regions the quality of the crop being produced is such that it does not require or justify as much effort in terms of pest and disease management. Finally, it needs to be borne in mind that Region may well be picking up a significant amount of unexplained variation in the data which may bias our results. As such these results do need to be treated with some degree care.

Overall, our nonparametric model has identified a different mix of explanatory variables that are statistically significant compared to the various parametric models. In particular, the nonparametric model has found that regional differences do exist. We have also found differences in terms of farm profitability which is an important change. Given that we make use of a fuller set of explanatory variables and as well as a large data set where continuous and categorical variables are combined without involving a large loss in number of observations retained for analysis, we are able to derive more robust estimates and we avoid the efficiency losses associated with other parametric and nonparametric methods.

6 Conclusions

In this paper we have examined the issue of technology adoption in relation to pest management. However, unlike the vast majority of papers in the literature that deal with technology adoption our focus has been on the number of technologies employed as opposed to the issue of whether or not a specific technology has been adopted. By employing survey data we have established several key determinants of the number of technologies adopted.

The preferred set of results presented in this paper have been generated by employing the nonparametric methods of Racine and Li (2004). By making use of a recently developed cross-validation method we have combined categorical and continuous data, thereby avoiding sample splitting problems. While computationally burdensome our nonparametric results do not require the imposition of an *a priori* rigid functional specification. This is a particular advantage in cases such as that considered here since we have an absence of a sound theoretical framework governing the choice of any such specification.

Our preferred model suggests that it is full-time, younger or less experienced arable farmers located primarily in the southern and eastern parts of the UK who employ the largest number of PM technologies on their farms. We find no evidence of an influence of educational level or level of profitability, contrary to our prior expectations. Perhaps more surprisingly, our results suggest that organic farmers, whom we would expect would need to rely heavily on some of the technologies considered in the survey, do not appear to adopt a larger number of PM approaches than do their conventional peers. Adoption of PM technologies appears then to be driven by agronomic and climatic factors, the warmer and the more intensively cropped areas have the higher the rate of adoption. We find that some evidence that farm size has a positive and significant impact on the number of techniques adopted as is expected from theory.

As presented, this relatively simple interpretation masks some important anomalies in our results and in addition ignores some key qualitative aspects specific to the adoption choices modelled here.

First, the spatial interpretation of the nonparametric results presented above ignores the case of the South East. If the reason for the north=low, south=high adoption count divide is to be explained on agronomic and climate grounds then it appears strange that farmers in the South East of England have adopted a smaller set of technologies than have their neighbours in the East Midlands. Three possible explanations exist, first, that farmers in the South East face less pressure from pests, weeds and disease and second, that the Rural Payments Agency applies different rules for agri-environmental schemes such as Environmental Stewardship employed here compared to neighbouring regions and lastly, that spatial network effects limit either the returns to individual innovators or promote a psychological barrier to innovation.

Results of count-data models do require some careful interpretation. We have mentioned that a key assumption made by these models is that a higher number of adoptions is always better. In the case of the adoption of non-essential technological components of technology portfolios, this assumption is strong. There is also the question of appropriateness of technologies adopted. For instance, some of the techniques we consider may be beneficial for the control of one crop protection problem such as insect pests, while another technology might be effective against weed problems. A farmer confronted with significant insect pest attack but with weed problems already under control would not be incentivised to adopt weed focused PM technologies from our list. Thus the choice of technology mix as well as technology portfolios adopted are both closely linked to the appropriateness of individual technologies, and interplay between technologies adopted. Interested readers can find a discussion of the portfolio adoption behaviour present in this data set in Bailey et al. (forthcoming). Nonparametric count data models allow us to sensibly aggregate such interlinked adoption decisions, without imposing any further rigid assumptions about functional forms or linkages between technologies within an adoption portfolio.

References

[1] Abadi Ghadim, A.K. and Pannell, D. J. (1999). A Conceptual Framework of Adoption of an Agricultural Innovation, *Agricultural Economics*, 21: 145-154.

[2] ADAS (2002). The Awareness, Use and Promotion of Integrated Crop and Pest Management Amongst Farmers and Growers, A Survey on behalf of DEFRA and the CPA, Agricultural Development and Advisory Service UK.

[3] Abdulai, A. and Huffman, W. E. (2005). The Diffusion of New Agricultural Technologies: The Case of Crossbred-Cow Technology in Tanzania,*American Journal of Agricultural Economics*, 87: 645-659.

[4] Bailey, A. S., Bertalia, M., Fraser, I. M., Sharma, A. and Dourain, E. (2009). Integrated Pest Managament Portfolios in UK Arable Farming: Results of a Farmer Survey, *Pest Management Science*, (forthcoming). Published Online: May 28 2009, DOI: 10.1002/ps.1790

[5] Chalak, A., Balcombe, K., Bailey, A. and Fraser, I. M. (2009). Pesticides, Preference Heterogeneity and Environmental Taxes, *Journal of Agricultural Economics*, 59(3), Forthcoming and Online.

[6] Cowen, R. and Gunby, P. (1996). Sprayed to Death: Path Dependence, lock-in and pest control Strategies, *The Economic Journal*, 106: 521-542.

[7] Feder, G., Just, R. E. and Zilberman, D. (1985). Adoption of Agricultural innovations in Developing Countries: A Survey, *Economic Development abd Cultural Change*, 33: 255-298.

[8] Fernandez-Cornejo, J. (1996). The Microeconomic Impact of IPM Adoption: Theory and Application, J*ournal of Agricultural and Resource Economics Review*, Oct: 149-160.

[9] Griliches, Z. (1957). Hybrid Corn: An Exploration in Economics of Technological Change, *Econometrica*, 46: 1251-1271.

[10] Hall, P., Li, Q. and Racine, J.S. (2007). Nonparametric Estimation of Regression Functions in the Presence of Irrelevant Regressors, *Review of Economics and Statistics*, 89(4): 784-78§9.

[11] Hayfield, T. and Racine, J.S. (2008). Nonparametric econometrics: The np package , *Journal of Statistical Software*, 27(5): 1-32.

[12] Henderson, D. J. and Millimet, D. L. (2005). Environmental Regulation and US State-Level Production, *Economics Letters*, 87: 47-53.

[13] Holland, J. and Oakley, J. (2007). Importance of Arthropod Pests and Their Natural Enemies in Relation to Recent Farming Practice Changes in the UK, Research Review 64, Home Grown Cereals Authority, London.

[14] Isgin, T., Bilgic, A., Forster, D.L. and Batte, M. T. (2008). Using Count Data Models to Determine the Factors Affecting Farmers' Quantity Decisions of Precision Farming Technology Adoption, *Computers and Electronics in Agriculture*, 62: 231-242.

[15] Kim, S. A., Gillespie, J. M. and Paudel, K. P. (2005). Count Data Analysis of the Adoption of Best Management Practices in Beef Cattle Production, Paper presented at the Southern Agricultural Economics Association Annual Meeting, Little Rock, Arkansas, February 5-9.

[16] Li, Q. and Racine, J. S. (2004). Cross-Validated Local Linear Nonparametric Regression, *Statistica Sinica*, 14(2): 485-512.

[17] Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princenton University Press, Princeton, NJ.

[18] Lohr, L. and Park, T. A. (2002). Choice of Insect Management Portfolios by Organic Farmers: Lessons and Comparative Analysis, *Ecological Economics*, 43: 87-99.

[19] Massoumi, E., Racine, J. S. and Stengos, T. (2007). Growth and Convergence: A Profile of Distribution Dynamics and Mobility, *Journal of Econometrics*, 136: 483-508.

[20] Min, I., Fang, C. and Li, Q. (2004). Investigation of Patterns in Food-Away-From-Home Expenditure for China, *China Economic Review*, 15: 457-476.

[21] R Development Core Team (2009). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http: //www.R-project.org.

[22] Racine, J. S. and Li, Q. (2004). Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data, *Journal of Econometrics*, 119: 99-130.

[23] Stiling, P. and Cornelissen, T. (2005). What makes a successful biocontrol agent? A meta-analysis of biological control agent performance. *Biological Control*, 34 (3): 236-246.

Figure 1: **Nonparametric Model Partial Regression Plots**

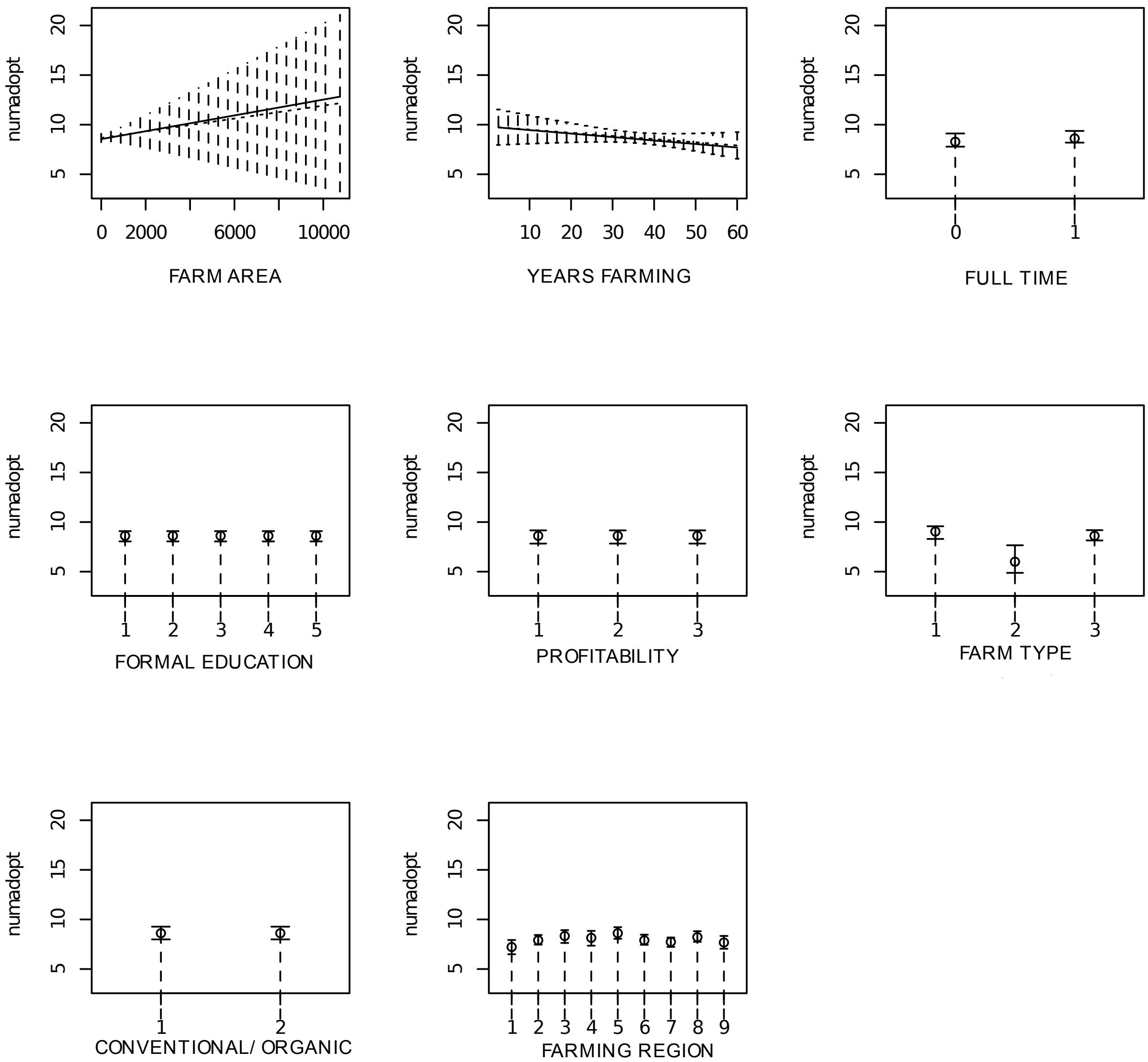

Figure 2: **Regional Adoption Pattern: Partial Regression Plot**