
Monograph:
Brown, S., Farrell, L. and Harris, M.N. (2008) Modelling the Incidence of Self-Employment: Individual and Employment Type Heterogeneity. Working Paper. Department of Economics, University of Sheffield ISSN 1749-8368

Sheffield Economic Research Paper Series 2008010

Sheffield Economic Research Paper Series

SERP Number: 2008010

ISSN 1749-8368

Sarah Brown*, Lisa Farrell and Mark N. Harris

Modelling the Incidence of Self-Employment: Individual and Employment Type Heterogeneity

September 2008

* Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Abstract:

Modelling the incidence of self-employment has traditionally proved problematic. Whilst the individual supply side characteristics of the self-employed are well documented, we argue that the literature has largely neglected demand-side aspects. We explore the determinants of self-employment using individual level data drawn from the U.S. *Survey of Consumer Finances (SCF)*. We present results from an econometric framework, the Parameterised Dogit model, that allows us to separately, and simultaneously, model individual heterogeneity (*i.e.* supply side) and employment type heterogeneity (*i.e.* demand-side) influences that determine self-employment. Our findings suggest that whilst individual characteristics are important determinants of self-employment, there are also factors which are specific to the type of employment that influence whether an individual is self-employed.

Key words: Discrete Choice Models; Dogit Models; Self-Employment.

JEL: J23; J33; C25; C10

Acknowledgments:
We are grateful to the U.S. Federal Reserve Board for supplying the *Survey of Consumer Finances*, 1989, 1992, 1995, 1998, 2001 and 2004. We are grateful to Tim Fry, Donald Poskitt, Stephen Pudney, Mike Shields and Karl Taylor for valuable comments. We are also grateful to participants at the Work, Pensions and Labour Economics Group Annual Seminar, October 2007, and Annual Conference, July 2008, for valuable comments and to Aurora Ortiz for excellent research assistance. The normal disclaimer applies.

I. Introduction and Background

Over the last three decades, there has been considerable interest amongst both academics and policy-makers in the determinants of self-employment. Such interest is not surprising given that the self-employed have emerged as an important part of the labour force in many countries including Canada, 10%; UK, 9%; and the US, 9% (Le, 1999). Furthermore, self-employment and entrepreneurship have been regarded as avenues for raising employment with self-employees and entrepreneurs creating their own jobs as well as potentially creating jobs for others thereby serving to alleviate unemployment and poverty.

A number of approaches have been developed to explore the determinants of self-employment, emphasizing to varying degrees sociological, psychological and economic influences. Recent literature has explored the choice between self-employment and paid employment (see Earle and Sakova, 2000), where individuals compare the utility derived from each sector and then decide which sector to enter. Unemployment push and pull factors also play an important role with displaced workers being pushed, or pulled, into self-employment by supply side considerations (Taylor, 1996). There has been a focus on the attributes of the self-employed concentrating on characteristics such as gender, ethnicity, education, marital status, age and number of children as well as financial factors such as wealth and unearned income. In general, empirical studies specify a reduced form Probit or Logit model of self-employment whereby the vector of explanatory variables contains a combination of individual and labour market characteristics. For comprehensive reviews of the existing empirical literature on self-employment, see Le (1999) and Parker (2004).

Modelling the incidence of self-employment however has proved problematic. Whilst the individual supply side characteristics of the self-employed are well documented, we argue that the literature to date has largely neglected, or indeed misspecified, demand-side aspects that are potentially important in determining self-employment. In this paper, we present results from a flexible econometric framework that allows us to separately, and

simultaneously, model individual heterogeneity (*i.e.* supply side influences) and employment type heterogeneity (*i.e.* demand-side influences) that potentially determine the type of individual who is self-employed. Demand-side influences, which characterize heterogeneity of employment types, may curb the extent to which individuals are free to choose their preferred type of employment. This estimation strategy allows us to distinguish between the differential effects of factors which lead to individual heterogeneity and those which lead to employment type heterogeneity. Thus, we are able to determine the impact of supply side factors on the probability of an individual being self-employed whilst also controlling for demand-side influences. Our focus on both supply-side and demand-side influences makes an interesting contribution to the existing literature, which has tended to focus on supply-side considerations. Such an approach is potentially particularly problematic in the context of policy directed towards encouraging self-employment. As argued by Parker (2004), it is important to explore both the supply of, and the demand for, self-employees in order to attain the socially optimal level of self-employment rather than simply assuming that an increase in self-employment is desirable.

II. Methodology and Data

The Dogit model of Gaudry and Dagenais (1979) extends the Logit framework for multinomial outcomes (MNL), traditionally used to model the incidence of self-employment (relative to other employment types), by introducing additional choice-specific parameters, θ_j, which can be interpreted as heterogeneity of the alternative, *i.e.,* labour market status (or type of employment) itself.[1] In addition, they may also capture unobserved individual heterogeneity, which is common to individuals within a chosen alternative. It is intuitive in this context to consider the choice-set generation framework of Manski (1977). To be specific, in the Dogit model an individual is assumed to be either restricted (sometimes referred to as being 'captive') to one of the *J* outcomes (employment type) or chooses freely

from the full choice set. Therefore, the available choice set faced by individual *i*, $B_i = B \,\forall i$, comprises *J*+1 sets, *J* single outcome 'captivity' sets and one set comprising all *J* outcomes over which the individual can subsequently exercise free choice. The choice set generation process itself can be represented as a random utility maximization model with utilities given by

$$U_{ik}^{(1)} = W_{ik} + \eta_{ik}, \quad i = 1,...,n; k = 1,...,J+1. \tag{1}$$

Under the assumptions that: η_{ik} are independently and identically distributed as a Type 1 Extreme Value variate; $W_{ik} = \log(\theta_k)$; and with the normalization that $W_{iJ+1} = 0$, the probability of individual *i* choosing a single outcome (captive) choice set is given by

$$P_{ij}^{(1)} = \frac{\theta_j}{1 + \sum_{k=1}^{J} \theta_k}, \tag{2}$$

and the probability that individual *i* chooses the full choice set is

$$P_{i,J+1}^{(1)} = \frac{1}{1 + \sum_{k=1}^{J} \theta_k}. \tag{3}$$

The probability that an individual chooses the specified outcome *j* from the full choice set is, in the second stage, derived from the standard random utility maximization model, RUM (Fry *et al*., 1993) of

$$U_{ij}^{(2)} = V_{ij} + \varepsilon_{ij}, \quad i = 1,...,n; k = 1,...,J+1 \tag{4}$$

where $U_{ij}^{(2)}$ is the utility that individual *i* gains from alternative *j* in this second stage, and V_{ij} and ε_{ij} are, respectively, the non-stochastic and stochastic components of utility. For simplicity, V_{ij} is specified as

$$V_{ij} = x_i' \beta_j, \tag{5}$$

[1] This section draws on Fry and Harris (2005).

and, under the assumption that the ε_{ij} independently follow a Type 1 Extreme Value distribution, the resulting probabilities have the standard MNL form (Maddala, 1983). So, utilizing the Manski framework, the Dogit model can be parameterised as

$$P_{ij}^{Dogit} = \frac{\theta_j}{1+\sum_{k=1}^{J}\theta_k} + \frac{1}{1+\sum_{k=1}^{J}\theta_k}\left(P_{ij}^{MNL}\right), \tag{6}$$

where P_{ij}^{MNL} are the simple Logit probabilities for multiple outcomes.[2] Using the indicator function d_{ij} where

$$d_{ij} = \begin{cases} 1 & if\ individual\ i\ chooses\ alternative\ j \\ 0 & otherwise \end{cases} \tag{7}$$

the parameters of the model are now estimated using the maximum likelihood criterion, where the log-likelihood function is

$$L(\varphi) = \sum_{j=1}^{J}\sum_{i=1}^{N} d_{ij} \ln P_{ij}^{Dogit}, \tag{8}$$

with $\varphi' = \left[\left(vec\beta_j\right)', \theta'\right]$ and P_{ij}^{Dogit} is given by equation (6). To explore the robustness of the findings of the existing empirical literature on the determinants of self-employment, we specify the Dogit model, which to our knowledge has not been previously used to model self-employment. Given the existence of labour market rigidities and demand-side influences, it is apparent that the introduction of the additional choice-specific parameters, θ_j, which account for employment type heterogeneity, makes a particularly interesting and potentially important contribution to this area.[3]

We analyse data drawn from the *Survey of Consumer Finances* (*SCF)*, which is a cross-section survey of the balance sheet, pension, income, demographic characteristics and

[2] To analyse the determinants of occupational choice, Brown *et al.* (2008) replace the MNL probabilities with those of the Ordered Generalized Extreme Values model, which allows for potential ordering of occupational outcomes.

[3] It should be noted that these parameters also potentially capture unobserved individual heterogeneity which varies by employment type.

use of financial institutions of U.S. families conducted since 1983 by the U.S. Federal Reserve Board. We pool the 1989, 1992, 1995, 1998, 2001 and 2004 cross-sectional surveys yielding a large nationally representative data set with 122,935 observations available for estimation (we include time dummy variables to take into account time varying changes in tastes). For the pooled cross-section data drawn from the *SCF*, we analyse the probability of being self-employed and distinguish between four types of employment status of individual *i* in year *t*: out of the labour force; laid off or unemployed; employed; and self-employed.[4]

We draw on the existing literature to specify the set of explanatory variables, including controls for individual characteristics such as gender, age, marital status and ethnicity. An individual's age may affect his/her propensity to become self-employed (see, for example, Calvo and Wellisz, 1980 and Kidd, 1993); it may act as a proxy to capture the effects of an individual's awareness, knowledge and experience in the labour market thereby reflecting general human capital. In general, studies have reported a non-linear relationship between self-employment and age (Rees and Shah, 1986) and we allow for this by including a quadratic in age. The role of ethnicity in determining the propensity to become self-employed has also attracted a great deal of attention. The issue of whether discrimination bars employment in certain sectors has been the subject of much debate (Hout and Rosen, 2000). Rees and Shah (1986) find that non-white individuals in the UK have a lower propensity to become self-employed whilst Brock and Evans (1986) find the reverse in the US. Marital status has been included in many empirical studies. As argued by Le (1999), marriage is assumed to represent stability and, as such, may provide a suitable background for risky self-employment. Related factors include household size and the number of children. Individuals

[4] Our sample comprises: out of the labour force (24.53%); laid off or unemployed (3.64%); employed (46.83%); and self-employed (25%). Individuals in the out of the labour force category include: students, homemakers, the disabled, the retired, sick leavers and voluntary workers. We adopt the definition of self-employment used by Fujii and Hawley (1991) who also use the *SCF* to analyse the determinants of self-employment, where the respondent has an active management role in his/her business, farm, professional practice or partnership as his/her main job. It should be noted that the proportion of self-employment is quite high which may reflect the emphasis on business ownership. Throughout the paper, we focus on the individual's first (*i.e.,* main) business. There are a small proportion of individuals who own more than one business (8%).

with dependent children, for example, may be less likely to bear the risk associated with self-employment. Since attitudes towards risk-taking have been a particular area of interest in the self-employment literature, which dates back to Knight (1921), we also include a proxy for attitudes towards risk preferences.[5]

Educational qualifications, which may act as a proxy for ability, have been incorporated into many empirical studies of self-employment. Studies reporting a positive relationship between educational attainment and the probability of self-employment include Rees and Shah (1986), Borjas (1986), Borjas and Bronars (1989) and Evans and Leighton (1989). Alternatively, higher levels of educational attainment may play a signalling role in the labour market with high educational qualifications serving to secure employment in the non-self-employed sector. Evidence supporting an inverse relationship between higher levels of educational qualifications and the propensity to become self-employed include Evans (1989), de Wit and Winden (1983) and Kidd (1993). In summary, the evidence regarding the relationship between education and self-employment remains inconclusive. To control for a different aspect of human capital, we include a control for whether the individual has received vocational training. We also control for (the natural logarithm of) household unearned income. Summary statistics for the variables used in our empirical analysis are presented in Table 1.

III Results

For comparison purposes, in Table 2 we present results from the MNL framework (in the form of estimated coefficients and marginal effects), which has traditionally been used in modeling the incidence of self-employment. The base category in this context is the "out of the labour market" group. For reasons of brevity and given the focus of our paper, we present

[5] In the *SCF*, individuals were asked the following question: *which of the following statements comes closest to describing the amount of financial risk that you are willing to take when you save or make investments? Take substantial financial risks expecting to earn substantial returns; Take above average financial risks expecting to earn above average returns; Take average financial risks expecting to earn average returns; Or not willing to take any financial risks*. We use the responses to this question to create a four point risk attitudes index, which is

the results for the self-employed category only.[6] We briefly comment on these results before turning to those from the Dogit framework.

It is apparent from Table 2 that being male is positively associated with self-employment as is being married, in accordance with the existing literature. Age appears to be positively related to self-employment, albeit at a diminishing rate, whilst being non-white is inversely associated with self-employment. Attitudes towards financial risk are found to be a statistically significant determinant of self-employment, with risk aversion being inversely associated with self-employment. Finally, all levels of educational attainment, with the exception of having a high school diploma or a college degree, are positively associated with self-employment, whereas having a high school diploma or a college degree are inversely associated with self-employment relative to having no education.

In Table 2, we also present the estimates from the Dogit framework. Again, we present estimated coefficients and the marginal effects for the self-employment category only. The patterns of the marginal effects across the MNL results and those from the Dogit framework are broadly in line in terms of the signs of the marginal effects. It is apparent that for some variables, such as being non-white and risk attitudes, the estimated marginal effects from the Dogit model are slightly larger.

It is particularly interesting to note that the distinction between the effects of the relatively low levels of education (high school diploma and college degree) and the high levels of education (Bachelor degree and above) are more pronounced in the context of the Dogit framework: the inverse association between self-employment and the low levels of education (relative to no having no education) is stronger in the Dogit estimates. Similarly, the positive association between self-employment and the relatively high levels of education

---

decreasing in risk aversion. Shaw (1996), who explores income growth and risk aversion, bases one of her empirical measures of risk aversion on this *SCF* question.

[6] We adopt this approach throughout the paper. Results relating to the three other outcomes (*i.e.,* not in the labour force, employment and unemployment) are available from the authors on request. However, the results generally accord with prior expectations and the existing literature.

is larger when we allow for employment type heterogeneity in the Dogit framework. This finding is especially interesting in the context of the somewhat inconclusive findings in the existing literature regarding the relationship between self-employment and educational attainment. In addition, our findings suggest that it is important to distinguish between types of educational attainment rather than simply including the years of education.

The estimated θ parameters, which account for employment type heterogeneity within the Dogit framework, are presented in Table 3: θ_{nilf} denotes the estimated θ parameter for the 'not in the labour force' category; θ_{unemp} that for the unemployed category; θ_{se} that for the self-employed category; and finally, θ_{emp} denotes the estimated θ parameter for the employed category. It is apparent that the sizes of the estimated θ parameters are monotonically increasing as we move from being not in the labour force to being an employee. The zero coefficient on θ_{nilf} may reflect the heterogeneity within this group ranging, for example, from those who are long term sick to the retired.[7] Hence, the absence of an influence from labour force status heterogeneity is perhaps not surprising in this case. The final column presents the extent of the captivity effect as determined by equation (2). It apparent that demand-side influences appear to be the most pronounced in the employee category. The captivity effect for the self-employment category is the second largest effect indicating that employment type heterogeneity is potentially important for this group of individuals.

Due to the requirement that in the Dogit model $\theta_j \geq 0, \forall j$, a simple Likelihood ratio (LR) test of the Dogit model versus the MNL model is not appropriate due to the one-sided nature of the alternative hypothesis. However, it is straightforward to calculate information criteria based on the maximised log-likelihood functions: for all of the Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and CAIC (corrected AIC) measures

(see, for example, Cameron and Trivedi, 1998, p. 183) the Dogit model dominates and would thus be the preferred model on statistical grounds. This is further evidenced by the individual significance of the captivity parameters (and also that, ignoring the one-sided nature of the alternative hypothesis, the LR test clearly rejects the MNL model).

IV Extensions

In order to explore the robustness of the empirical results presented in Tables 2 and 3 above, we explore two extensions. Firstly, we distinguish between two different types of self-employment, own account status (*i.e.* self-employees who do not employ others) and employer status (*i.e.* self-employees who do employ others). Secondly, we explore a parameterised version of the Dogit model.

IV.1 Own Account and Employer Status Self-employment

Recent literature has explored the division within self-employment between own-account and employer status (see Earle and Sakova, 2000). Hence, we also analyse a model with five types of employment status by distinguishing between these two types of self-employee: *i.e.* those who do not employ others and those who do employ others. Out of those individuals reporting self-employment, 19% have no employees. Our empirical results are presented in Table 4. For brevity, we only again report the marginal effects. It is apparent from Table 4 that, in general, the marginal effects follow the same pattern as those in Table 2 with a more pronounced effect in the case of employer status. There are some differences however across the findings for own account and employer status: for example, number of children, household size and unearned income are all inversely associated with own account status yet positively associated with employer status self-employment. Turning to education, having a high school diploma is inversely associated with employer status (as in Table 2), yet statistically insignificantly related to own account status, whereas having a college degree is positively associated with own account status yet negatively associated with employer status. As in Table 2, the

[7] Given the zero coefficient on θ_{nilf}, this is then set to zero *a priori* and the model re-estimated.

remaining levels of educational attainment are positively associated with both types of self-employment.

It is apparent that the patterns of the marginal effects from the MNL and Dogit models are generally consistent in terms of sign and size. Focusing on education, both models suggest that having a high school diploma has a statistically insignificant effect on the probability of being an own account self-employee and an inverse association with employer status self-employment. Having a college degree is positively (inversely) associated with own account status (employer status), with this positive association being slightly more pronounced in the Dogit model relative to the MNL model.

The estimated θ parameters, which account for employment type heterogeneity within the Dogit framework, are presented in Table 5, where θ_{se}^{own} denotes the estimated θ parameter for the own account self-employment category and θ_{se}^{emp} represents that for employer status self-employment. It is apparent that as in Table 3, the sizes of the estimated θ parameters are monotonically increasing as we move from being not in the labour force to being an employee, with the estimated captivity effect being largest for employer status rather than own account status self-employment.

IV.2 The Parameterised Dogit Model

As an extension to the Dogit framework described above, we also analyse a "parameterised" version of the Dogit model whereby we parameterise the θ s. In a model of labour market status, a relatively standard set of observed individual characteristics is likely to directly affect the (second-stage) utilities of the individual via the index functions described in equations (4) and (5). As argued above, however, demand-side influences and labour market rigidities may lead to individuals being restricted (or "captive") to particular types of employment or conversely excluded from particular types of employment. The question arises as to whether these captivity effects are constant across individuals? That is, whether the heterogeneity of the various types of labour market states that exist will vary in its effect across individuals?

To explore this possibility, we allow the captivity parameters, θ, to vary by observed factors **z** which potentially characterize employment type heterogeneity, such that

$$\theta_j = \exp(\boldsymbol{z}'\gamma_j), \tag{9}$$

where the use of the exponential transformation ensures non-negativity of the θ parameters, required for the probabilities of equation (6) to be properly defined (Gaudry and Dagenais, 1979).[8] Such a generalization appears to be particularly appropriate when modeling labour market status, as it is potentially possible to identify status specific factors that are likely to influence an individual's observed labour market status or employment type. Furthermore, including demand-side variables as standard regressors does not take into consideration their true impact in terms of tying individuals to particular types of employment. This could lead to misspecification and erroneous inference. To further explore the contribution of modeling employment type heterogeneity to our understanding of who are the self-employed, we analyse a parameterised Dogit model.

In the context of this extremely flexible econometric framework, we allow the explanatory variables described above to either influence the probability via the $\boldsymbol{x}$-vector, via $\boldsymbol{z}$-vector (*i.e.,* via the θ parameters) or via both. For purposes of comparison, we allow all the covariates in Table 2 to influence the probability of self-employment (and all other employment types, for that matter) via the $\boldsymbol{x}$ -vector. We then allow human capital (*i.e.,* education and training), ethnicity, risk attitudes and the year dummy variables, to also influence the probability of self-employment via the captivity parameters.

Our chosen set of factors for the θ parameters reflects a range of issues raised in the existing literature on self-employment. As discussed above, the relationship between education and self-employment remains inconclusive in the existing literature: one line of argument takes education as a measure of ability and predicts a positive relationship between

[8] It should be noted that the proposed econometric framework is sufficiently flexible to allow overlap between the *z*'s and the *x*'s.

self-employment and education, whilst alternatively, self-employment may be regarded as a means to escape unemployment for the less educated. Given the conflicting findings in the existing literature, it appears fruitful to explore the role of education within this flexible econometric framework.

With respect to ethnicity, as suggested in the existing literature, employer based discrimination may push ethnic minorities into self-employment and thus employment types may be heterogeneous in terms of ethnic groups. However, it is important to note that, in this respect, this is potentially a demand-side factor which may impact on the incidence of self-employment among ethnic minority groups. The year dummy variables potentially capture any demand-side influences, which may affect opportunities for self-employees due to, for example, changes in the macro economy. Finally, attitudes towards risk have been a recurring theme in the theoretical literature, where those individuals who choose and succeed in self-employment have been found to be relatively less risk averse (see, for example, Van Praag and Cramer, 2001).

One omission in the empirical analysis so far relates to the argument that an individual may become self-employed once he/she has accumulated the necessary financial resources. Detailed analyses of the importance of capital constraints for the probability of becoming self-employed have been conducted by Blanchflower and Oswald (1998), Dunn and Holtz-Eakin (2000) and Johansson (2000). Given that capital constraints may be regarded as a demand-side influence, we also control for the effect of financial resources by including natural logarithm of household wealth in the vector $\boldsymbol{z}$.[9]

Table 6 presents the estimated coefficients for the self-employment category from the parameterised Dogit model, whilst Table 7 presents the associated marginal effects where the first column presents the overall marginal effect (the combination of the influences via $\boldsymbol{x}$ and

[9] Our measure of wealth includes: value of land, buildings, farms or ranches owned by the respondent; value of home, holiday houses or other properties; value of owned cars and other vehicles; financial assets; net of mortgages and loans.

z), the second and the third columns break down the overall marginal effect into the portion determined by $\boldsymbol{x}$ and the portion determined by z, respectively. Note that the LR test clearly rejects the simple Dogit model in favour of its parameterised counterpart; moreover again with regard to all of the AIC, BIC and CAIC metrics, the parameterised Dogit model would be statistically our preferred model.

It is apparent from Table 6 that the estimated coefficients of the θ parameters are, with the exception of training and two of the year controls, statistically significant at the 1% level. The high degree of statistical significance of the θ parameters indicates their important role in explaining employment type heterogeneity. We focus our discussion on the marginal effects presented in Table 7. Turning initially to the explanatory variables which only operate via $\boldsymbol{x}$, it is apparent that being male is positively related to self-employment across the MNL, Dogit and parameterised Dogit estimates, with the largest effect from gender being found within the parameterised Dogit framework. In contrast, the size of the age effect is broadly in line across the three econometric specifications. Similarly, the effects relating to the other household influences are broadly consistent across the three models.

With respect to ethnicity, which is allowed to influence the probability of self-employment via $\boldsymbol{x}$ and z, it is apparent that the inverse relationship between self-employment and ethnicity is more pronounced within the parameterised Dogit framework, with a moderate yet statistically significant employment type heterogeneity effect operating in the opposite direction to the relatively large inverse individual effect.

Turning to education, the inverse relationship between high school diploma and the probability of self-employment as indicated by the overall marginal effect is more pronounced in the parameterised Dogit results than the results presented in Table 2. Once again, the inverse individual heterogeneity effect outweighs a positive and statistically significant employment type heterogeneity effect. The findings related to having a college degree are particularly interesting since the MNL and Dogit frameworks both indicate an inverse

relationship between having a college degree and self-employment. Within the parameterised Dogit context, however, an overall positive effect is found with the inverse effect from employment type heterogeneity being outweighed in magnitude by a positive individual heterogeneity effect. This pattern is also followed in the case of the three levels of educational attainment above a college degree, with in each case the overall influence on self-employment being much more pronounced within the parameterised Dogit framework. This is also the case with the influence of attitudes towards risk on the probability of self-employment. With respect to wealth, a positive effect on the probability of self-employment is found relating to employment type heterogeneity which is consistent with the argument that capital constraints may act as a barrier to self-employment.

Finally, we summarise the predicted θ parameters from the parameterised Dogit results in Table 8, as well as the extent of the captivity effect (evaluated at sample means of all covariates). It is apparent that although the captivity effect for self-employment prevails within this framework, it is somewhat outweighed in magnitude by that of the unemployed and employed categories. It is interesting to note that once we allow employment type heterogeneity to vary by observed characteristics, all of these captivity effects become significantly larger, with the one exception of the self-employed (with captive probabilities of 0.019 compared to 0.020, respectively). Overall, our findings accord with those in the existing literature and, in addition, suggest that the effect of some influences, such as education, on the probability of self-employment may have been under-stated in the existing literature.

VI. Final Comments

The importance of the self-employment sector of the labour force has become apparent over recent times, with self-employees comprising a significant percentage of the labour force in many countries. Since self-employees create their own jobs, as well as potentially creating jobs for others, self-employment has been regarded as an important means to stimulate economic growth and reduce unemployment and poverty. Therefore, there has been keen

interest amongst both academics and policy makers in predicting what type of individual is likely to become self-employed. In this paper, we have presented the findings of an extremely flexible econometric model, which as yet has not been applied to modelling the incidence of self-employment, which importantly allows us to separate the influences of individual and employment type heterogeneity. Our findings suggest that employment type heterogeneity is potentially important in modelling self-employment and, furthermore, that the influences of certain individual characteristics, such as education, may have been underestimated in the previous empirical literature on self-employment, which may account for the somewhat inconclusive results relating to the relationship between educational attainment and self-employment reported in the existing literature.

References

Blanchflower, D., and A. Oswald (1998) 'What makes an entrepreneur?' *Journal of Labor Economics*, 16, 26–60.

Borjas, G. I. (1986) 'The Self-Employment Experience of Immigrants.' *The Journal of Human Resources,* 21,485-506.

Borjas, G. I. and S. G. Bronars (1989) 'Consumer Discrimination and Self-Employment.' *Journal of Political Economy*, 97,581-605.

Brock, W. A. and D. S. Evans (1989) 'The Economics of Small Business: Their Role and Regulation in the U.S. Economy.' *New York: Holmes and Meier.*

Brown, S., M. N. Harris and T. R. L. Fry (2008) 'Untangling Supply and Demand in Occupational Choice.' *Economics Letters*, 99, 2, 414-417.

Calvo, G. and S. Wellisz (1980) 'Technology, Entrepreneurs and Firm Size.' *Quarterly Journal of Economics*, 85, 663-678.

Cameron, C. and P. Trivedi (1998) Regression Analysis of Count Data, Econometric Society Monographs, Cambridge University Press, Cambridge, U.K.

De Wit, G. and V. F. A. A. M. Winden (1989) 'An Empirical Analysis of Self-Employment in the Netherlands.' *Small Business Economics*, 1, 263-272.

Dunn, T. and D. Holtz-Eakin (2000) 'Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links.' *Journal of Labor Economics*, 18(2), 282-305.

Earle, J. S. and Z. Sakova (2000) 'Business Start-ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies.' *Labour Economics*, 7(5), 575-601.

Evans, D. S. and L. S. Leighton (1989) 'Some Empirical Aspects of Entrepreneurship.' *American Economic Review*, 79, 519-535.

Evans, M. D. R. (1989) 'Immigrant Entrepreneurship: Effects of Ethnic Market Size and Isolated Labor Pool.' *American Sociological Review*, 54,950-962.

Fry, T., R. L. Brooks, B. Comley and J. Zhang (1993) 'Economic Motivations for Limited Dependent and Qualitative Variable Models.' *Economic Record*, 69,193-205.

Fry, T. R. L. and M. N. Harris (2005) 'The Dogit Ordered Generalized Extreme Value Model.' *Australian and New Zealand Journal of Statistics*, 47, 531-542.

Fujii, E. T. and C. B. Hawley (1991) 'Empirical Aspects of Self-Employment.' *Economics Letters*, 36, 323-329.

Gaudry, M. and M. Dagenais (1979) 'The Dogit Model.' *Transportation Research- Part B*, 13B, 105-112.

Hout, M., and H. Rosen (2002) 'Self-Employment, Family Background and Race.' *Journal of Human Resources,* 35(4), 670-692.

Johansson, E. (2000) 'Self-Employment and Liquidity Constraints: Evidence from Finland.' *Scandinavian Journal of Economics*, 102(1), 123-134.

Kidd, M. P (1993) 'Immigrants Wage Differentials and the Role of Self-Employment in Australia.' *Australian Economic Papers*, 32, 92-115.

Knight, F. H. (1921) Risk, Uncertainty and Profit, *New York: Houghton Mifflin*.

Le, A-T. (1999) 'Empirical Studies of Self-Employment.' *Journal of Economic Surveys*, 13(4), 381-416.

Maddala, G.S. (1983) 'Limited Dependent and Qualitative Variables in Econometrics', Econometric Society Monographs Number 3, Cambridge University Press.

Manski, C. (1977) 'The Structure of Random Utility Models.' *Theory and Decision*, 8, 229-254.

Parker, S. C. (2004) 'The Economics of Self-Employment and Entrepreneurship', Cambridge University Press.

Rees, H. and A. Shah (1986) 'An Empirical Analysis of Self-Employment in the UK.' *Journal of Applied Econometrics,* 1, 95-108.

Shaw, K. L. (1996) 'An Empirical Analysis of Risk Aversion and Income Growth.' *Journal of Labor Economics*, 14(4), 626-53.

Taylor, R. (1996) 'Earnings, Independence or Unemployment: Why Become Self-Employed?' *Oxford Bulleting of Economic and Statistics*, 30, 194-204.

Van Praag, C. M. and J. S. Cramer (2001) 'The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion.' *Economica*, 68, 45-62.

Table 1: Summary Statistics

Variable	Mean	Standard Deviation
Male	0.7851	0.4107
Age	50.244	16.3050
Non White	0.1931	0.3947
Married	0.6235	0.4845
Separated/Divorced	0.0300	0.1706
Widowed	0.1125	0.3160
Number of Children	1.5509	1.8507
Household Size	2.6563	1.4358
Log Unearned Income	7.8795	4.7065
1989	0.1277	0.3337
1992	0.1584	0.3651
1995	0.1746	0.3796
1998	0.1750	0.3799
2001	0.1806	0.3847
2004	0.1837	0.3872
No Education	0.2444	0.4297
High school Diploma	0.3133	0.4638
College	0.0438	0.2046
Bachelors Degree	0.2104	0.4076
Masters Degree	0.1015	0.3019
PhD	0.0867	0.2814
Training	0.1490	0.3561
Risk Preference	0.9574	0.8663
Log Total wealth	4.958	5.034

Table 2: Modelling the Incidence of Self-Employment

	Multinomial Logit Model		Dogit Model	
Variable	Coefficients (S. E.)	M.E. (S.E)	Coefficients (S. E.)	M.E. (S.E)
Male	1.3420	0.1977	1.5040	0.2092
	(0.0409)	(0.0059)	(0.0498)	(0.0067)
Age	2.1260	0.2558	2.7850	0.2594
	(0.0498)	(0.0069)	(0.0697)	(0.0810)
Age Squared	-0.2530	-0.0228	-0.3241	-0.0230
	(0.0047)	(0.0007)	(0.0070)	(0.0008)
Non White	-0.9120	-0.1332	-1.0040	-0.1397
	(0.0316)	(0.0044)	(0.0359)	(0.0048)
Married	0.4128	0.0642	0.4941	0.0674
	(0.0357)	(0.0046)	(0.0395)	(0.0049)
Separated/Divorced	-0.1852	0.0124	-0.1897	0.0153
	(0.0704)	(0.0107)	(0.0790)*	(0.0011)*
Widowed	-0.2343	0.0174	-0.3195	0.0097
	(0.0477)	(0.0076)	(0.0576)	(0.0086)*
Number of Children	-0.0304	-0.0001	-0.0199	0.0003
	(0.0062)	(0.0009)	(0.0070)	(0.0010)*
Household Size	0.1469	0.0156	0.1357	0.0164
	(0.0098)	(0.0012)	(0.0105)	(0.0012)
Log Unearned Income	-0.07317	0.0098	-0.0738	0.0090
	(0.0031)	(0.0004)	(0.0034)	(0.0003)
1992	0.1440	0.0546	0.1523	0.0562
	(0.0385)	(0.0050)	(0.0429)	(0.0052)
1995	-0.0082	0.0073	-0.01427	0.0077
	(0.0390)	(0.0050)	(0.0441)*	(0.0053)*
1998	-0.1637	-0.0077	-0.2029	-0.0110
	(0.0391)	(0.0050)	(0.0435)	(0.0053)*
2001	-0.1216	-0.0076	-0.1435	-0.0081
	(0.0388)	(0.0050)	(0.0432)	(0.0053)*
2004	0.0555	0.0015	0.0386	0.0006
	(0.0386)	(0.0050)	(0.0430)*	(0.0052)*
High school Diploma	-0.1257	-0.0129	-0.1445	-0.0150
	(0.0305)	(0.0043)	(0.0340)	(0.0045)
College	0.3238	-0.0179	0.2542	-0.0246
	(0.0601)	(0.0075)	(0.0660)	(0.0079)
Bachelors Degree	0.8778	0.0708	0.9323	0.0741
	(0.0321)	(0.0041)	(0.0354)	(0.0043)
Masters Degree	0.6925	0.0364	0.7288	0.0378
	(0.0391)	(0.0049)	(0.0423)	(0.0051)
PhD	1.6010	0.1898	1.6710	0.2038
	(0.0428)	(0.0051)	(0.0472)	(0.0006)
Training	-0.0506	-0.0319	-0.0685	-0.0343
	(0.0331)	(0.0047)	(0.0372)*	(0.0050)
Risk Preference	0.6256	0.0804	0.6643	0.0840
	(0.0133)	(0.0017)	(0.0148)	(0.0065)
Constant	-5.2380	-1.1410	-6.9220	-1.1500
	(0.1396)	(0.0173)	(0.1917)	(0.0192)
Log Likelihood		-1.020E+005		-1.071E+005

Note: * denotes that the variable is not statistically significant at the 1% level

Table 3: The Estimated θ Parameters

Incidence of Self-Employment

Estimated θ	Coefficient	Standard Error	'Captivity' Effect
θ_{nilf}	0.0000	0.0000	0.0000
θ_{unemp}	0.0014	0.0005	0.0013
θ_{se}	0.0207	0.0017	0.0192
θ_{emp}	0.0537	0.0032	0.0499

Table 4: Modelling the Incidence of Self-Employment: Own Account & Employer Status

	Multinomial Logit Model		Dogit Model	
	Own Account Status	Employer Status	Own Account Status	Employer Status
Variable	M.E. (S.E)	M.E. (S.E)	M.E. (S.E)	M.E. (S.E)
Male	0.0201 (0.0028)	0.1856 (0.0055)	0.0195 (0.0029)	0.1984 (0.0063)
Age	0.0519 (0.0036)	0.1994 (0.0059)	0.0629(0.0041)	0.1911(0.0068)
Age Squared	-0.0046 (0.0003)	-0.0177 (0.0006)	-0.0060(0.0004)	-0.0166 (0.0007)
Non White	-0.0398 (0.0025)	-0.0896 (0.0037)	-0.0422(0.0027)	-0.0938 (0.0041)
Married	0.0066 (0.0024)	0.0565 (0.0039)	0.0082(0.0024)	0.0582 (0.0042)
Separated/Divorced	0.0095 (0.0048)*	0.0014 (0.0098)*	0.0095(0.0049)*	0.0019 (0.0109)*
Widowed	-0.0053 (0.0038)*	0.0267 (0.0067)	- 0.0074(0.0041)*	0.0216 (0.0076)
Number of Children	-0.0044 (0.0005)	0.0032 (0.0007)	-0.0040(0.0006)	0.0031 (0.0010)
Household Size	-0.0023 (0.0006)	0.0165 (0.0009)	-0.0028(0.0007)	0.0177 (0.0010)
Log Unearned Income	-0.0013 (00002)	0.0107 (0.0003)	-0.0015(0.0002)	0.0103 (0.0003)
1992	0.0135 (0.0027)	0.0392 (0.0041)	0.0130(0.0027)	0.0417 (0.0043)
1995	0.0061 (0.0028)	0.0012 (0.0042)*	0.0062(0.0028)	0.0016 (0.0043)*
1998	-0.0010 (0.0028)*	-0.0067 (0.0042)*	- 0.0017(0.0028)*	-0.0092 (0.0043)
2001	0.0024 (0.0028)*	-0.0101 (0.0041)	0.0025(0.0027)*	-0.0108 (0.0043)
2004	0.0028 (0.0028)*	-0.0021 (0.0041)*	0.0021(0.0028)*	-0.0024 (0.0043)*
High school Diploma	-0.0022 (0.0023)*	-0.0101 (0.0036)	- 0.0027(0.0024)*	-0.0115 (0.0038)
College	0.0110 (0.0037)	-0.0294 (0.0065)	0.0097(0.0036)	-0.0328 (0.0070)
Bachelors Degree	0.0147 (0.0023)	0.0522 (0.0034)	0.0156(0.0023)	0.0554 (0.0036)
Masters Degree	0.0132 (0.0027)	0.0212 (0.0041)	0.0126(0.0027)	0.0231 (0.0042)
PhD	0.0400 (0.0027)	0.1407 (0.0042)	0.0419(0.0029)	0.1545 (0.0050)
Training	0.0050 (0.0025)*	0.0356 (0.0042)	0.0041(0.0025)*	-0.0375 (0.0043)
Risk Preference	0.0133 (0.0001)	0.0634 (0.0014)	0.0134(0.0001)	0.0670 (0.0016)
Constant	-0.2388 (0.0025)	-0.9877 (0.0146)	-0.2453 (0.0094)	-0.9792 (0.0159)
Log Likelihood		-1.161E+005		-1.158E+005

Table 5: The Estimated θ Parameters

Own Account versus Employer Status

Estimated θ	Coefficient	Standard Error	'Captivity' Effect
θ_{nilf}	0.0000	0.0000	0.0000
θ_{unemp}	0.0014	0.0005	0.0013
θ_{se}^{own}	0.0062	0.0008	0.0058
θ_{se}^{emp}	0.0121	0.0013	0.0113
θ_{emp}	0.0540	0.0032	0.0503

Table 6: Modelling the Incidence of Self-Employment: Parameterised Dogit Model

Variable	Coefficients	Standard Errors
Individual Heterogeneity		
Male	1.3800	0.0455
Age	1.7910	0.0586
Age Squared	-0.2304	0.0053
Non White	-0.9233	0.0365
Married	0.3721	0.0401
Separated/Divorced	-0.2342	0.0791
Widowed	-0.1179	0.0531
Number of Children	-0.0304	0.0068
Household Size	0.1322	0.0114
Log Unearned Income	-0.0813	0.0035
1992	0.1805	0.0456
1995	-0.0460	0.0444*
1998	-0.2472	0.0445
2001	-0.1920	0.0443
2004	0.0544	0.0447*
High school Diploma	-0.3950	0.0345
College	0.4023	0.0800
Bachelors Degree	0.8429	0.0380
Masters Degree	0.6851	0.0469
PhD	1.4330	0.0484
Training	-0.0227	0.0350*
Risk Preference	0.6330	0.0152
Constant	-3.7610	0.1703
Employment Type Heterogeneity		
Non White	-1.3790	0.2062
1992	-0.0903	0.1780*
1995	-0.6682	0.1814
1998	-0.6628	0.1955
2001	-0.6451	0.2079
2004	-0.2286	0.2045*
High school Diploma	-14.9900	0.6552
College	1.7020	0.3661
Bachelors Degree	3.0890	0.2306
Masters Degree	3.1620	0.2495
PhD	3.6970	0.2469
Training	0.3841	0.3333*
Risk Preference	0.3354	0.0714
Log Wealth	-1.0480	0.0197
Constant	0.3125	0.3035*
Log Likelihood		-1.002E+005

Note: * denotes that the variable is not statistically significant at the 1% level.

Table 7: Parameterised Dogit Model: Marginal Effects			
Variable	**Overall Marginal Effect**	***x***	***z***
Male	0.2815	0.2815	-
	(0.0090)	(0.0090)	
Age	0.2599	0.2599	-
	(0.0111)	(0.0111)	
Age Squared	-0.0246	-0.0246	-
	(0.0011)	(0.0011)	
Non White	-0.1966	-0.1972	0.0006
	(0.0071)	(0.0071)	(0.0002)
Married	0.0599	0.0599	-
	(0.0072)	(0.0072)	
Separated/Divorced	0.0123	0.0123	-
	(0.0157)*	(0.0157)*	
Widowed	0.0154	0.0154	-
	(0.0111)*	(0.0111)*	
Number of Children	-0.0021	-0.0021	-
	(0.0014)	(0.0014)*	
Household Size	0.0174	0.0174	-
	(0.0019)	(0.0019)	
Log Unearned Income	0.0071	0.0071	-
	(0.0005)	(0.0005)	
1992	0.0651	0.0645	0.0006
	(0.0086)	(0.0087)	(0.0002)
1995	-0.0057	-0.0063	0.0006
	(0.0086)*	(0.0086)*	(0.0002)
1998	-0.0323	-0.0328	0.0005
	(0.0086)	(0.0087)	(0.0003)*
2001	-0.0295	-0.02974	0.0003
	(0.0086)	(0.0086)	(0.0002)*
2004	-0.0101	-0.0101	0.0000
	(0.0087)	(0.0087)*	(0.0001)*
High school Diploma	-0.0406	-0.0607	0.0201
	(0.0071)	(0.0067)	(0.0031)
College	0.0923	0.0956	-0.0033
	(0.0200)	(0.0203)	(0.0007)
Bachelors Degree	0.2598	0.2640	-0.0041
	(0.0108)	(0.0109)	(0.0008)
Masters Degree	0.2380	0.2425	-0.0046
	(0.0147)	(0.0148)	(0.0008)
PhD	0.5008	0.5049	-0.0042
	(0.0231)	(0.0232)	(0.0008)
Training	-0.03491	-0.0342	-0.0007
	(0.0064)	(0.0064)	(0.0003)
Risk Preference	0.1220	0.1225	-0.0005
	(0.0028)	(0.0028)	(0.0001)
Log Wealth	0.0010	-	0.0010
	(0.0002)		(0.0002)
Constant	-1.1390	-1.1320	-0.0066
	(0.0294)	(0.0294)	(0.0012)

Note: * denotes that the variable is not statistically significant at the 1% level.

Table 8: Predicted θ Parameters		
θ	Estimated θ	'Captivity' Effect
θ_{nilf}	0.0366	0.0144
θ_{unemp}	0.4048	0.1591
θ_{se}	0.0512	0.0201
θ_{emp}	1.0516	0.4133