Who are the Self-employed? A New Approach

Sarah Brown[a], Lisa Farrell[b] and Mark N. Harris[c,d]

[a]Department of Economics, University of Leicester, United Kingdom.

[b]Department of Economics, University of Melbourne, Australia.

[c]Department of Econometrics and Business Statistics,
Monash University, Australia.

[d]Central European University, Budapest, Hungary.

December 18, 2002

Abstract

Modelling the incidence of self-employment has traditionally proved problematic. Whilst the individual supply side characteristics of the self-employed are well documented, the literature has largely neglected (or misspecified) demand side aspects. In this paper we present results from an econometric framework that allows us to separately, and simultaneously, model the supply and demand side characteristics that determine employment outcomes. We show that whilst individual characteristics are important determinants of the type of employment contract that individuals hold, there are also important contract specific factors that influence the nature of the contract an individual is employed under. Our results suggest that workers may be "captive" to a particular type of employment because of the sector in which they work, the number of hours they prefer to work and their ethnicity. The results are based on a new estimator, the parameterised DOGEV model, which allows for ordering and correlation in the observed alternatives, and for captivity within an observed alternative.

Keywords: Self-employment, captivity
JEL classification: J23 J33 C25 C10
Acknowledgements: We acknowledge financial support from the University of Melbourne and the University of Leicester. We are grateful to the ESRC Data Archive for supplying the Family Expenditure Surveys, 1997 to 2000. We wish to thank seminar participants at Monash University, Tim Fry and Mike Shields for helpful comments.

1 Introduction

Recent years have seen the self-employed emerge as an important group of workers in many developed countries, accounting for sixteen percent of the workforce in Australia, ten percent in Canada, nine per cent in the UK and nine percent in the US in 1999 (Le 1999). The increasing proportion of self-employed individuals in the labour force has led to an interest amongst economists and policy makers in the characteristics of the self-employed population. Yet modelling the incidence of self-employment has traditionally proved problematic. Whilst the individual supply side characteristics of the self-employed are well documented, the literature to date, has largely neglected (or misspecified) the demand side aspects that are important in determining self-employment. In this paper we present results from an econometric framework that allows us to separately, and simultaneously, model the supply and demand side characteristics that determine employment outcomes.

The literature concerning the profile of the self-employed has, in the main, concentrated on looking at the characteristics of individuals who are self-employed compared to those employed as fixed wage workers. We extend this approach by additionally considering workers employed on profit related pay (PRP) contracts. Such contracts allow some proportion of remuneration to be conditional upon a measure of performance. These have become much more commonplace in recent years and represent an interesting intermediate contract between the earnings certainty of fixed wage employment, and the more uncertain earnings associated with self-employment.

If individuals were identical in terms of their ability, tastes and prefer-

ences, and if there were no demand side rigidities in the labour market, in the absence of capital constraints, one would anticipate a pooling equilibrium with all workers flocking to one of the three contacts (self-employment, PRP or fixed wage employment).[1] In reality, there is a spectrum of ability, tastes and preferences, labour market rigidities (such as hours constraints) do exist and workers are faced with capital constraints, which together imply a separating equilibrium, with the expected utility of employment across each of the three types of employment being equalized. This paper focusses on the factors that determine the type of individual who is self-employed. We investigate the possibility that individuals may not be strictly free to choose their preferred type of employment, but may be "captive" to particular contract types due to a number of demand side factors which characterise the heterogeneity of employment contracts. Such factors here are: the nature of the employment sector; the length of working week the contract implies; and the individual's ethnic origin. There are obvious welfare considerations and associated policy issues if certain types of individual are being pushed into risky self-employment due to labour market rigidities or discrimination, rather than entering self-employment due to personal tastes and preferences.

The novelty of our approach is therefore twofold. Firstly, in contrast to the existing research which assumes that those not in self-employment form a homogeneous population, we set our analysis within a wider framework by focusing on a range of employment contract types which are explored collectively rather than in isolation.

[1] In what follows, we use the expressions 'type of employment' and 'type of employment contract' interchangably.

Secondly, we apply a new estimator, the Parameterised DOGEV model. This model allows for ordering and correlation in the observed alternatives, and for captivity within an observed alternative. Captivity to particular employment types is likely due to labour market rigidities and may be characterised by heterogeneity of the contract type. That is, we allow the choice alternatives themselves to have observable heterogeneity which may potentially capture individuals within a certain type of employment, separately from the effects of observable individual characteristics. Thus we are able to determine the impact of supply side factors on the probability of the individual being employed under a given contract type whilst controlling for demand side effects.

We argue that the employment types are necessarily ordered by the extent of earnings uncertainty. Fixed wage employment, for example, implies relatively stable pay. In contrast, a self-employed worker would expect to experience greater earnings fluctuations. PRP contracts, comprising an element of both fixed and variable pay, provide an intermediate contract between these two extremes. In what follows, we presume that self-employment is relatively more risky than PRP, which is itself relatively more risky than fixed wage employment. Our statistical model allows us to investigate, and test, this hypothesis.

The plan of this paper is as follows. Section 2 reviews the literature to date on the determinants of self-employment. Section 3 introduces our statistical framework and presents the Parameterised DOGEV model. Section 4 outlines the data used in the analysis. Section 5 discusses the findings from the model and Section 6 summarises and presents some final comments.

2 Background

Recent years have heralded a resurgence of interest amongst both academics and policy makers in the determinants of self-employment and its role as a potential solution to unemployment and poverty (especially in times of economic downturns). A number of approaches have been developed to explain the supply and demand of self-employment, emphasizing to varying degrees sociological, psychological and economic factors.

The basic economic argument is that individuals decide whether or not to enter self-employment on the basis of the relative utilities on offer. Such an approach encapsulates unemployment push and pull factors, with displaced workers being pushed or pulled into self-employment by supply side considerations (Taylor 1996). Relative returns, however, are but one part of the story. It has long been recognized that returns to self-employment are intrinsically riskier than the returns to salaried employment. An interesting issue is what type of individual is attracted into self-employment?

The most recent research has focused on the attributes of the self-employed in order to address the question "who are the self-employed?" focusing on characteristics such as gender, ethnicity and family background (Le 1999). In general, studies specify a reduced form Probit or Logit equation of self-employment whereby the vector of explanatory variables contains a combination of personal and labour market characteristics.

Turning initially to personal characteristics, marital status has been incorporated into many empirical studies. As argued by Le (1999), marriage is assumed in the economics literature, to represent stability and, as such,

may provide a suitable background for "risky" self-employment. Moreover, Blanchflower and Oswald (1990) and Bernhardt (1994) find that having a working spouse enhances the probability of self-employment - this may include financial stability. Similarly, Schiller and Crewson (1997) find evidence of intra-couple risk pooling with a husband's primary employment increasing the probability that a wife will be observed in self-employment. Related factors include the presence of children – individuals with young children, for example, may be less likely to bear to risk associated with self-employment.

An individual's age may also affect his/her propensity to become self-employed via a number of different channels (see, Calvo and Wellisz 1980, Kidd 1993). For instance, age 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. Alternatively, as an individual becomes older, he/she may have accumulated the financial resources required for self-employment – hence age may capture effects related to financial, as well as human, capital. 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). In general, studies have reported a non-linear relationship between self-employment and age (Rees and Shah 1986).

The role of ethnicity in determining the propensity to become self-employed has also attracted a great deal of attention in the literature. 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 U.K. have a lower propensity to become

self-employed whilst Brock and Evans (1986) find the reverse in the U.S.. More evidence reported by Clark and Drinkwater (2000), indicates higher rates of self-employment amongst ethnic minorities in England and Wales. In addition, it appears to be the case that ethnic enclaves provide a market for self-employed immigrants catering for individuals from the same ethnic background (Le 1999).

The role of educational qualifications has been incorporated into many empirical studies, being a key determinant of success, or otherwise, in the labour market. Educational attainment may act as a proxy for ability – individuals of higher ability may make better managers, which in turn may enhance their probability of becoming self-employed. 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 (1989) and Kidd (1993). In summary, the evidence regarding the relationship between education and the propensity to become self-employed remains inconclusive.

Another important labour market influence is local labour market conditions – if self-employment essentially provides a step out of unemployment, we might expect to see a higher rate of self-employment in areas characterised by relatively high unemployment. Evans and Leighton (1989), report

evidence supporting this view whilst Blanchflower and Oswald (1990) find inconclusive evidence. This effect may be evidence by variables pertaining to region of residence and industry/sector of employment, in that they capture the effects of local and general labour market conditions.

Finally, Blanchflower and Oswald (1998) and Blanchflower and Oswald (1990), argue that the psychological factors associated with entrepreneurial drive play a key role in the propensity to become self-employed. Evans and Leighton (1989) focus on one such psychological factor – willingness to be in charge of one's own destiny – and find that this is positively correlated with self-employment propensity.

A related point concerns an individual's preference for hours of work. In general, fixed wage employment entails accepting the working week dictated by the employer (although in some cases an employee may be able to exercise some discretion, albeit within the parameters specified by the employer, for example "flexi-time" schemes). In contrast, the self-employee can determine the number of hours of work, presumably to satisfy their entrepreneurial drive, thereby exerting more control over his/her destiny.

To summarise, the general approach adopted to profile the self-employed entails specifying a Probit or Logit model to predict the probability of self-employment relative to fixed wage employment, with a set of explanatory variables representing a mixture of labour supply (such as family status) and demand (such as sector of employment) influences. We extend the existing literature by allowing for three types of employment contract namely; fixed wage employment, performance related pay and self-employment. The inclusion of the intermediate category allows for the fact that, in reality, the choice

between self-employment and non self-employment, is not as stark as that depicted in the existing literature – intermediate contracts do exist which have characteristics of both fixed wage employment and self-employment. In addition, the existing literature does not distinguish between the supply side and demand side influences despite the fact that the labour demand influences may lead to contract specific effects, which should be modelled explicitly, and hence differently, to individual heterogeneity. In the next section, we suggest an appropriate econometric model that allows for such considerations.

3 The Econometric Model

3.1 The DOGEV Model

A new discrete model recently proposed by Fry and Harris (2002) forms the basis for the econometric analysis. Known as the DOGEV model, it is based upon separate, independent, generalisations of the Logit model for multiple outcomes. One of the generalisations embodies the OGEV probabilities (Small 1987), which allow for ordering in the data and correlations of alternatives in close vicinity. This is important as it is our contention that employment contract types can be naturally ranked according to the associated earnings uncertainty, from self-employment at one extreme to fixed wage employment at the other. Moreover, it is likely that neighbouring alternatives will indeed be correlated - using the same dataset as that used in this paper, Brown, Farrell, Harris, and Sessions (2002) find strong evidence of ordering across types of employment.

The second extension which the DOGEV model embodies, is the Dogit

model of Gaudry and Dagenais (1979). This expands on Logit probabilities by the introduction of additional choice-specific parameters, θ_j. The interpretation of these parameters varies across application, but in general they can be regarded as heterogeneity of the alternative itself, as opposed to observed individual heterogeneity. In addition, they may also represent unobserved individual heterogeneity which is common to individuals within a chosen alternative. For example, if the majority of individuals choosing a particular employment type, work a similar number of hours and the number of hours worked is a missing variable in the data set, the θ parameter for this particular outcome will embody the effect of this variable. Once more, such an extension appears very appropriate with regard to modelling types of employment contract. For example, it is quite likely that certain individuals will be captive to particular types of employment given the known labour market rigidities (for example, the possibility of discrimination).

It is intuitive in this context to consider the choice-set generation set-up of Manski (1977). Specifically, in the Dogit model an individual is assumed to be either captive to one of the J outcomes (here employment contract 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 from which "free choice" is (subsequently) exercised by the individual. The choice set generation process itself can be represented as a random utility maximisation model with utilities given by

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

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

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

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

$$P_{iJ+1} = \frac{1}{1 + \sum_{k=1}^{J} \theta_k}. \tag{3}$$

For the outcome selection process, the probability that an individual chooses the specified outcome j from a single outcome choice set is one. 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 maximisation model, RUM (Fry, Brooks, Comley, and Zhang 1993), of

$$U_{ij}^{(2)} = V_{ij} + \varepsilon_{ij} \tag{4}$$

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 this utility. For simplicity, V_{ij} is typically specified as

$$V_{ij} = \mathbf{x}_i' \boldsymbol{\beta}_j, \tag{5}$$

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

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

where P_{ij}^{L} are the simple Logit probabilities for multiple outcomes.

Fry and Harris (2002) suggest utilising this basic set-up, but to replace the Logit probabilities with those of the OGEV model, which are given by

$$\begin{aligned} P_{ij}^{OGEV} = & \frac{\exp\left(\rho^{-1} V_{ij}\right)}{\sum_{r=1}^{J+1} \left(\exp\left(\rho^{-1} V_{i,r-1}\right) + \exp\left(\rho^{-1} V_{ir}\right)\right)^{\rho}} \\ & \times \\ & \left[\left(\exp\left(\rho^{-1} V_{i,j-1}\right) + \exp\left(\rho^{-1} V_{ij}\right)\right)^{\rho-1} \right. \\ & \left. + \left(\exp\left(\rho^{-1} V_{ij}\right) + \exp\left(\rho^{-1} V_{i,j+1}\right)\right)^{\rho-1} \right], \end{aligned} \tag{7}$$

where the parameter ρ accounts for any ordering/correlation in the data.

Akin to a moving average process, the OGEV (and hence DOGEV) model(s) account for ordering in the data by allowing a correlation between alternatives in close proximity, which decreases the further are the two alternatives apart. Although ρ is not strictly a correlation coefficient - which has no closed form expression in these models (Small 1987) - it is inversely related to this such that when $\rho = 1$, the OGEV (DOGEV) model collapses to the Logit (Dogit) one.

Therefore the full functional form for the DOGEV choice probabilities are given by

$$P_{ij}^{DOGEV} = \frac{\theta_j}{1 + \sum_{k=1}^{J} \theta_k} + \frac{1}{1 + \sum_{k=1}^{J} \theta_k} \times P_{ij}^{OGEV} \tag{8}$$

This specification is convenient in that simple parameter restriction tests of $\rho = 1$ and $\boldsymbol{\theta} = \mathbf{0}$, are tests against the nested sub-models of Dogit and

OGEV (and jointly, Logit), respectively. Such parameter restrictions can be easily tested for using usual testing paradigms.

Using the indicator function d_{ij} where

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

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

$$L(\boldsymbol{\phi}) = \sum_{j=1}^{J} \sum_{i=1}^{N} d_{ij} \ln P_{ij}^{DOGEV}.$$

with $\boldsymbol{\phi}' = \left[\left(\text{vec}\boldsymbol{\beta}_j \right)', \boldsymbol{\theta}', \rho \right]$ and P_{ij}^{DOGEV} being given by equation (8).

3.2 The Parameterised DOGEV Model

In a model of employment contracts, a relatively standard set of observed individual characteristics are likely to directly affect the (second-stage) utilities of the individual via the index functions described in equations (5) and (4). Labour market rigidities, and in particular demand side constraints, instigate the likely captivity of individuals to particular types of employment, as stated above. However, an important question, is are these captivity effects constant across individuals? In other words, is there heterogeneity of the various types of employment, that will vary in its effect across individuals? For example, consider hours worked. Certain individuals are likely to have a penchant for a particular length of working week. However, such divergences in hours worked are likely to be specific to particular types of employment and predominantly controlled (or set) by the employers - for example fixed wage contracts are more likely to be associated with the "standard" working week.

Therefore, to allow for this, here we follow Harris, Fry, and Webster (2002) by considering an important extension to the basic DOGEV model. The so-called "Parameterised" DOGEV model, allows the captivity parameters $\boldsymbol{\theta}$, to vary by observed factors $\mathbf{z}$ which characterize employment contract heterogeneity (or demand side effects), such that

$$\theta_j = \exp\left(\mathbf{z}'\boldsymbol{\gamma}_j\right), \tag{9}$$

where the use of the exponential transformation ensures non-negativity of the $\boldsymbol{\theta}$ parameters, required for the probabilities of equation (8) to be properly defined (Gaudry and Dagenais 1979). Such a generalisation appears appropriate in a model explaining observed types of employment contracts, as it is possible to identify *contract specific factors* that are likely to impact on the type of employment contract an individual holds. We return to these in Section 5.

4 The Data

Our data is drawn from the Family Expenditure Survey (FES) for the UK, which is a nationally representative survey that has been conducted on an annual basis since 1957. Some 10,000 households are selected each year and the average response rate is approximately 70%. The main aim of the survey is to provide a reliable source of information on household expenditure, income and other aspects of household finances.

We use pooled data from the 1997/98, 1998/99 and 1999/00 surveys. Our sub-sample comprises of working adults aged between 18 and 65 who are either self-employed, employed under a fixed wage contract or on a con-

Table 1: Distribution of the Sample by Employment Contract

Self-employed		PRP		Fixed Wage		Total	
Number	%	Number	%	Number	%	Number	%
1287	9.4%	3,623	26.5%	8,765	64.1%	13,675	100%

tract characterized by a bonus scheme thereby introducing an element of PRP. The bonus schemes include: Christmas bonuses; productivity bonuses; profit related bonuses; loyalty bonuses; dividends; incentive schemes and performance/sales bonuses. Clearly the nature of these bonuses is somewhat diverse and as such may create different incentive mechanisms in the workplace. For simplicity we will treat those employed on PRP contracts as a homogeneous group.

Sample statistics are set out in Tables 1 - 3. It is apparent from Table 1 that fixed wage contracts are the dominant form of employment (64.1%) and self-employment (9.41%) is the least common.[2] Table 2 shows there are interesting differences in the incidence of employment type across individual characteristics. For example, men are relatively more likely to hold employment contracts associated with earnings uncertainty - self-employment. The incidence of self-employment is higher amongst married relative to single, separated, divorced and widowed, respondents. This may reflect a form of risk pooling with married people being attracted to self-employment because they can offset some income risk with other household members (for a full discussion of these issues see Brown, Farrell, and Sessions (2001) and Schiller and Crewson (1997)). With respect to ethnicity, PRP contracts are more

[2]The 9.4% self-employment figure ties in closely with that stated earlier of 9% for the U.K. suggesting that we have a representative working sample of the working population.

heavily concentrated among whites. Whilst the incidence of self-employment is low amongst blacks, the incidence of self-employment is relatively high among Asians, tying in with the findings of (Borooah and Hart 1999).[3]

One might also hypothesize that both the number and the age of children will affect a parent's willingness to become self-employed. We therefore consider the number of pre-school and school-age children in the household and find that whilst the former are distributed relatively evenly across the three types of employment, the latter are highest amongst self-employed workers. This may reflect the fact that self-employed workers are, on average, older than workers on PRP or fixed wage contracts.

The age profile of individuals employed on PRP contracts appears to be "$n-$shaped". This may be due to that fact that such contracts have been more widely introduced over the last decade and, as such, may be reflecting a cohort, rather than an age, effect. The age profile of individuals on fixed wage contracts is skewed towards the youngest age group (*i.e.,* those aged less than twenty), suggesting that the earnings uncertainty associated with PRP and self-employment may be prohibitively high for individuals with little labour market experience. In addition, they are less likely to have acquired the necessary financial capital to start a business. The proportion of individuals in self-employment increases with age, a finding that is consistent with the hypothesis that older, displaced workers might turn to self-employment given their relatively low probability of re-employment. It might also reflect the ability of older workers to acquire the necessary start-up capital for, and to

[3]Small effective sample sizes, prohibit separate identification of all ethnic minority effects in the econometric analysis.

Table 2: Contract Type by Individual Characteristics

	Self-employed	PRP	Fixed Wage
Gender			
Males	11.77	28.19	60.04
Females	4.67	23.07	72.26
Family Status			
Single	4.15	26.18	69.67
Married/cohabit	10.28	26.79	62.39
Separated/widowed/divorced	9.44	24.50	66.06
Pre-school Children*	0.22	0.23	0.18
School-Age Children*	0.75	0.47	0.52
Ethnicity			
White	9.28	26.78	63.94
Black	6.04	19.23	74.73
Asian	15.17	20.40	64.43
Age			
17<Age<19	1.74	15.12	83.14
20<Age<29	3.55	28.70	67.75
30<Age<39	8.84	29.98	61.18
40<Age<49	11.92	24.34	63.74
Age>50	14.17	23.39	63.44
Education Level			
Less than GCSE	14.06	22.36	63.58
GCSE	9.29	26.94	63.77
Further Education	7.38	28.48	64.14
Higher Education	7.78	27.28	64.94
Housing Tenure			
Rented- local authority	6.11	22.17	71.72
Rented - private	9.82	22.16	68.02
Mortgaged	9.02	28.44	62.54
Owned	13.89	22.20	68.92
Hours of work*	50.77	39.67	39.92

Note: Numbers are expressed as a percentage of the total number of individuals within the three types of employment for each individual characteristic, except for those denoted by * which represent the average for each types of employment

better absorb the income uncertainty associated with, self-employment, on account of their longer accumulation of wealth.

Individuals in self-employment have relatively few formal qualifications. PRP contracts, on the other hand, are concentrated amongst those with formal school qualifications and higher, whilst individuals holding fixed wage contracts are relatively evenly spread across all levels of schooling. It would appear that whilst educational attainment plays an important role in explaining the probability of holding PRP contacts or being self-employed, it may not be so important in explaining why individuals hold fixed wage contracts.

In relation to housing tenure, fixed wage employees are most likely to be found living in local authority housing, whilst the self-employed are more likely to own their homes outright and PRP workers are most likely to live in homes that are mortgaged. This may reflect the fact that housing equity is often used as collateral to secure loans necessary to start up a small business.

Finally, in terms of employer determined aspects of employment contracts, we can see that self-employed individuals have a much longer working week than those employed under fixed wage or PRP contracts. This is consistent with the known labour market rigidities in terms of hours of work. Fixed wage and PRP employees are not free to determine the number of hours that they work. Hours of work are usually a contractually specified aspect of the employment contract, although overtime may be available. We observe fixed wage and PRP workers to have a standard working week of approximately 40 hours, whilst the self-employed are observed to have a much longer working week of around 50 hours (and also with a much higher standard deviation).

Table 3 shows the relationship between contract types and employment

Table 3: Contract Type by Employment Sector

	Fixed Wage	PRP	Self-employed	Total
	Sample sizes			
Retail	1917	1043	451	3411
Manufacturing	2990	1553	522	5065
Services/public	3858	1027	314	5199
Total	8765	3623	1287	13675
	Percentages			
Retail	56	31	13	100
Manufacturing	59	31	10	100
Services/public	74	20	6	100

sector - another supply side factor important in determining observed employment contract outcomes, namely the sector of employment. That is, employment in a given sector captures an element of employment contract heterogeneity. The table shows that fixed wage contracts are the dominant type of employment across all three (broad) sectors of the economy. However there are interesting differences regarding the frequency of PRP and self-employment across the sectors. Contracts other than fixed wages are very rare within the service sector and the self employed are more common in the retail and manufacturing sectors, both of which also have a high frequency of PRP.

5 The Results

The results are set out in Tables 4 and 5 below. For comparison purposes, we also report results from multinomial Logit estimations, along with those from the Parameterised DOGEV. As discussed in Section 2 the standard approach to modelling the determinants of self-employment is to undertake a Logit or

Probit analysis where variables such as the sector the individual works in are included as standard regressors.[4] We argue here that these are demand side factors which reflect employment contract heterogeneity and should be modelled explicitly, and separately, to individual heterogeneity. Contract heterogeneity implies that individuals may be captive to a particular contract type irrespective of their personal characteristics. The instruments we use to control for employment type heterogeneity are: the sector of employment; hours of work; and ethnic origin.

It is clear that these variables capture demand side aspects. That is, certain sectors are more suited to different types of employment contracts. Where large economies of scale and/or large sunk costs exist, we would not expect to find a high prevalence of self-employed workers. Similarly the services/public sector will logically be dominated by fixed wage contracts. Employment contracts will therefore be a function of the sector that the individual works in. Thus, in addition to observed personal characteristics, individuals employed in different sectors are likely to be captive, to varying extents, to particular employment types.

Hours of work are usually contractually specified and hours inflexibility is a well documented labour market rigidity. Within fairly narrow bands, employees cannot, generally, freely choose the number of hours they wish to work (although some flexibility is generated by the availability of overtime and flexi-time, but this too is employer driven and controlled). There is much evidence which suggests that self-employment requires working longer

[4] An Ordered Probit model is inappropriate here, as our descriptive statistics above clearly suggest the coefficient vector ought not to be restricted to be equivalent across outcomes.

hours than other types of employment. Thus employment types are heterogeneous in terms of the number of hours they imply. An individuals ability/willingness to supply labour will therefore make then captive to particular contract types. Those unable/unwilling to work long hours are less likely to be observed to be self-employed.[5]

Finally, as suggested in the literature, we argue that employer based discrimination may push ethnic minorities into self-employment and thus employment contracts may be heterogeneous in terms of ethnic groups. However, it is important to note that, in this respect, this is a clear *demand side* factor which will impact on the incidence of self-employment among ethnic minority groups.

Including demand side variables as standard regressors does not take into consideration their true impact in terms of tying workers to particular types of employment. We argue this leads to misspecification and potentially erroneous inference. In order to understand the contribution of modelling outcome heterogeneity makes to our understanding of *who are the self-employed?*, we present a comparison of the Parameterised DOGEV results with the standard multinomial Logit specification.

Prior to discussing the estimated coefficients on the explanatory variables we will present a comparison of the two models. Table 1 shows the sample proportions across types of employment and the within-sample predicted proportions from the multinomial Logit and the Parameterised DOGEV models.

[5]The data indicates that there is a larger variance of hours worked amongst the self-employed. However, for any self-employed individual this does not necessarily imply a high variance across average weekly hours worked. Therefore these data can inform us about an individual's preference for abolute hours, and not variance of such.

Table 4: Multinomial Logit Estimates

Reference Category: Fixed Wage Contracts

	PRP Contract			Self-Employed		
Variable	Coef			Coef		
Constant	-1.981	(0.32)	**	-10.310	(0.57)	**
Demographic						
Male	0.353	(0.05)	**	0.453	(0.09)	**
Age	0.915	(0.15)	**	1.388	(0.27)	**
Age squared	-0.122	(0.02)	**	-0.124	(0.03)	**
Married	0.088	(0.06)		0.236	(0.13)	*
Separated/widowed/divorced	0.133	(0.10)		0.217	(0.17)	
Pre school children	0.071	(0.04)		0.181	(0.07)	**
School age children	-0.147	(0.03)	**	0.156	(0.04)	**
Education						
GCSE	0.097	(0.07)		-0.061	(0.09)	
Further Education	0.238	(0.07)	**	-0.093	(0.11)	
Higher Education	0.131	(0.07)	*	-0.182	(0.11)	
Housing Tenure						
Rented private	-0.021	(0.10)		0.752	(0.18)	**
Mortgaged	0.284	(0.07)	**	0.575	(0.14)	**
Owned	0.124	(0.09)		0.956	(0.16)	**
Year of Survey						
1998/99	-0.263	(0.05)	**	-0.323	(0.08)	**
1997/98	-0.213	(0.05)	**	-0.218	(0.08)	**
Household Income	0.286	(0.04)	**	0.019	(0.07)	
Region						
Wales	-0.146	(0.11)		-0.043	(0.17)	
Scotland	-0.070	(0.13)		0.197	(0.20)	
North England	-0.021	(0.08)		-0.041	(0.13)	
Midlands	-0.168	(0.08)	**	-0.135	(0.13)	
South England	0.048	(0.05)		-0.250	(0.09)	**
Industry and hours of work						
Manufacturing	-0.085	(0.05)		-0.248	(0.08)	**
Services/public	-0.753	(0.05)	**	-0.953	(0.09)	**
Hours of work	-0.191	(0.04)	**	0.995	(0.03)	**
Non-white	-0.266	(0.11)	**	0.284	(0.16)	*
Number of Observations	13675					
Log Likelihood	-10500					

Standard errors in parentheses. ** and * significant at 5 and 10% level, respectively (two-sided).

Table 5: Parameterised DOGEV Estimates

	PRP Contract			Self-Employed		
Variable	Coef			Coef		
Constant	-3.360	(1.05)	**	-8.588	(1.83)	**
Male	0.675	(0.19)	**	0.406	(0.28)	
Age	1.276	(0.51)	**	2.841	(0.79)	**
Age squared	-0.184	(0.07)	**	-0.264	(0.09)	**
Married	0.153	(0.16)		1.024	(0.34)	**
Separated/widowed/divorced	0.092	(0.26)		0.930	(0.42)	**
Pre school children	-0.123	(0.07)		0.272	(0.17)	
School age children	0.437	(0.18)	*	0.237	(0.10)	**
Education						
GCSE	-0.323	(0.18)	*	-0.063	(0.22)	
Further Education	-0.025	(0.17)		-0.128	(0.28)	
Higher Education	-0.125	(0.17)		0.080	(0.28)	
Housing Tenure						
Rented private	-0.216	(0.27)		1.533	(0.47)	**
Mortgaged	0.506	(0.21)	**	1.781	(0.39)	**
Owned	0.469	(0.26)	*	2.502	(0.48)	**
Year of Survey						
1998/99	-0.413	(0.15)	**	-0.574	(0.20)	**
1997/98	-0.397	(0.15)	**	-0.363	(0.20)	*
Household Income	0.437	(0.18)	**	-4.719	(1.04)	**
Region						
Wales	-0.668	(0.34)	*	-0.881	(0.49)	
Scotland	-0.074	(0.30)		0.658	(0.54)	
North England	-0.151	(0.19)		-0.158	(0.27)	
Midlands	-0.433	(0.24)	*	-0.432	(0.31)	
South England	-0.007	(0.13)		-1.009	(0.31)	
Captivity Parameters						
Fixed Wage						
Constant	11.860	(0.75)	**			
Manufacturing	-0.040	(0.16)				
Services	0.387	(0.20)	**			
Hours of work	-2.720	(0.19)	**			
Non-white	0.354	(0.29)				
PRP						
Constant	14.120	(0.89)	**			
Manufacturing	-0.096	(0.17)				
Services	-0.473	(0.20)	**			
Hours of work	-3.456	(0.23)	**			
Non-white	-0.012	(0.30)				
Self-Employed						
Constant	-2.777	(0.36)	**			
Manufacturing	-0.187	(0.15)				
Services	-0.713	(0.17)	**			
Hours of work	0.348	(0.06)	**			
Non-white	0.706	(0.26)	**			
ρ	0.562	(0.27)	**			
Number of Observations	13675					
Log Likelihood	-10350					

Standard errors in parentheses. ** and * significant at 5 and 10% level, respectively (two-sided).

Table 6: Sample proportions and predicted outcomes

	Fixed wage	PRP	Self-employed	Correct
Sample	0.641	0.265	0.094	
Multinomial Logit	0.945	0.016	0.039	0.65
Parameterised DOGEV	0.938	0.011	0.051	0.66

We can see that both specifications over predict the dominant outcome, fixed wages. However, as was our contention, the Parameterised DOGEV model does much better at predicting outcomes within the self-employed category (although marginally worse in the PRP one).[6] Indeed, the poor ability of the multinomial Logit to predict outside of the dominant category is worrying. Clearly, explicitly accounting for employment contract heterogeneity greatly increases our ability to understand observed employment outcomes. In terms of predictive power the Parameterised DOGEV model is better, predicting 66% of observations correctly; compared to 65% for the multinomial Logit specification. However the predictive capacity of the models is misleading given the dominance of fixed wage contracts in the data - a naïve model predicting fixed wage employment for the entire sample would correctly predict 64% of the observations! A better criterion for judging the performance of the models is their ability to predict outcomes outside of the dominant category and the Parameterised DOGEV clearly does better than the multinomial Logit in this respect, especially with regard to self-employment.

To further consider the within-sample prediction accuracy of the Parameterised DOGEV model Table 7 presents a simulated hit and miss table. A

[6]Neither models predict the PRP category particularly well. One suspects that this is due to a paucity of variables pertaining to the workplace. Datasets containing such variables are workplace orientated and, by definition, will not include the self-employed. However, it is an important intermediate category as it allows, for instance, us to identify earnings' uncertainty.

common measure of predictive accuracy in discrete choice models is the "hit-miss" table. The predictions underlying hit-miss tables are given by assigning individuals the outcome associated with their highest predicted probability across the alternatives and then comparing this to the observed outcome. As is generally the case with discrete choice models (Duncan and Harris 2002) both of these models tended to over predict the most frequently chosen alternative (as noted above). This is because construction of traditional hit and miss tables implicitly ignores the stochastic elements of the underlying economic model. Therefore we present simulated hit and miss tables where we explicitly take into account both the stochastic and non-stochastic elements of the underlying utility function (over 1,000 random draws). For comparison purposes, the case of a random assignment of individuals types of employment according to observed sample proportions is also presented. The Parameterised DOGEV quite clearly predicts better than simple random assignment. So, it appears that accounting for potential captivity within a particular type of employment (and so modelling it explicitly as opposed to simply including these variables as standard regressors in a multinomial Logit framework), is particularly important in understanding why individuals are observed in particular types of employment.

Figure 1 further illustrates the predictive power of the Parameterised DOGEV model by graphing: the actual sample proportions by type of employment; the average predicted probability across the individuals for each outcome; total probabilities predicted at the sample means of the explanatory variables; and the amount of the latter accounted for by the captivity

Table 7: Hit and Miss Tables
RANDOM ASSIGNMENT

	Predicted			
Actual	Fixed wage	PRP	Self-employed	Total
Fixed wage	5618	2322	825	8765
PRP	2322	960	341	3623
Self-employed	825	341	121	1287
Total	8765	3623	1287	13675

PARAMETERISED DOGEV SIMULATED

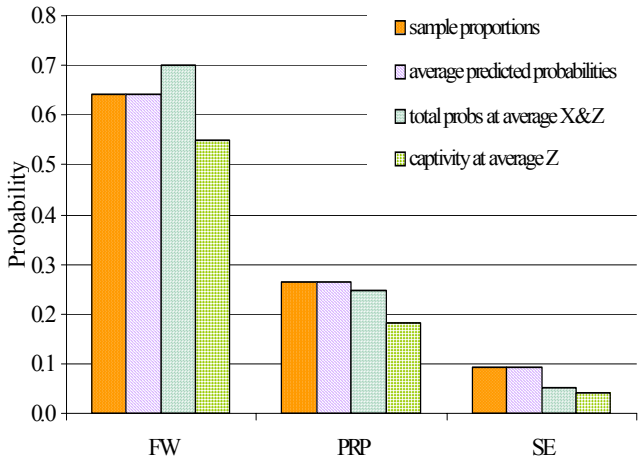
	Predicted			
Actual	Fixed wage	PRP	Self-employed	Total
Fixed wage	5837	2296	632	8765
PRP	2301	1069	253	3623
Self-employed	624	263	400	1287
Total	8762	3628	1285	13675

within each employment outcome.[7] Under these two tests of predictive power we can see the model clearly performs well. The average probability of being in each employment contract accurately matches the sample proportions. Whilst calculating the predicted probabilities at the sample means of the variables does slightly over predict the number of fixed wage employees and slightly under predict the number of PRP and self-employed workers, the orders of magnitude are small. Moreover, the diagram illustrates the substantial contribution made to the total probabilities arising from the demand side (or captivity) factors.

Another short fall of the multinomial Logit specification is the fact that it can't account for the proposed presence of ordering in the alternatives (according to earnings' uncertainty). The Parameterised DOGEV, on the

[7]This differs from Table 1 which is based on observed individual characteristics as opposed to sample means.

Figure 1: Sample Proportions; Predicted Probabilities - average and at sample means (by total and captive)



other hand, does. A simple test for the null hypothesis of no ordering, is undertaken as a test of $\rho = 1$. We clearly reject the null, confirming our hypothesis that employment contracts are ordered according to their degree of earnings uncertainty, and therefore it is important in modelling observed type of employment, to account for this. Note that ρ is statistically significantly different from 0 and 1 (at the 95% level). Ignoring the ordering in our data would therefore result in misspecification and erroneous potentially inferences. Note also, that this confirms previous results of Brown, Farrell, Harris, and Sessions (2002) and Fry and Harris (2002).

In terms of explanatory variables, a comparison of the multinomial Logit estimates against the Parameterised DOGEV estimates shows a fairly consistent story, although the significance of some variables varies across the specifications. One should, however, note the high degree of significance of

the captivity parameters in the Parameterised DOGEV model suggesting that they are controlling for important demand side factors and thus do capture employment contract heterogeneity. This is also reflected in the large contribution they make to the total probability of being in any employment class (as is highlighted by the diagrams above and below). A simple likelihood ratio test of the Parameterised DOGEV model against the DOGEV specification showed that the parameterisation of the captivity terms, according to the variables discussed above does make a significant contribution to the likelihood function. Hence we will concentrate the discussion of the results on the Parameterised DOGEV specification.

Given the complexity of the implicit marginal effects of the Parameterised DOGEV model and, for reasons of clarity, we present the implied probabilities for different realisations of our explanatory variables. Specifically, we will consider the impact of individual heterogeneity in terms of age and housing tenure (two variables that past studies have found significant in predicting the incidence of self-employment) and of contract heterogeneity in terms of employment sector, hours of work and ethnicity - setting all other variables to their sample means.

The diagrams have been constructed so that one can distinguish the contribution to the total probabilities generated by the individual and contract specific heterogeneity. Note that as we change $\mathbf{x}$ (individual heterogeneity variables) total probabilities change via equations (5) and (8), although the captivity probabilities are unaltered. On the other hand, when we change $\mathbf{z}$ (contract heterogeneity variables), captive probabilities are directly affected via equations (9) and (8), but also the non-captive ones via a scale factor

implied by these equations. That is, changes in $\mathbf{x}$ change the total probabilities through the impact on P_{ij}^{OGEV} and hence affect only the second term of equation (8). Whilst changes in $\mathbf{z}$ influence θ_j and so affect both terms in (8): the first term directly and the second term via a scale factor.

Regarding age, we find that individual heterogeneity plays almost no role in the probability of being self-employed when young, but becomes increasingly important for workers aged 40 plus. This is consistent with the literature that concludes that age is positively related to the propensity to become self-employed. This finding is consistent with any of the following hypotheses: i) older workers are "pushed" into self-employment as a result of a lack of other employment opportunities in later life, ii) older individuals may becomes self-employed once they have accumulated sufficient human capital or, iii) older individuals have acquired the necessary and financial capital to become self-employed (Le 1999).

Figure 2: Predicted Probability: Age

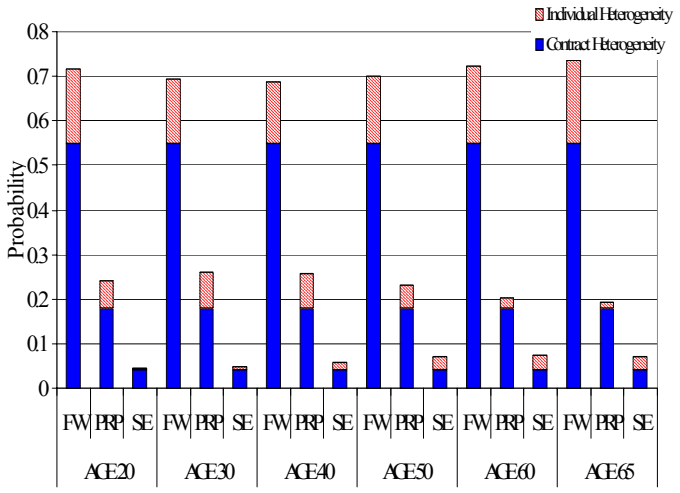
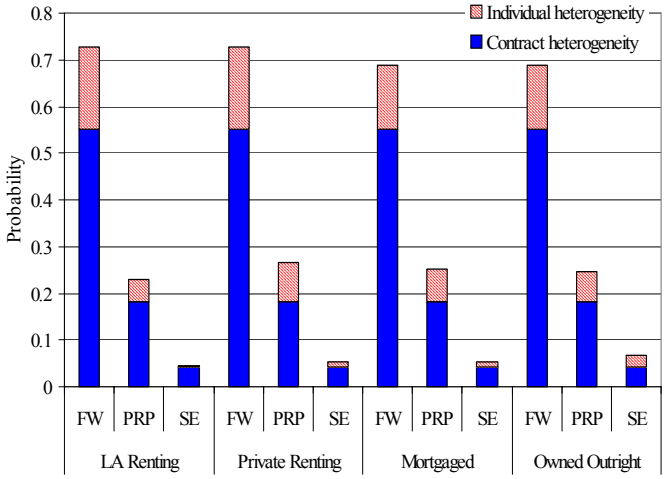


Figure 3: Predicted Probabilities: Housing Tenure

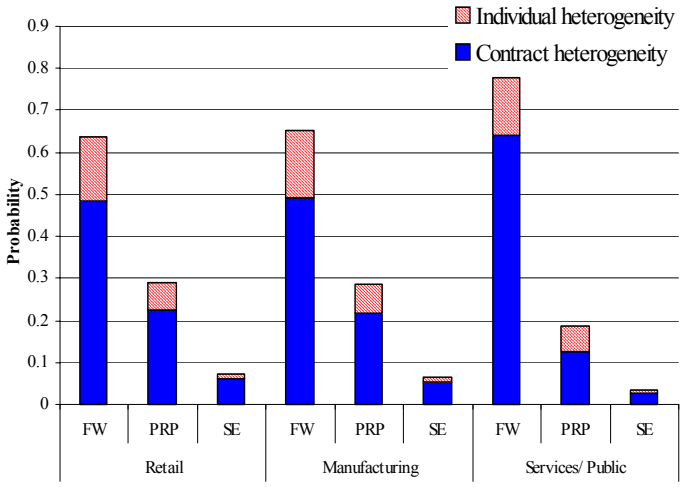


Secondly, we graph the predicted probabilities conditional on housing tenure (see Figure 3). We can see that these findings are consistent with the hypothesis that capital/asset accumulation allows entry into self-employment, given that housing equity is a common source of loans security. Home ownership is positively associated with self-employment, whilst renting is positively associated with fixed wage employment.[8]

Turning to the results regarding the impact of demand side effects in terms of their ability to capture workers within particular types of employment, we find that fixed wage workers are most likely to be found in the service/public sector and the increased probability is largely due to the captivity parameters. PRP and self-employment in contrast are least likely to be found in this sector (see Figure 4).

[8] *LA* renting in Figure 3 refers to Local Authority rented homes - government provided housing.

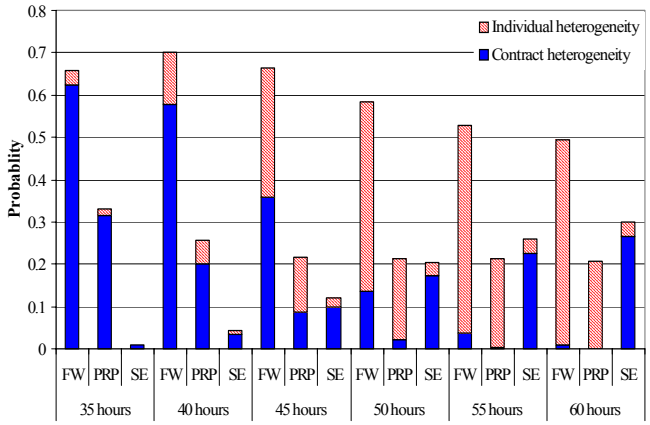
Figure 4: Predicted Probabilities: Sector



In relation to the number of hours worked, we find the largest degree of variation in the contribution of the individual and contract heterogeneity parameters across the range of possible hours worked and the types of contracts. This is illustrated in Figure 5.

Fixed wage workers are most likely to be observed working a standard working week (approximately 40 hours), PRP workers have a slightly shorter working week (35 hours per week) and the self-employed appear to have longer working hours (the probability of being self-employed increases as hours increases). Captivity variables dominate the probability of being employed on a fixed wage contract for low hours of work, however, their contribution to the total probability diminishes rapidly as the number of hours worked increases. The same pattern is observed for those employed under PRP contracts; whilst the converse is true for the self-employed. The self-employed are much more likely to be observed working long hours and

Figure 5: Predictied Probabilities: Hours

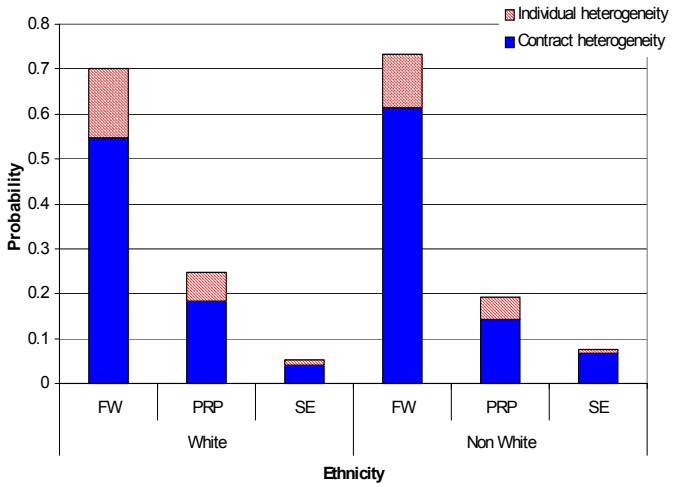


logically the contract heterogeneity parameters dominate this probability.

Figure 6 shows how employment type varies with ethnic origin. There have been a number of studies that have shown that ethnic minorities have a higher probability to be self-employed (see, Clark and Drinkwater 2000, Brock and Evans 1986), and this finding is reflected in our results. The diagram shows that non-whites are much more likely to be self-employed than the other ethnic groups, Borooah and Hart (1999) find similar evidence. In contrast, contract heterogeneity plays a large role in determining the probability of being a fixed wage worker for whites. The significance of ethnic origin as a captivity parameter is consistent with the hypothesis that labour market discrimination pushes non-whites into self-employment.

In summary, our results show that demand side factors are important determinants of the type of employment contract an individual holds. Whilst individual heterogeneity explains some of the story, demand side factors are

Figure 6: Predicted Probabilities: Ethnicity



highly influential. Previous literature has concentrated on explaining who the self-employed are in terms of individual heterogeneity, our results suggests that demand side factors (such as labour market rigidities and discrimination) also play an important role in explaining self-employment. Moreover, simply including demand side variables as standard regressors in a model of individual heterogeneity leads to poorer predictive power when compared to a specification that explicitly and simultaneously controls for the captivity impact of the demand side factors.

6 Conclusions

Economists have in general found it hard to predict who the self employed are based on models of individual heterogeneity. Our results suggest that this is not surprising. Whilst individual characteristics are important identifiers of employment contracts, aspects of contract specific heterogeneity also play a

very important role. We have shown that modelling outcome heterogeneity as individual heterogeneity leads to poor predictions when compared to a model that allows us to separate out the effects of individual and outcome heterogeneity. Moreover, our results suggest that the existing literature has drawn conclusions from potentially misspecified models.

We present a model where individual heterogeneity represents labour supply side characteristics and contract heterogeneity represents supply side characteristics. We argue that workers are captive to certain types of employment as a result of the sector in which they work, the number of hours that they work and their ethnic origin. We have shown that the self-employed are captive to self-employment due to their ethnicity, by some strong preference to work longer hours and as a result of the sector they work in. Moreover, we show that the probabilities of being in any particular employment type are heavily driven by supply side factors. Our results concerning individual heterogeneity are consistent with the existing literature in as much as the self-employed are more likely to be older and have housing equity. Additionally we have confirmed previous results in the literature confirming that types of employment are ranked in order of earnings uncertainty.

References

BERNHARDT, I. (1994): "Comparative Advantage in Self-Employment and Paid Work," *Canadian Journal of Economics*, 27, 273–289.

BLANCHFLOWER, D. G., AND A. J. OSWALD (1990): "What Makes a Young Entrepreneur?," *Centre for Labour Economics, London School of*

Economics, Discussion Paper Number 373.

——— (1998): "What Makes an Entrepreneur?," *Journal of Labor Economics*, 16(1), 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.

Borooah, V. K., and M. Hart (1999): "Factors Affecting Self-Employment Among Indian and Black Caribbean Men in Britain," *Small Business Economics*, 13(2), 111–129.

Brock, W. A., and D. S. Evans (1986): "The Economics of Small Business: Their Role and Regulation in the U.S. Economy," *New York: Holmes and Meier.*

Brown, S., L. Farrell, M. Harris, and J. Sessions (2002): "Risk Preference and Employment Contract Type," Discussion paper, Department of Economics, University of Melbourne, Australia.

Brown, S., L. Farrell, and J. Sessions (2001): "Employment Contract Matching: An Analysis of Dual Earner Couples and Working Households," *University of Leicester, Department of Economics, Working Paper*, 01/9.

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

Clark, K., and S. Drinkwater (2000): "Pushed Out or Pulled In? Self-Employment Among Ethnic Minorities in England and Wales," *Labour Economics*, September, 603–28.

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.

Duncan, A., and M. Harris (2002): "Simulating the Effect of Welfare Reforms Among Sole Parents in Australia," *The Economic Record*.

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.

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. Brooks, B. Comley, and J. Zhang (1993): "Economic Motivations for Limited Dependent and Qualitiative Variable Models," *Economic Record*, 69, 193–205.

Fry, T., and M. Harris (2002): "The Dogev Model," Discussion paper, 7/2002 Monash University, University of Melbourne, Australia.

GAUDRY, M., AND M. DAGENAIS (1979): "The Dogit Model," *Transportation Research - Part B*, 13B, 105–112.

HARRIS, M., T. FRY, AND E. WEBSTER (2002): "Occupational Choice and Demand Side Effects," *mimeo MIAESR*.

HOUT, M., AND H. ROSEN (2000): "Self-Employment, Family Background and Race," *Journal of Human Resources*, 35(4), 670–92.

JOHANSSON, E. (2000): "Self-Employment and Liquidity Constraints: Evidence from Finland," *Scandinavian Journal of Economics*, 102(1), 123–34.

KIDD, M. P. (1993): "Immigrant Wage Differentials and the Role of Self-Employment in Australia," *Australian Economic Papers*, 32, 92–115.

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*. Cambridge University Press, Cambridge, U.K.

MANSKI, C. (1977): "The Structure of Random Utility Models," *Theory and Decision*, 8, 229–254.

REES, H., AND A. SHAH (1986): "An Empirical Analysis of Self-Employment in the UK," *Journal of Applied Econometrics*, 1, 95–108.

SCHILLER, B., AND P. CREWSON (1997): "Entrepreneurial Origins: A Longitudinal Inquiry," *Economic Inquiry*, 25, 523–531.

SMALL, K. (1987): "A Discrete Choice Model for Ordered Alternatives," *Econometrica*, 55, 409–424.

TAYLOR, M. (1996): "Earnings, Independence or Unemployment: Why Become Self-Employed?," *Oxford Bulletin of Economics and Statistics*, 30, 194–204.