The Economic Journal, 105 (*March*), 333–344. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.

DYNAMIC COUNT DATA MODELS OF TECHNOLOGICAL INNOVATION*

Richard Blundell, Rachel Griffith and John Van Reenen

This paper examines the application of count data models to firm level panel data on technological innovations. The model we propose exhibits dynamic feedback and unobserved heterogeneity. We develop a fixed effects estimator that generalises the standard Poisson and negative binomial models allowing for dynamic feedback through both the firm's stock of knowledge and its product market power. By using the long pre-sample history of innovation information this "entry stock" estimator is shown to control for correlated fixed effects and is compared with an alternative nonlinear GMM estimator. We find evidence of history dependence in innovation activity although variables reflecting the company's economic environment are also found to play a major role.

Technological innovation is an inherently dynamic and nonlinear process. Empirical models seeking to track its progress should share these features. Count data models, where the variable of interest is a non-negative integer, are commonly used to analyse innovation headcounts. However, they are not typically formulated to deal with the dynamic feedback that is suggested by theory. In this paper we examine alternative approaches to modelling innovation counts when there are important dynamics and unobservable heterogeneity.

We model a count of the number of innovations commercialised by a firm in a year as a function of a firm's market power and its tangible and knowledge capital stock. There is clearly going to be some feedback mechanism between market power and innovation – a successful innovation is likely to lead to an increase in a firm's market share. In addition, any representative sample of companies is likely to display a wide range of innovative activity. The majority of companies make few innovations while a small group are involved in a high level of activity. This difference is unlikely to be solely attributable to observable differences across companies. Unobservable permanent heterogeneity is, therefore, an important feature of any empirical model of innovation activity.

In cross-section data firm specific heterogeneity is reflected in a larger number of zero innovation counts than the standard Poisson and negative binomial models would predict. One solution is to use zero-inflated or positive count data models as in Crepon and Duguet (1993) and Silva and Silva (1994). These models allow a different process to describe the number of positive counts from that determining whether or not a count occurs. However, panel data – where there is repeat information on innovations for each company – should

* The authors would like to thank James Banks, Stephen Bond, Alan Duncan, Bronwyn Hall, Frank Windmeijer, participants at the SPES microeconometric seminar in Paris and two anonymous referees. We are obliged to the ESRC Data Archive and SPRU researchers for providing the innovations data tape and to José Montalvo for providing us with his Gauss code. This work is part of a programme of research at the ESRC Centre for Microeconomic Analysis of Fiscal Policy at the Institute for Fiscal Studies. Errors remain, as ever, our own.

provide a more robust alternative as well as allowing explicit examination of the dynamic feedback. The panel data count models developed by Hausman *et al.* (1984) allow for permanent or fixed effects but unfortunately rely on a strict exogeneity assumption for the explanatory variables. Although plausible in some contexts, strict exogeneity will clearly not hold in panel data models where there are important feedback mechanisms.

The worry over dynamic feedback has motivated more recent work on methods allowing for both fixed effects and weakly exogenous (or predetermined) regressors which mirror the dynamic estimation methods for standard panel data models developed by Holtz-Eakin *et al.* (1988) and Arellano and Bond (1991). The papers by Chamberlain (1993) and Montalvo (1993), which extend GMM panel data methods to nonlinear count data models, are the outstanding contributions in this area. They are also of particular relevance for models of technological innovation where past success will have an influence on the probability of current success. We investigate these estimators and also provide an alternative strategy which uses the pre-sample history of the variable of interest to control for the permanent unobservable differences across firms. This alternative method is attractive whenever there is a long history of the endogenous variable, but only a shorter time series for the conditioning variables.

The application we develop focuses on the impact market structure has on innovation. Many empirical studies have shown that large firms tend to have higher rates of R&D and innovate more (e.g. Scherer, 1967) though on the whole empirical work in this area remains inconclusive. One reason for the divergence in empirical results lies in the long-standing unobserved differences between firms. These include such things as the different appropriability conditions of research efforts and technological opportunities facing firms, which are almost impossible to measure.[1]

The policy implication drawn from these results is generally that the static efficiency losses associated with monopoly are offset by gains in dynamic efficiency. However, work at the industry level (e.g. Geroski, 1990) suggests that concentration has a dampening effect on innovations. In Blundell *et al.* (1993) we examine whether ability or incentives play a larger role in driving firms to innovate and find that, while higher market share firms innovate more, firms in competitive industries tend to have a greater probability of innovating. This suggests that firms with large market shares which increase the level of industry concentration would depress the aggregate level of innovative activity. The findings of this paper support this claim.

The structure of the paper is as follows. Section I lays out the Poisson and negative binomial models. Section II discusses alternative methods of controlling for fixed effects. Section III describes the data and in Section IV the results obtained using the various estimators are given. Section V presents some conclusions.

[1] The 'Yale studies', however, represent an attempt to do just this by questioning senior R&D managers about the appropriability conditions in their lines of business (see Levin *et al.* 1987). Questions still remain regarding the subjectiveness and representativeness of this exercise.

I. COUNT DATA MODELS FOR INNOVATION ACTIVITY

In modelling a discrete variable it is well known that the classical linear model is inadequate.[2] For count data the linear exponential – or log-link – family provides a good alternative. We look at two models within this family – the Poisson and the negative binomial. These models have been discussed extensively in Gourieroux *et al.* (1984*a*, *b*) and used by Hausman *et al.* (1984), henceforth GMT (1984) and HHG (1984) respectively, and Cameron and Trivedi (1986).

The common first moment condition for these models is,

$$E(Y_{it}) = e^{X'_{it}\beta}. \tag{1}$$

In our innovation model[3] we write

$$X'_{it}\beta = \theta_0 + \theta_1 M_{it-1} + \theta_2 K_{it-1} + \theta_3 \tau_{it-1} + \theta_4 G_{it-1} + \eta_i + v_t, \tag{2}$$

where M_{it} is a set of measures of market structure (including firm and industry level variables), K_{it} is firms' tangible capital stock, τ_{it} measures knowledge capital in the firms' industry, G_{it} is the firms' accumulated knowledge stock, η_i is a firm specific fixed effect and v_t is a time specific effect.

The simplest form of this model is one in which the dependent variable follows a Poisson distribution, so the variance of Y_{it} is set equal to the mean, and unobservable heterogeneity η_i is ruled out. The negative binomial model provides a useful generalisation allowing for heterogeneity in the mean function and thereby relaxing the variance restriction. However, the heterogeneity allowed for in this way is independent of the regressors and cannot be correlated over time. It therefore cannot be represented by a correlated permanent or fixed effect, η_i. For this reason these standard models are unlikely to provide a suitable representation of innovation behaviour. Nevertheless, as the generalisations for dynamic panel data draw directly on these specifications, it is worth briefly outlining them. In addition they provide a useful base model for comparative purposes.

A. *The Poisson and Negative Binomial Models*

In the Poisson model the conditional probability density function for firm i in year t is given by,

$$\Pr(Y_{it} = y_{it} \mid X_{it}) = \frac{e^{-\lambda_{it}}\lambda_{it}^{y_{it}}}{y_{it}!} \tag{3}$$

with $\lambda_{it} = E(Y_{it})$. As noted above, the mean of Y_{it} is equal to its variance, i.e.

$$E(Y_{it}) = V(Y_{it}) = e^{X'_{it}\beta}. \tag{4}$$

The exponential form is a convenient specification because it ensures non-negativity; however, the variance restriction is unlikely to hold.

It is useful to note that the first moment condition alone can provide consistent estimates of the parameters β. This underlies the Pseudo Maximum

[2] The distribution of residuals is heteroscedastic non-normal, and the predicted probabilities can be above unity.

[3] See Blundell *et al.* (1993) for the derivation of this model.

Likelihood (PML) estimator of GMT (1984) which is based on maximising the objective function,

$$-\sum_{t=1}^{T}\sum_{i=1}^{N} e^{\mathbf{X}'_{it}\boldsymbol{\beta}}+\sum_{t=1}^{T}\sum_{i=1}^{N} Y_{it}\mathbf{X}'_{it}\boldsymbol{\beta} \tag{5}$$

which fits the sample analog of moment restriction (1). This estimator exploits the first moment restriction underlying the Poisson model but relaxes the variance restriction. However, if we are willing to make more specific parametric assumptions about the way the variance differs from the mean then it is possible to obtain more efficient estimators. This is precisely the aim of the negative binomial models.

The relationship between the mean and variance in the negative binomial models may be specified in a number of different ways. Two specifications that have been used by HHG (1984) and GMT (1984) allow the variance to be linear or quadratic in the mean. We take the latter approach and give the mean as (1) and the variance as

$$\mathrm{V}(Y_{it}) = e^{\mathbf{X}'_{it}\boldsymbol{\beta}}+\alpha e^{2\mathbf{X}'_{it}\boldsymbol{\beta}}. \tag{6}$$

A consistent estimator of α can be obtained (GMT, 1984) from an OLS regression using (6) inserting the PML estimate of $\boldsymbol{\beta}$ from the Poisson model. The corresponding PML estimates of $\boldsymbol{\beta}$ for the negative binomial model can then be obtained by maximising the objective function

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\left[Y_{it}\mathbf{X}'_{it}\boldsymbol{\beta}-\left(\frac{1}{\alpha}+Y_{it}\right)\ln\left(1+\alpha e^{\mathbf{X}'_{it}\boldsymbol{\beta}}\right)\right]. \tag{7}$$

B. *Dynamic Feedback and Fixed Effects*

Apart from any feedback through the $\mathbf{X}_{it}$, the innovation process in the models described above is independent over time. Although the negative binomial model allows for some heterogeneity, it was shown that this could not take the form of the permanent unobservable effects η_i in (2). The possible presence of a fixed effect should be examined before its precise modelling is considered and one way in which permanent heterogeneity will display itself is through persistent serial correlation in the unobservables. HHG (1984) suggest testing the specification of the Poisson model using the generalised residual based test

$$\boldsymbol{\Sigma} = \frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i\varepsilon'_i)$$

where ε_i is a $T\times T$ matrix of the following generalised residuals for each firm.

$$\epsilon_{it} = \frac{y_{it}-\lambda_{it}}{\sqrt{\lambda_{it}}}.$$

Non-zero off-diagonal elements in $\boldsymbol{\Sigma}$ will indicate serial correlation. Firm specific effects will induce persistent serial correlation.

If fixed effects are present but are uncorrelated with the regressors then the models described above remain consistent. However, in our model of technological innovation this is unlikely to hold. We first examine an estimator proposed by HHG (1984) which allows us to control for correlated fixed effects but requires strict exogeneity of the regressors. In our dynamic model such an exogeneity assumption cannot hold. This is then compared with methods that allow for correlated fixed effects while requiring only that the explanatory variables be predetermined.

HHG (1984) develop a fixed effects specification that is based on the idea of the conditional maximum likelihood estimator for the panel data logit model. Suppose that we separate out the individual effect η_i from the X_{it} in (2) and write $\lambda_{it} = e^{\eta_i + \mathbf{X}'_{it}\beta}$ in (3). Assume, that Y_{it} is independently distributed conditional on $\mathbf{X}_{it}$ and η_i. HHG (1984) use a conditional maximum likelihood approach where they condition on the sum over time of the dependent variable. This works because the Poisson distribution is a member of the exponential family and as such ΣY_{it} is a sufficient statistic for $T\tilde{\lambda}_i = \Sigma\tilde{\lambda}_{it}$. The fixed effect η_i is eliminated from the likelihood function, conditional on $\mathbf{X}_{it}$ and $\Sigma_{t=1}^{T} y_{it}$. The conditional likelihood becomes

$$\mathrm{L}(\boldsymbol{\beta}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\Gamma(y_{it}+1) - \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}\ln\left[\sum_{s=1}^{T} e^{-(\mathbf{X}_{it}-\mathbf{X}_{is})\beta}\right]. \quad (8)$$

The share of occurrences (innovations in our case) for firm i in a given year is defined by $S_{it} = y_{it}/\Sigma y_{it}$. This model explains the share of total occurrences in each year given the firms' total number of occurrences in T years. However, as noted above, this specification assumes strict exogeneity of the explanatory variables.

An alternative estimator that allows for weakly exogenous regressors and correlated fixed effects can be obtained by adapting the PML moment estimator above. Suppose we write the conditional mean of innovations in period t based on $\mathbf{X}_{it}$ information *up to period t* as

$$\mathrm{E}(Y_{it} \mid X_{i1}, \ldots, X_{it}, \eta_i) = e^{\mathbf{X}'_{it}\beta + \eta_i}. \quad (9)$$

Instead of using conditional likelihood (8) to eliminate η_i in this nonlinear mean specification, Chamberlain (1993) suggests using a quasi-differenced transformation that generates the following moment restriction

$$\mathrm{E}[Y_{it} - Y_{it+1}\, e^{(\mathbf{X}_{it}-\mathbf{X}_{it+1})'\beta} \mid X_{i1}, \ldots, X_{it}] = 0 \quad \text{for} \quad t = 1, \ldots, T-1. \quad (10)$$

As the conditioning set in (10) is dated t or earlier, this moment condition remains valid under weak exogeneity of the regressors (see Blundell *et al.* (1994) for details) and is much like that found in standard panel data estimation with weak exogeneity and correlated fixed effects (see Arellano and Bond (1991), for example). The use of this moment condition also follows from the standard optimal GMM estimator theory (see Hansen (1982), for example). However, moment condition (10) is nonlinear in the parameters of interest $\boldsymbol{\beta}$ which requires a non-linear adaption of the standard estimator (see Montalvo (1993) and Blundell *et al.* (1994)).[4]

[4] In the results presented below we implement this nonlinear moment estimator using the matrix programming language GAUSS. The code is available from the authors upon request.

II. FIXED EFFECTS AND PRE-SAMPLE INFORMATION

An alternative possibility that we adopt is to attempt to measure the unobserved heterogeneity directly. We argue that the main source of unobserved heterogeneity in our innovation model lies in the different knowledge stocks with which firms enter our sample. The 'permanent' capacities of companies successfully to commercialise new products and processes should be reflected in the pre-sample history of innovative success. Essentially this means including a variable in the regression that approximates the build-up of knowledge of the firm at its point of entry into the sample. This technique is usually not feasible for firm panels because only a relatively short time-series dimension is available. However, it is possible for data like ours where we have a long history of the dependent variable.

We argue that pre-sample innovation activity provides a good approximation of the unobservable heterogeneity component η_i above. To see this, first consider the case where we have a fully observed latent variable S_{it} which represents innovation search activity and follows a process of the form,

$$S_{it} = \omega_1 S_{it-1} + \omega_2 x_{it-1} + \eta_i + \epsilon_{it}. \tag{11}$$

Since the x_{it-1} are weakly exogenous, we assume the feedback mechanism has the form,

$$x_{it} = \delta_1 x_{it-1} + \delta_2 S_{it-1} + v_{it}. \tag{12}$$

Eliminating x_{it} from (11) and assuming stationarity, the final form (see Wallis, 1973) expression for S_{it} may be written

$$(1-\delta_1 \mathrm{L})(1-\omega_1 \mathrm{L}) S_{it} = \omega_2 \delta_2 \mathrm{L}^2 S_{it} + \omega_2 \mathrm{L} v_{it} + \eta_i + \epsilon_{it}, \tag{13}$$

where L is the lag operator.

Taking expectations over t for a given i we have the equilibrium condition,

$$(1-\omega_1)(1-\delta_1)\bar{S}_i = \omega_2 \delta_2 \bar{S}_i + \eta_i, \tag{14}$$

where $\bar{S}_i = \mathrm{E}(S_{it})$ and where we have assumed that $\mathrm{E}(\epsilon_{it}) = \mathrm{E}(v_{it}) = 0$. From (14) and using stationarity it is clear η_i is proportional to $\bar{S}_i$, in particular

$$\bar{S}_i = \eta_i[(1-\omega_1)(1-\delta_1) - \omega_2 \delta_2]^{-1}. \tag{15}$$

If we have sufficient time series observations on S_{it} (but not on x_{it}) we can approximate $\bar{S}_i$ using the long pre-sample history of S_{it} and this, in turn, can proxy η_i.

In our case (11) does not describe the process we observe. What is observed is the non-negative count of innovations Y_{it}, not the search activity S_{it}. Nevertheless, the average presample history of Y_{it} will proxy the unobserved individual effect η_i except insofar as the count of innovations is bounded below by zero. As a result we use both the pre-sample average innovation count plus a dummy to indicate a zero value of pre-sample innovation activity. This enables PML methods to be adapted for individual heterogeneity without the need for quasi first differencing as was required for the GMM estimator described above. Since many of the X_{it} are likely to be 'slow moving' using this measure of η_i is likely to improve precision considerably while continuing to allow for dynamic feedback through weakly exogenous explanatory variables.

III. THE INNOVATIONS DATA

The data set is composed of firm level information from company accounts, industry level variables and a count of innovations from the Science Policy Research Unit (SPRU).[5] Innovations (Y) is the count of 'technologically significant and commercially important' innovations commercialised by British firms between 1972 and 1982. The vast majority of observations in our sample contain zero innovations with 5% having a non-zero value.

Around a third of the firms innovated at some point. The number of innovations are spread out fairly evenly over time but appear broadly pro-cyclical (the largest falls are in 1973/4 and the early 1980s). The sample contains 375 firms listed on the London International Stock Exchange[6] in all eleven years from 1972 to 1982 who principally operated in manufacturing. Descriptive statistics on the variables are shown in Table 1.

Table 1

Descriptive Statistics

Variable		All firms	Innovating firms	Non-innovating firms
Firm level variables				
Innovations	*Y*	0·0987	0·3083	—
		0·6180	*1·0627*	—
Knowledge stock	*G*	0·3401	1·0628	—
		1·6831	*2·8441*	—
Fixed effect	*FE*	0·0755	0·2358	—
		0·3528	*0·5926*	—
Fixed effect – dummy	*FEdum*	0·2667	0·8333	—
		0·4423	*0·3728*	—
Market share	*MS*	0·0378	0·0809	0·0175
		0·0868	*0·1245*	*0·0502*
Tangible capital stock	*Capital*	0·1354	0·3237	0·0468
		0·4428	*0·7275*	*0·1216*
Industry level variables				
Concentration	*Conc*	0·4160	0·4235	0·4124
		0·1720	*0·1606*	*0·1770*
Import penetration	*Imp*	0·2454	0·2530	0·2418
		0·1382	*0·1289*	*0·1422*
Union density	*Union*	0·6549	0·6805	0·6429
		0·1136	*0·0983*	*0·1183*
Producer knowledge stock	*G-Prod*	0·3434	0·4304	0·3025
		0·4941	*0·5320*	*0·4699*
User knowledge stock	*G-User*	0·1296	0·1597	0·1154
		0·1434	*0·1457*	*0·1402*

Note: numbers in italics are standard errors.

[5] These data were collected by SPRU in three waves over a period of fifteen years. The full SPRU dataset runs from 1945 to 1983 and includes over 4,500 innovations. The data are described in more detail in Blundell *et al.* (1993), Geroski *et al.* (1993), Robson *et al.* (1988) and Pavitt *et al.* (1987).

[6] The fact that we restrict our attention to listed firms is of some consequence. Previous research has found that smaller firms were often highly innovative although they reported no formal R&D. Medium-sized firms have done relatively poorly given their size. It is likely that this non-linearity will not be picked up when conditioning on publicly quoted firms. It is our belief that public firms are qualitatively different from private firms so we can generalise only to this population from the data.

As described in Section II, a possible measure of the fixed effect (FE) is the average number of innovations by the firm in the period from 1945 to 1971

$$FE_i = \frac{\sum_{1945}^{1971} y_{it}}{27}$$

and a dummy variable ($FEdum$) equal to one if the firm had ever innovated prior to 1971 ($FE_i > 0$) or acquired a firm which had ever innovated pre-sample. The dummy captures the fact that firms who sometimes innovate may be qualitatively different from those who never innovate.

Knowledge stock (G) is the depreciated sum of past innovations and is defined by

$$G_{it} = y_{it} + (1-\delta)\, G_{it-1}, \tag{16}$$

where we have taken the depreciation rate to be 30%.[7]

Market share (MS) is the proportion of sales in firms' principal operating industry. Tangible capital stock (*Capital*) is the replacement value of firm's capital stock which is constructed as in Blundell *et al.* (1992) using perpetual inventory methods. Concentration (*Conc*) is the proportion of sales in firms' principal operating industry that are represented by the five largest domestic firms, and Import penetration (*Imp*) the proportion of sales represented by imports. Union density (*Union*) is the proportion of workers in the firm's principal operating industry that are members of a trade union.

The User and producer knowledge stock (*G-User* and *G-Prod* respectively) are constructed as in (16) but use the count of innovations used in the firms' principal operating industry and the proportion of those produced. These pick up rivalry and spill-over effects.

IV. RESULTS

Table 2 shows the estimates obtained from PML estimation of the Poisson and over dispersion (negative binomial) models and from the nonlinear GMM estimator. Numbers in italics are standard errors.

The signs of the coefficients in column (i) are generally as we expect. Market share enters positively and is very significant. Concentration has a negative impact on the probability of innovation suggesting that the impact of competition on the aggregate number of industry innovations is positive. Firms appear to innovate more in booms to capture the increase in demand as indicated by the negative effects of the recession dummies in 1973–4 and 1980–2.[8] The industry knowledge variables enter with opposite signs. *G-User* is the stronger and positive effect which perhaps reflects complementarities in the production of innovation or rivalry effects. The diagnostics indicate that the

[7] We experimented with other values (15, 25, 50%) and found that the precise rate made very little difference.

[8] The last two variables are time dummies for 1973–4 and the early 1980s respectively. This restriction (against a full set of time dummies) was accepted by a test statistic of 9·02 using an LR test $\sim \chi^2_{(8,\, 0\cdot05)} = 15\cdot5$.

Table 2
Results

Innovation	(i) Poisson	(ii) Poisson	(iii) Poisson	(iv) Over dispersion
Years	1972–82	1972–82	1972–82	1972–82
Observations	4,125	4,125	4,125	4,125
Log likelihood	−1127·687	−979·005	−776·265	
Constant	−5·7523	−4·5228	−4·8831	−4·8993
	0.5185	*0·4854*	*0·5358*	*0·5491*
MS_{t-1}	7·7034	4·9974	2·1637	2·2071
	0·4166	*0·5522*	*0·6379*	*0·6785*
$Conc_{t-1}$	−1·2016	−1·5430	−2·1148	−2·0503
	0·4476	*0·4559*	*0·5000*	*0·5272*
$Imports_{t-1}$	0·2371	−0·3491	0·3045	0·3408
	0·4714	*0·5209*	*0·5610*	*0·6076*
$Union_{t-1}$	3·4957	3·1058	1·2301	1·2326
	0·7046	*0·6884*	*0·7115*	*0·7368*
$Capital_{t-1}$	0·1275	0·0750	−0·0875	−0·0945
	0·0547	*0·0628*	*0·0666*	*0·0740*
$G\text{-}Prod._{t-1}$	−0·7169	−0·2472	−0·6332	−0·6501
	0·2745	*0·2977*	*0·3087*	*0·3345*
$G\text{-}User_{t-1}$	7·6677	3·2261	2·8049	2·8693
	0·9597	*1·0917*	*1·1583*	*1·2595*
G_{t-1}	—	0·1252	0·0318	0·0397
		0·0071	*0·0142*	*0·0176*
FE	—	—	0·6733	0·6562
			0·0885	*0·1040*
FEdum	—	—	3·0325	3·0194
			0·2290	*0·2295*
1973/4	−0·3286	−0·3196	−0·4136	−0·4483
	0·1557	*0·1562*	*0·1547*	*0·1647*
1980s	−0·9138	−1·2075	−0·7737	−0·8654
	0·1318	*0·1437*	*0·1428*	*0·1570*
Serial correlation[a]	198·91		108·71	
Over dispersion parameter[b]				0·1582
				0·0510

Notes:
[a] $\sim \chi^2_{(45,\,0\cdot05)} = 62\cdot15$. See Blundell *et al.* (1994) for more detail.
[b] This is a t-test of the dispersion parameter α in equation (6). The numbers reported are the OLS estimate of α with heteroskedastic-consistent standard error in italics.

Poisson may be inappropriate, however, as we suspect that fixed effects and over-dispersion are features of our data. The serial correlation test (Hausman and Newey, 1984) is based on the off diagonal elements of the residual correlation matrix. In column (i) we reject the null hypothesis of no serial correlation. The pattern of autocorrelation (see Blundell *et al.* (1994)) indicates a long memory and it seems likely that this is due to the presence of fixed effects.

In column (ii) we add a measure of firm knowledge stock – effectively a lagged dependent variable.[9] This variable enters the equation positively and

[9] We have experimented with putting in several lags of innovations, thereby not imposing the structure implied by (16). While tests of this restriction are not accepted, the inclusion of these lags in an unrestricted form does not substantially alter the sign or magnitude of our variables of interest.

significantly as would be expected. The capital stock and number of innovations produced within the industry are both driven to insignificance.

In column (iii) we include our measured fixed effect variables. These are both well determined. Controlling for firm specific effects reduces the market share coefficient by almost half although it remains significant. The size of the knowledge stock coefficient is also dramatically reduced. The strong effect of industry union density is driven into insignificance by controlling for the fixed effect. The basic message is that even when we control for unobserved firm heterogeneity it still appears that dominant firms have a higher propensity to innovate. However, there do appear to be offsetting effects at the industry level – more competitive industries (lower concentration) tend to generate a higher aggregate number of innovations. The inclusion of our measured fixed effects dramatically reduces serial correlation as is evidenced by the fall in the Hausman and Newey test statistic. An analysis of the pattern of residual correlation (see Blundell *et al.* (1994)) strongly suggests that we have eliminated the problem of unobservable permanent effects.

In column (iv) we allow for over-dispersion as in the negative binomial models (6) above. It is clear that the assumptions imposed by the Poisson model do not hold – the dispersion parameter, α, is significantly different from zero. Finally, we can compare these estimates with those obtained using the Chamberlain nonlinear GMM method of controlling for fixed effects. They support the conclusions regarding market share and concentration – market share has a coefficient of 3·3142 (with standard error 0·7940) and concentration – 5·9362 (2·8826). The coefficients are less precisely determined which is not a surprising result. Differenced (or quasi-differenced) models are notoriously noisy when the fundamental information in the explanatory variables move slowly over time. Indeed, several of the other variables were not well determined and we found it difficult to estimate their coefficients accurately. Given that the parameter estimates of interest support our preferred specification from column (iv), our pre-sample fixed effect estimator may be preferable.

V. CONCLUSIONS

This paper has examined various count data models of technological innovation. In particular it has focused on the modelling importance of unobserved heterogeneity with dynamic feedback mechanisms. Economic theory suggests that innovation activity is an inherently dynamic and nonlinear process between heterogeneous firms. Standard ways of dealing with these problems generally rely on the assumption of strict exogeneity which is clearly inappropriate for the innovation process.

Two alternative methods of relaxing the exogeneity assumption in dynamic count data models are considered, that proposed by Chamberlain and Montalvo and an alternative that uses the long pre-sample history of the variable of interest to control for permanent inter-firm differences. The availability in our firm level data of information on innovations from 1945, but accounting data from only 1972, means that the latter is particularly attractive for our application. This "entry stock" estimator is shown to adequately control for fixed effects, virtually eliminating persistent serial correlation in our model. It therefore provides a precise and computationally simple procedure for estimating count data models with dynamic feedback.

Applying these estimators to the traditional question of the influence of market structure on innovative activity renders interesting results. Dominant firms tend to innovate more, even after controlling for fixed effects, though the impact of market share is substantially reduced. Counteracting this is the fact that industry concentration dampens innovative activity. Since it is the case that for a given market size when one firm's market share increases another correspondingly decreases, the overall level of industry innovation remains unchanged. However, to the extent that growing dominance increases industrial concentration the level of aggregate innovation will tend to fall. Thus the anti-trust authorities should remain wary of arguments that monopoly power is the price of a dynamically efficient economy.

Institute for Fiscal Studies and University College London

References

Arellano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.' *Review of Economic Studies*, vol. 58, pp. 277–98.

Blundell, R., Bond, S., Devereux, M. and Schiantarelli, F. (1992). 'Investment and Tobin's Q.' *Journal of Econometrics*, vol. 51, pp. 233–57.

Blundell, R., Griffith, R. and Van Reenen, J. (1993). 'Knowledge stocks, persistent innovation and market dominance: evidence from a panel of British manufacturing firms.' IFS Working Paper No. W93/19.

Blundell, R., Griffith, R. and Van Reenen, J. (1994). 'Dynamic count data models of technological innovation.' IFS Working Paper No. W94/10.

Cameron, A. and Trivedi, P. (1986). 'Econometric models based on count data: comparisons and applications of some estimators and tests.' *Journal of Applied Econometrics*, vol. 1, pp. 29–53.

Chamberlain, G. (1984). 'Panel data.' In *Handbook of Econometrics* (ed. Z. Griliches and M. Intrilagator). Amsterdam: North Holland.

Chamberlain, G. (1993). 'Feedback in panel data models.' Mimeo, Econometric Society Summer Meeting, Boston, June.

Cohen, W. and Levin, R. (1989). 'Empirical studies of innovation and market structure.' In *The Handbook of Industrial Organisation*, vol. I (ed. R. Schmalensee and J. Willig). Amsterdam: North Holland.

Crepon, B. and Duguet, E. (1993). 'Research and development, competition and innovation: what patent data show.' INSEE Working Paper G9314.

Geroski, P. A. (1990). 'Innovation, technological opportunity and market structure.' *Oxford Economic Papers* no. 42, pp. 586–602.

Geroski, P., Machin, S. and Van Reenen, J. (1993). 'The profitability of innovating firms.' *Rand Journal of Economics*, vol. 24, pp. 198–211.

Goureroux, C., Monfort, A. and Trognon, A. (1984*a*). 'Pseudo maximum likelihood methods: application to poisson models.' *Econometrica*, vol. 52(3), pp. 701–20.

Goureroux, C., Monfort, A. and Trognon, A. (1984*b*). 'Pseudo maximum likelihood methods: theory.' *Econometrica*, vol. 52(3), pp. 681–700.

Hansen, L. (1982). 'Large sample properties of methods of moments estimators.' *Econometrica*, vol. 50, pp. 1029–54.

Hausman, J., Hall, B. and Griliches, Z. (1984). 'Econometric models for count data and an application to the patents–R&D relationship.' *Econometrica*, vol. 52, pp. 909–38.

Hausman, J. and Newey, W. (1984). Appendix to Hausman, Hall and Griliches (1984).

Holtz-Eakin, D., Newey, W. and Rosen, H. (1988). 'Estimating vector autoregressions with panel data.' *Econometrica*, vol. 56, pp. 1371–95.

Levin, R., Klevorick, A., Nelson, R. and Winter, S. (1987). 'Appropriating the returns from industrial R&D.' *Brookings Papers on Microeconomics*, pp. 783–831.

Montalvo, J. (1993). 'Patents and R&D at the firm level: a new look.' Mimeo, Harvard University.

Pavitt, K., Robson, M. and Townsend, J. (1987). 'The size distribution of innovating firms in the UK: 1945–83.' *Journal of Industrial Economics*, vol. 35, pp. 297–316.

Schmalensee, R. and Willig, J., *The Handbook of Industrial Organisation*, vol. II. Elsevier Science Publishers BV.

Robson, M., Townsend, J. and Pavitt, K. (1988). 'Sectoral patterns of production and use of innovations in the UK: 1945–83.' *Research Policy*, vol. 17, pp. 1–14.

Scherer, F. (1967). 'Market structure and the employment of scientists and engineers.' *American Economic Review*, vol. 57, pp. 524–31.

Silva, J. and Silva, J. (1994). 'Overdispersion in the positive Poisson regression model.' Mimeo.

Wallis, K. (1973). *Topics in Applied Econometrics*, pp. 121–7. London: Gray-Mills Publishing.