0203305

DOES THE ENCLAVE EFFECT HELP

EXPLAIN ETHNIC MINORITY SELF-

EMPLOYMENT IN THE UK?

Does The Enclave Effect Help Explain Ethnic Minority Self-Employment in the UK?

Abstract

This paper aims to offer a possible explanation for the disproportionate number of ethnic minorities in self-employment, through focusing on the enclave effect. The self-employment decision is influenced by ethnic-specific attributes, local labour market conditions as well as sectoral earnings differentials; all of which are incorporated into a probit model. The estimation of earnings equations for each sector enables a calculation of the self-employment/paid-employment earnings differential, which, when incorporated with information on ethnic density facilitates the prediction of the probability of an individual entering self-employment. This paper finds that contrary to popular belief, individuals residing in ethnic enclaves (areas with a high percentage of their own ethnic group) are less likely to become self-employed. As expected the earnings differential exerts a powerful positive influence on the self-employment decision, as do many of the human capital characteristics.

Word Count = 5328

I. Introduction

Self-employment is considered to be particularly important for ethnic minorities. It offers an alternative to the potential discrimination they face in paid-employment, in the form of lower wages. The disproportionate number of ethnic minorities in self-employment is seen as a rational response to such labour market conditions. Data from the 2001 Census shows that the self-employment rate for whites is 11.2%, but the self-employment rate of non-whites is 12.5%.

The focus of this analysis is based on the effect of residing in an ethnic enclave, on the probability of an individual from a particular ethnic group becoming self-employed. An enclave is described as a concentration of individuals from the same ethnic background who reside within a specific geographic area. I show whether the enclave argument is a significant factor in determining the destination of an individual into self-employment or paid-employment[1]. The ethnic density of the area an individual resides in is not the only contributing factor in determining an individual's choice of sector[2]; I also control for factors such as human capital variables, sectoral earnings differences and local labour market conditions. Thus allowing a more complete and thorough analysis.

The enclave effect can be either positive or negative. If positive; living in an area with many co-ethnics increases the likelihood of entering self-employment, and vice-versa for the negative effect. These arguments are based on very different hypotheses. It is usually argued that enclaves give rise to a protected market in which individuals from particular ethnic groups are able to trade with each other in their preferred language.

---

[1] If enclaves are important and if they affect the self employment decision, then it is expected that the proportion of an areas population accounted for by a particular ethnic group will influence the likelihood of an individual from that group to be in self employment.
[2] Throughout this paper when I refer to an individuals "choice of sector" I am referring to the choice between paid-employment and self-employment.

This then offers a comparative advantage in serving the needs of the ethnic group, as one would assume the self-employed individual as part of that ethnic group, would have a greater knowledge and understanding of the specific needs, tastes and preferences of customers. Such markets include foodstuffs and clothing. The ethnic enclave could offer a self-sustaining economic environment, thus, providing niche markets for such culture specific or ethnic goods. Such an economic environment could encourage individuals to enter self-employment. Yuengert (1995)[3] suggests that access to informal start-up capital may be easier for minorities, due to family and cultural ties within the community, thus encouraging ethnic entrepreneurship. Another factor which could explain the positive hypothesis is that in any specialised area (an ethnic enclave) there is a need for a certain number of businesses to cater for the specific cultural needs of its inhabitants. This could lead to numerous businesses being established to serve the specific ethnic group enclave, for example doctors surgeries, food & clothing stores, health stores, lawyers, accountants and travel agencies etc, therefore implying a greater ratio of ethnic specific businesses with respect to the enclave population[4]. This is in comparison with predominantly white areas.

Several counter arguments have been proposed. Aldrich & Waldinger (1990)[5] say that enclaves could spark too much competition amongst entrepreneurs in the supply of services in a fairly restricted and small market. Also, enclaves could also offer better opportunities for ethnic minorities in paid-employment, if ethnic employers offer non-discriminatory employment opportunities. This overcomes the problem of

---

[3] Yuengert, Andrew M. 1995. *Testing Hypothesis of Immigrant Self-Employment*, Journal of Human Resources 30(1):194-204
[4] Other businesses could also be required which need ethnic entrepreneurs to run them in order to communicate in the same language, and share the same values as others in the enclave. These could include specialist professional services firms.
[5] Aldrich, H & Waldinger, R (1990). *Ethnicity and entrepreneurship*. Annual Review of Sociology 16, 111-135.

4

discrimination from whites in the paid-employment sector. Aldrich & Waldinger (1990) also suggest that expansion into the wider market may be more difficult, thus constraining the potential for growth of ethnic businesses in enclaves. These factors could combine to reduce the likelihood of an individual in an ethnic enclave choosing self-employment over paid-employment.

These very different hypotheses make the potential finding of the analysis more interesting. In the next section, I provide a brief but informative overview of the literature, followed by an outline of the model and methods used in my analysis. I then provide a detailed analysis of the dataset, followed by a section on the results and finally the conclusions.

II. Literature Review

In reviewing the literature, I have focussed more broadly on the relationship between ethnic minorities and self-employment in order to identify key variables affecting the likelihood of self-employment. This is partly due to the lack of direct studies into the enclave effect, and also because many factors studied by other authors will form an important part of my analysis[6]. As stated earlier, there are a disproportionate number of self-employed ethnic minorities in the UK, with no agreed reason as to why. There is a broad range of literature attributing the relationship to many factors, ranging from year of immigration to religion[7]. The models and methods used are similar between papers, and in many cases involve the need to add extra variables to regressions to tailor to specific needs[8].

Clark & Drinkwater (2000)[9] provide a rigorous analysis of many factors believed to be the causes of high ethnic self-employment. Due to the range of factors considered, this paper provides a useful focus, around which, to compare the conclusions form across the literature. They use the 'Fourth National Survey of Ethnic Minorities', sampling a total of 5196 ethnic minorities (both male and female) aged 16 or over, as well as 2867 whites. The sample of ethnic minorities is much greater than in other papers using regular government surveys, providing a great advantage considering most conclusions are about the ethnic minority groups. They found substantial variation in self-employment rates across the six ethnic groups, highlighting the importance of treating them separately and

---

[6] This is due to the large number of other variables included in the analysis which can be significant determinants of the probability of entering self-employment.

[7] Much of the research focuses on the US labour market. Although useful in guiding my research, the conclusions they draw are not applicable to the UK, due to the labour markets being characterised by different factors.

[8] Where I refer to the "standard model", this will be explained in the methodology section.

[9] Clark K & Drinkwater S (2000). *Pushed out or pulled in? SE among EM in England and Wales.* Labour Economics, 7, 603-628

not collectively. Their analysis is based on a series of 'push' and 'pull' factors. 'Push factors' are believed to drive ethnic minorities into self-employment, and 'pull factors' drag individuals into self-employment.

PUSH FACTORS	PULL FACTORS
- Labour market discrimination	- Enclaves
	- Religion
	- Language
	- Immigrant status

Labour market discrimination comes in the form of lower wages for non-whites working in the same sector as whites. It reduces the opportunity cost of self-employment for non-whites, pushing them out of paid-employment and making self-employment a more attractive option. Blackaby et al (1998)[10] use decomposition analysis[11] to decompose earnings differences into productivity and discrimination components. They find non-whites, on average, earn 11% less in paid-employment than whites. Decomposition analysis can attribute only 5% of the difference to human capital differences; therefore there is scope for discrimination as a push factor. Many authors have found employer discrimination to be prevalent in the UK, but in the US, Fairlie & Meyer (1996)[12] conclude that the high returns from self-employment pull ethnic minorities into self-employment more than discrimination pushes them into it.

---

[10] Blackaby et al (1998). *The ethnic wage gap and employment differentials in the 1990s: Evidence for Britain.* Economic Letters, 58, 97-103

[11] Model is adapted from the pioneers of decomposition analysis: Blinder, A.S (1973), '*Wage Discrimination: Reduced Form and Structural Variables.*' Journal of Human Resources, 8, 436-455 and Oaxaca, R (1973), '*Male-Female Wage Differentials in Urban Labour Markets'* International Economic Review, 14, 693-709.

[12] Fairlie R & Meyer B (1996). *Ethnic and Racial SE Differences and possible explanations.* Journal of Human Resources31, 757-795

Pull factors considered by Clark & Drinkwater (2000) tend to be cultural values, the importance of which Rafiq (1992)[13] points out: "Culture is important in any discussion of entrepreneurship because it can determine the attitudes of individuals towards entrepreneurship". Religion is a factor that has acquired little attention, mainly because of its colinearity with ethnicity. But, Rafiq (1992), and Clark & Drinkwater (2000), found that entrepreneurship is more likely among devotees of the Sikh, Muslim and Hindu faiths, who seem to value entrepreneurship more than others.

The research regarding ethnic enclaves by Clark & Drinkwater (2000) finds little evidence to suggest enclaves help explain the high ethnic minority self-employment in the UK[14]. Borjas (1986)[15], using 1970 and 1980 US Censuses data adopts the standard model used in analysing the relationship between ethnic minorities and high self-employment rates, but, adds a variable for ethnic density to consider the enclave effect, and separately adds a vector of dummy variables for year since migration, to show the effect that immigration status has on self-employment propensity. He finds that residing in an enclave increases the probability of self-employment by 1%-2% for Hispanic, Mexican and Cuban immigrants over native born men[16]. Clark & Drinkwater (2000), and in the US, Fairlie & Meyer (2000), also find that individuals with low English fluency are less likely to be self-employed, but if they reside in an enclave where they can trade with people from the same ethnic group, then Evans (1989)[17] finds that self-employment tends to be higher.

---

[13] Rafiq, M (1992). *Ethnicity and enterprise: a comparison of Muslim and non-Muslim owned Asian businesses in Britain.* New Community 19, 43-60.
[14] Clark & Drinkwater attribute it to the relatively deprived nature of such areas.
[15] Borjas G (1986). *The SE experience of Immigrants.* Journal of Human Resources 21, 485-506
[16] In the same paper Borjas also shows that the probability of becoming self-employed increases with the length of time since migration. This is because initially the cost of entry into self-employment is likely to deter immigrants.
[17] Evans, MDR, 1989. *Immigrant Entrepreneurship: effects of ethnic market size and isolated labor market pool.* American Sociological Review 54, 950-962.

There are also other important factors analysed, which do not fall under push or pull factors. Dunn & Holtz-Eakin (1996)[18] examine the effect of parental self-employment and asset levels on self-employed propensity. They found a strong positive effect of parental self-employment on the probability of becoming self-employed but only a weak positive effect for parental wealth[19]. Interestingly, Lentz & La-Bard (1990)[20] found that 53% of their sample of self-employed people had self-employed parents.

In addition to employer discrimination, studies have considered consumer and lending discrimination, which may drive ethnic minorities out of self-employment. Borjas & Bronars (1989)[21] study consumer discrimination, and find that relative gains of entering self-employment are reduced for ethnic minorities because they have to compensate white consumers by lowering prices charged for goods/services. The model is based on finding the reservation price that makes a white individual indifferent between purchasing from a black seller and continuing to search. I believe this is an unrealistic model, because consumers are unlikely to know the race of the seller before purchasing a good/service nor are they likely to construct reservation prices.

Lending discrimination is studied by Coate & Tennyson (1992)[22], under the assumption that lenders are unable to observe entrepreneurial ability. Individuals from a group discriminated against in the labour market will receive less favourable terms in the credit market. This is because lenders know that for discriminated against individuals, the

---

[18] Dunn, T & Holtz-Eakin, D (1996). *Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links.* Syracuse University, 1996.

[19] Parental self-employment was included because parents may pass on the family business or informal business managerial experience. The reason for including parental wealth is that it may provide an access to start up-capital.

[20] Lentz, B & Laband, D (1990). Entrepreneurial Success and Occupational Inheritance among Proprietors. Canadian Journal of Economics 23, 563-579.

[21] Borjas, G & Bronars, S. *Consumer Discrimination and Self-employment*. Journal of Political Economy 97, 581-605.

[22] Coate, S & Tennyson, S. *Labour Market Discrimination, Imperfect Information and Self-Employment.* Oxford Economic Papers 44, 272-288.

opportunity cost of entering self-employment is lower, therefore they are willing to take more risks. Ethnic minority groups will be charged higher interest rates, thus reducing the expected returns from self-employment. This represents spillover effects from labour market discrimination.

The above presents a wide range of potential factors believed to affect an individual's sectoral choice. Many can be incorporated into the analysis, but others are constrained by data limitations. There are several problems with the literature I have tried to address in this paper. Firstly, many authors make the mistake of categorising all ethnic minorities as a single homogeneous group, which conceals important inter-group differences. There is also an issue with the under-reporting of self-employed incomes (due to tax advantages), but this is a problem I fear will be unavoidable. Another issue is with the size of the datasets used, with only Clark & Drinkwater (2000) really managing to attain a large sample size for ethnic minorities. They do so using a specialised dataset, one which has not been updated since 1993. Most authors use a grouped earnings variable, which hides important information, therefore making the analysis less accurate. Also, in analysing the enclave effect, authors tend to split the UK into 'north' and 'south', which I believe are too broad areas to categorise an enclave. I try to address this problem by using eleven different regions.

III. Methodology

For the purpose of the analysis I use the model which is used across the board by most authors. An individual has two choices to make. He/she first has to decide whether to be in the labour market. If they decide to be in the labour market they then need to decide whether to be in paid-employment or self-employment. I assume that the individual has already decided to participate in the labour market, leaving a choice of which sector to work in. The model itself begins with a formulation of the self-employment/paid-employment choice problem, drawing upon Knight's (1921)[23] notion that individuals respond to the risk adjusted relative earnings opportunities in each sector. Therefore for each individual, I predicted what their earnings would be if they worked in the other sector[24]. In order to calculate predicted earnings I used a standard Mincer wage equation of the following format:

Equation 1[25]

$$\log earnings^j_i = b^j_o + b^j_1 Age_i + b^j_2 Age^2 1000_i + b^j_3 Male_i + b^j_4 Married_i +$$
$$b^j_5 Disability_i + b^j_6 YrsEdu_i + b^j_7 Edlev\_otherqual_i + b^j_8 Edlev\_high\deg_i +$$
$$b^j_9 Edlev\_1st\deg_i + b^j_{10} Edlev\_profqual_i + b^j_{11} Edlev\_Alev_i + b^j_{12} Edlev\_GCSE_i +$$
$$b^j_{13} Ind\_cons_i + b^j_{14} Ind\_serv_i + b^j_{15} Ind\_na_i + b^j_{16} Ind_i + b^j_{17} Ethnic_i + b^j_{18} reg\_NW_i +$$
$$b^j_{19} reg\_YaH_i + b^j_{20} reg\_EM_i + b^j_{21} reg\_WM_i + b^j_{22} reg\_EE_i + b^j_{23} reg\_Lon_i +$$
$$b^j_{24} reg\_SE_i + b^j_{25} reg\_SW_i + b^j_{26} reg\_Wales_i + b^j_{27} reg\_S\cot_i + e_i$$

i = 1,…,n. j=SE/PE. Log earnings depend on a vector of accumulated human capital variables, as well as controls for industry, ethnicity and region. The earnings differential

---

[23] Knight , F.H (1921), *Risk, Uncertainty and Profit,* Houghton, Mifflin, New York
[24] For example, for a self-employed individual I calculate their predicted wage if they worked in paid-employment, and for paid-employed individuals I calculate their predicted wage if they were in self-employment.
[25] For a full description of variables used please refer to Appendix 1

is the difference between log predicted earnings ($\ln \overline{Y}$) in self-employment and paid-employment for each individual:

Equaiton.2

$$EarnDiff \quad = \ln \overline{Y}_{SE_i} - \ln \overline{Y}_{E_i} .$$

As well as the earnings differential, I also produced data on ethnic density. The ethnic density variable was produced by finding the ethnic breakdown of each region as a percentage of the total population of the region[26]. This provided a continuous value, with every individual having an ethnic density associated with them depending on their ethnic group and region of residence. The earnings differential and ethnic density form a crucial part of the probit equation below, predicting an individual's sectoral choice.

Equaiton.3 [27]

$$Z_i^* = a_0 + a_1 (\ln \overline{Y}_i^{SE} - \ln \overline{Y}_i^{PE}) + a_2 Age_i + a_3 Age^2 / 1000_i + a_4 Male_i +$$
$$a_5 Married_i + a_6 Disability_i + a_7 DepChil_i + a_8 UnempRate_i + a_9 GDP_i +$$
$$a_{10} YrsEdu_i + a_{11} Edlev\_otherqual_i + a_{12} Edlev\_high\deg_i +$$
$$a_{13} Edlev\_1st\deg_i + a_{14} Edlev\_profqual_i + a_{15} Edlev\_Alev_i + a_{16} Edlev\_GCSE_i +$$
$$a_{17} Ind\_cons_i + a_{18} Ind\_serv_i + a_{19} Ind\_na_i + a_{20} EthDensity_i + h_i$$

The value of Z* is an index of self-employment propensity. If it exceeds 0 it is assumed to equal 1, in which case the individual chooses self-employment. By running the probit model above, it is possible to compute marginal effects for each variable. It is the marginal effects which provide interpretable information, unlike the coefficients of the probit, which only provide information on the sign of the variable and its significance. They key marginal effect to look for is the EthDensity variable, which coupled with its

---

[26] The ethnic breakdown was taken from the 2001 Census, as this was the only reliable dataset available to gain such information matching the regions under consideration.
[27] For a full description of variables used please refer to Appendix 1

sign and significance will allow me to see whether the enclave effect is positive or negative and the degree to which it affects the self-employment decision.

IV. Data Section

The dataset used for the analysis is the General Household Survey (GHS)[28]. The GHS provides an extensive amount of information on approximately 13000 individuals each year, incorporating over 6000 variables. The sample is selected from the Postal Address File (PAF) which covers 97% of UK private households, providing a representative sample. The major advantage of the GHS is that it provides detailed income information, and has extensive detail in terms of topics covered and questions asked.

Due to the fact that ethnic minorities only represent a small proportion of the total UK population, the number of ethnic minorities in each year's dataset is small. To overcome this problem I have combined four datasets from 1998-1999, 2000-2001, 2001-2002 and 2002-2003[29].

After combining the datasets it was important to delete certain individuals and unnecessary variables. The analysis has been restricted to individuals between the age of 18 and 65, because these are the ages between which we would expect to find most data on self-employed individuals. Individuals considered economically inactive, unemployed or in full-time education are also deleted from the dataset. This has left a total of 34778[30] individuals, of which only 1029 are from an ethnic minority background (3% of sample). Ideally I would have liked this to be higher, but limitations in the dataset due to the re-development of the GHS made this difficult. From the surveys available, the GHS was

---

[28] The General Household Survey (GHS) is a continuous national survey of people living in private households conducted on an annual basis, by the Social Survey Division of the Office for National Statistics. The main aim of the survey is to collect data on a range of core topics, covering household, family and individual information. This information is used by government departments and other organizations for planning, policy and monitoring purposes, and to present a picture of households, family and people in Great Britain.

[29] The gap in 1999-2000 is because the survey was being re-developed.

[30] Of the sample, 4294 are classified as self employed (12.3%). A further breakdown is provided in tables 1 and 2.

the only one which provided extensive earnings data for both paid-employed and self-employed individuals. The Labour Force Survey (LFS) was an alternative, but it does not provide earnings information for self-employed individuals. The LFS would have provided a larger sample of ethnic minorities, because it is a quarterly dataset with a larger sample size.

Table.1

Self-employment rates as a percentage of those in employment

	White	Mixed	Other	Indian	Pakistani	Bangladeshi	Black	Chinese
n	32947	599	203	282	190	202	303	52
Total	12.12%	16.69%	16.70%	11.00%	28.42%	13.86%	6.61%	19.23%
Male	16.43%	23.68%	22.43%	13.41%	36.22%	16.39%	11.59%	12.50%
Female	7.40%	8.63%	7.63%	12.70%	10.00%	2.42%	25.00%	10.41%

Table.2

Self-employment rates as a percentage of those in self-employed (categorised)

	White	Total Ethnic
n	33749	1029
Total	12.30%	13.90%
Male	16.60%	18.61%
Female	7.44%	7.93%

Table 1 provides an in-depth breakdown of the percentage of individuals in self-employment for each ethnic group (for both males and females) and table 2 categorises the percentage of white's vs. non-whites in self-employment. The reason for table 2, is due to the fact that ethnic minorities represent such a small percentage of the sample that, in the Mincer wage equations it has become necessary to categorise the ethnic groups as white and non-whites. If they were left as in table 1 then the inferences made would not be very accurate or reliable.

15

Interestingly, from table 1, we see all the ethnic groups apart from Blacks and Indians have a higher proportion of individuals in self-employment than whites. This coincides with what others have found, apart from the Indian case. Other surveys have found Indians to have one of the highest proportions of individuals in self-employment[31]. Table 2 highlights that the pattern of self-employment among ethnic minorities (13.90%) is 1.7% higher than for whites (12.20%), which I would have expected to be higher[32]. Males tend to exhibit higher self-employment rates than females for both whites and non-whites. Figure.1 provides an informative graphical interpretation of the economic activity of each ethnic group. It shows the percentage of each group in paid-employment and self-employment.

Figure 1: Economic Activity categorised by ethnic group:

----
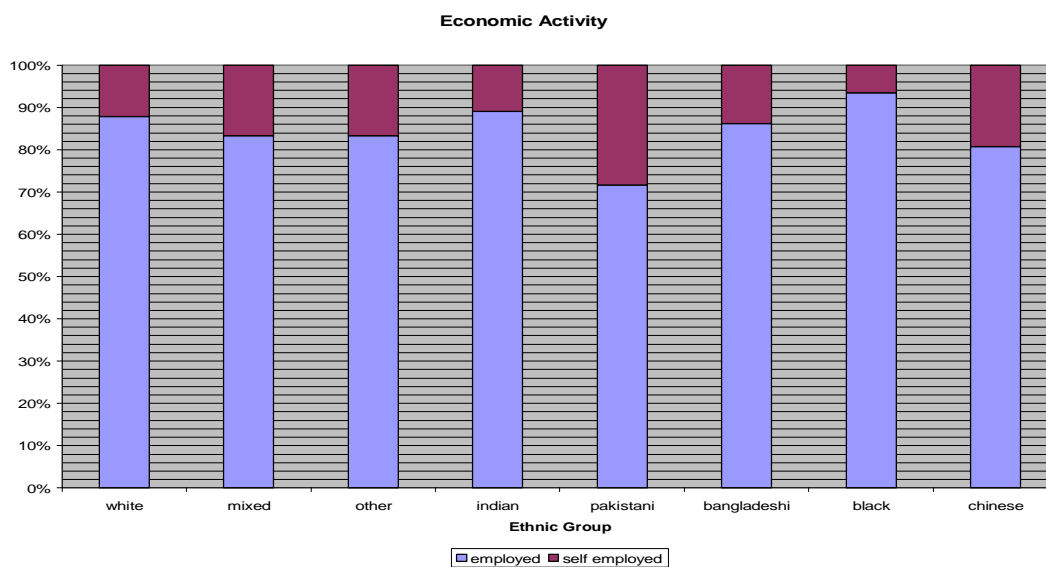
[31] Data from the ONS (Office of National Statistics) suggests my finding goes against what is true in reality. Therefore the discrepancy must be because of the small sample of Indians (ethnic minorities in general) in the survey. This suggests I am correct to treat ethnic minorities as a single homogenous group for the Mincer wage equation.

[32] Clark & Drinkwater (2000) found the difference was 5.5%, but accurate inferences are constrained by the size of the dataset, which could explain the discrepancies.

Descriptive statistics for all variables used are provided in appendix 2. The most striking element of the dataset comes from the earnings information. On average an individual in self-employment earns 30680p compared to 31836p a week for an individual in paid-employment. This represents a negative earnings differential of 1156p (paid-employed individuals earn on average 3.77% more than self-employed individuals), which is unexpected, as theory suggests individuals look to self-employment due to the higher potential gains. The discrepancy could be attributed to the renowned under reporting of self-employed earnings, mainly due to tax advantages. This general relationship of individuals earning more in paid-employment than self-employment, seems to apply for each of the ethnic groups, apart from for 'Other' and 'Bangladeshis', as depicted by figure.2[33]. Earnings information itself is very volatile; more so for self-employed individuals, as shown by the high standard deviations for variables 'grmainjb' and 'logwage' in appendix 2.

Figure.2 Average Gross Weekly Earnings for Each Ethnic Group and Sector

---
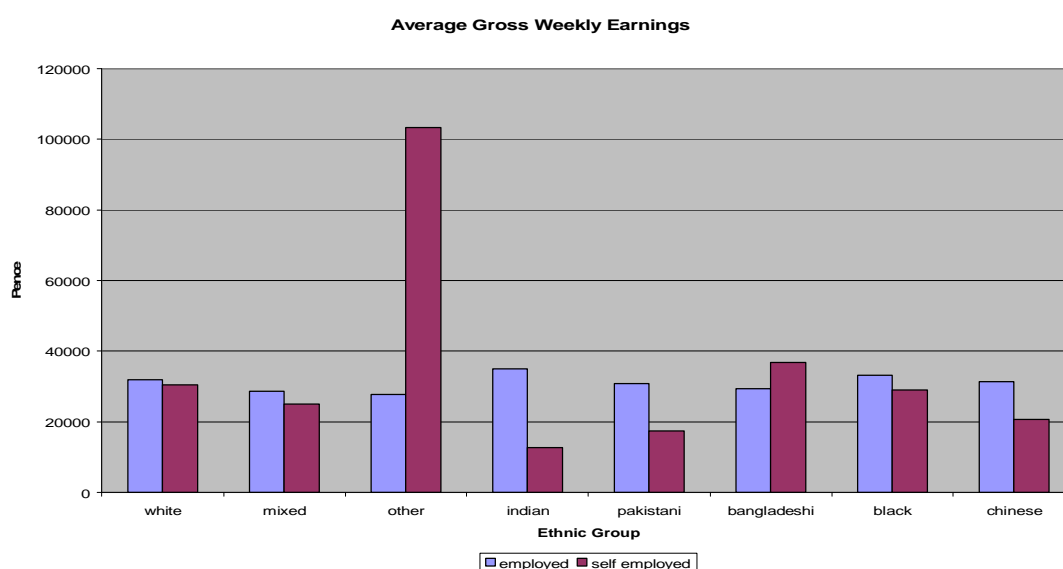
[33] The average self-employed earnings of the 'other' ethnic group seem to be wildly over predicted. I have found no outliers in the data which leaves me mystified as to why the average is so high relative to the rest of the ethnic groups.

Another significant feature is that the ethnic density of areas where self-employed individuals reside averages at 1.85% less than where paid-employed individuals reside. This obviously coincides with the negative, rather than positive, effect of enclaves. If instead we look only at ethnic minorities (i.e. excluding whites) the same effect becomes apparent[34]. Figure.3 is a graph showing the ethnic make-up (excluding whites) of each of the eleven regions in the sample. The main features of the graph are that London and the West Midlands have the largest proportions of ethnic minorities, which is in accordance with reality. The ethnic densities are fairly similar across the other regions, apart from the East Midlands, which has substantially more Indians, and Yorkshire & Humberside which has substantially more Pakistanis. Also, relative to the other ethnic groups, the Chinese and Bangladeshis have much smaller densities across most regions.

---

[34] Self-employed ethnic minorities live on average in areas with 2.58% of co-ethnics, and paid-employed individuals live on average in areas with 3.64% of co-ethnics.

Figure 3 [35] [36]

Figure 3: Ethnic Density by Region (Excluding Whites)

---
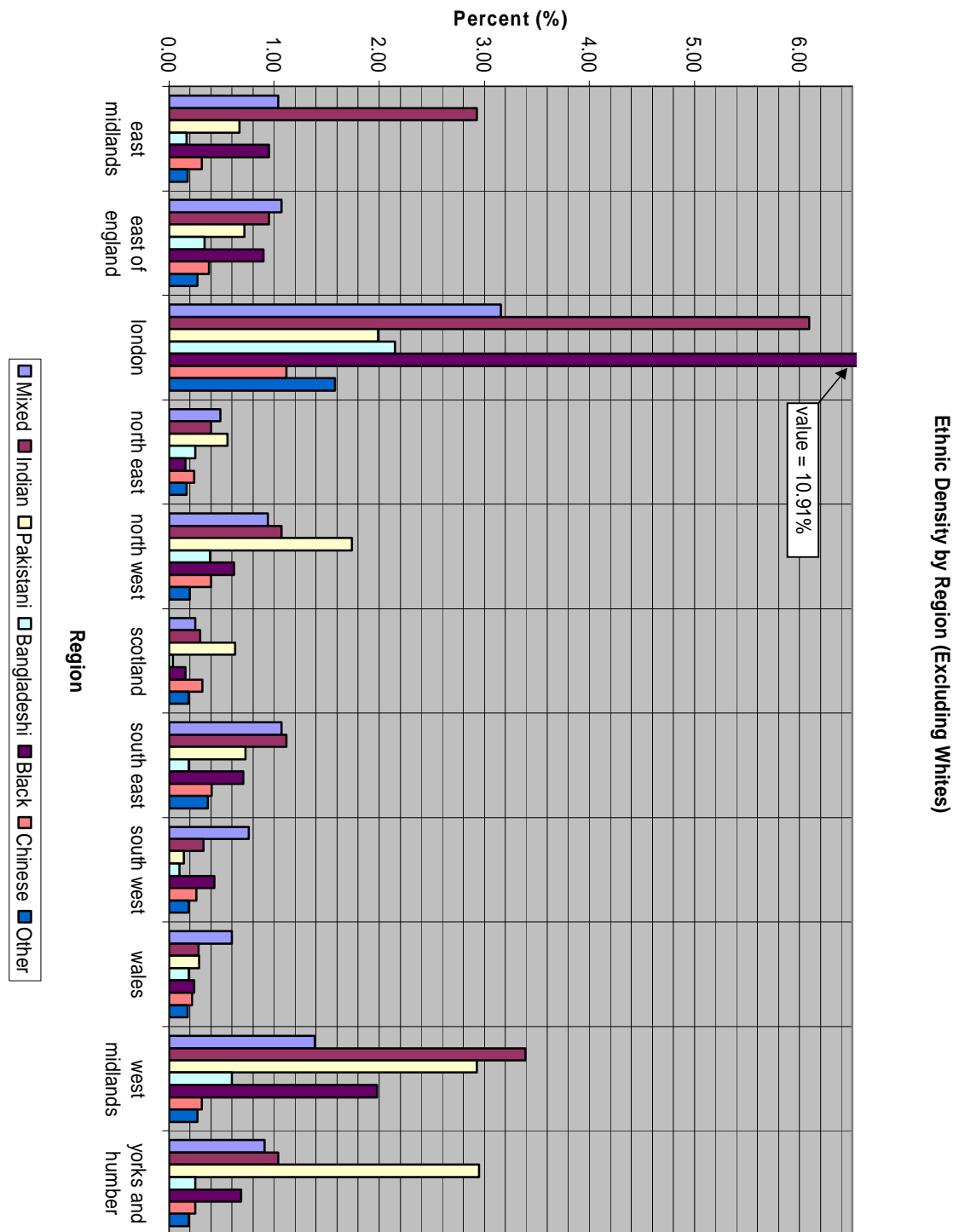
[35] The figure is drawn without the white ethnic group, because the percentages for white ethnic density are mainly in the 90's (apart form London which has a white ethnic density of 68.08%), which means that the minority groups bar's would simply look like small dots on the graph, had whites been included. Excluding whites from the graph allows a much clearer image and comparison between the minority groups.

[36] For a tabular description of the ethnic density of each region (including whites), please refer to appendix 5.

Interesting relationships can also be extracted by looking at the correlation between variables. The correlation matrix is too large to provide here; but there are not many significant relationships to present (i.e. relationships with a value greater than +/-0.10). I have found that males are more likely to go into either the manufacturing or construction sectors and less likely to go into the services sector. Again there is a negative relationship between ethnic density and self-employment, given by a correlation of -0.106. There is a correlation of 0.292 between an individual's logwage and years in full time education. This is expected because education forms an important part of an individual's human capital, and a superior education is generally believed to coincide with higher earnings. Ethnic density is negatively correlated with all regions except for London, with whom it has a positive correlation of 0.688. The most significant correlations in the matrix come from relationships between London and GDP (0.987) and London and unemployment rate (0.687). These relationships would be expected, due to the large number of ethnic minorities in London, the high unemployment rates and the mass accumulation of wealth.

V. Data Analysis

As outlined in section III, the first stage of the analysis involves calculating the sectoral earnings difference for each individual. This involved using a Mincer wage equation of the type in equation 1, reproduced below:

$$\log earnings^{j}{}_{i} = b_{o}^{j} + b_{1}^{j} Age_{i} + b_{2}^{j} Age^{2} 1000_{i} + b_{3}^{j} Male_{i} + b_{4}^{j} Married_{i} +$$
$$b_{5}^{j} Disability_{i} + b_{6}^{j} YrsEdu_{i} + b_{7}^{j} Edlev\_otherqual_{i} + b_{8}^{j} Edlev\_high\deg_{i} +$$
$$b_{9}^{j} Edlev\_1st\deg_{i} + b_{10}^{j} Edlev\_profqual_{i} + b_{11}^{j} Edlev\_Alev_{i} + b_{12}^{j} Edlev\_GCSE_{i} +$$
$$b_{13}^{j} Ind\_cons_{i} + b_{14}^{j} Ind\_serv_{i} + b_{15}^{j} Ind\_na_{i} + b_{16}^{j} Ind_{i} + b_{17}^{j} Ethnic_{i} + b_{18}^{j} reg\_NW_{i} +$$
$$b_{19}^{j} reg\_YaH_{i} + b_{20}^{j} reg\_EM_{i} + b_{21}^{j} reg\_WM_{i} + b_{22}^{j} reg\_EE_{i} + b_{23}^{j} reg\_Lon_{i} +$$
$$b_{24}^{j} reg\_SE_{i} + b_{25}^{j} reg\_SW_{i} + b_{26}^{j} reg\_Wales_{i} + b_{27}^{j} reg\_S\cot_{i} + e_{i}$$

The equation was estimated by OLS (separately for self-employed and paid-employed sectors). The results are presented below in table 3, and in full detail in appendix 3a.

Table.3

SELF-EMPLOYED EARNINGS EQUATION						PAID-EMPLOYED EARNINGS EQUAITON				
logwage	Coef.	Std. Err.	t	P>t		logwage	Coef.	Std. Err.	t	P>t
age	0.107	0.020	5.360	0.000		age	0.076	0.003	25.120	0.000
ageSq1000	-1.175	0.224	-5.260	0.000		ageSq1000	-0.862	0.036	-23.730	0.000
male	0.799	0.062	12.930	0.000		male	0.616	0.010	61.480	0.000
married	0.134	0.062	2.160	0.031		married	-0.035	0.011	-3.190	0.001
disability	-0.094	0.061	-1.540	0.123		disability	-0.032	0.011	-2.940	0.003
yrsedu	0.018	0.009	2.030	0.042		yrsedu	0.001	0.002	0.740	0.459
edlev_othe~l	0.166	0.121	1.370	0.172		edlev_othe~l	0.161	0.023	6.840	0.000
edlev_high~g	0.751	0.149	5.020	0.000		edlev_high~g	0.995	0.026	37.700	0.000
edlev_1stdeg	0.678	0.105	6.460	0.000		edlev_1stdeg	0.846	0.020	42.750	0.000
edlev_prof~l	0.396	0.100	3.960	0.000		edlev_prof~l	0.566	0.018	30.790	0.000
edlev_ALev	0.388	0.094	4.140	0.000		edlev_ALev	0.417	0.017	24.350	0.000
edlev_GCSE	0.308	0.084	3.650	0.000		edlev_GCSE	0.274	0.016	17.130	0.000
ind_cons	0.212	0.095	2.220	0.026		ind_cons	-0.001	0.024	-0.040	0.970
ind_serv	-0.003	0.081	-0.040	0.968		ind_serv	-0.177	0.012	-14.440	0.000
ind_na	-0.100	0.397	-0.250	0.800		ind_na	-0.032	0.087	-0.036	0.719
ethnic_All~s	-0.142	0.158	-0.900	0.369		ethnic_All~s	-0.093	0.030	-3.160	0.002
reg_NW	0.119	0.178	0.670	0.505		reg_NW	0.088	0.026	3.330	0.001
reg_YaH	0.045	0.186	0.240	0.810		reg_YaH	0.076	0.027	2.790	0.005
reg_EM	-0.158	0.189	-0.084	0.403		reg_EM	0.108	0.028	3.830	0.000
reg_WM	-0.170	0.188	-0.900	0.366		reg_WM	0.085	0.027	3.120	0.002
reg_EE	0.090	0.179	0.510	0.614		reg_EE	0.198	0.027	7.360	0.000

reg_Lon	0.236	0.175	1.350	0.178	reg_Lon	0.400	0.027	14.940	0.000
reg_SE	0.162	0.170	0.950	0.342	reg_SE	0.258	0.025	10.230	0.000
reg_SW	0.047	0.177	0.260	0.792	reg_SW	0.042	0.027	1.550	0.121
reg_Wales	-0.305	0.202	-1.510	0.132	reg_Wales	0.064	0.031	2.070	0.039
reg_Scot	0.010	0.189	0.050	0.957	reg_Scot	0.144	0.027	5.240	0.000
_cons	6.150	0.481	12.800	0.000	_cons	7.898	0.067	117.880	0.000

The important information from table 3 comes from the sign on the coefficient and the t-value. The final column P>t gives the likelihood that the stated coefficient will exceed the critical t-value (for a further discussion of the diagnostic tests refer to appendix 3b).

The results obtained are also comparatively standard for this type of work. The base case is a single white female without a disability or any qualifications, working in the manufacturing industry and living in the north-east of England. Making comparisons between variables is not important here, it is important to get a feel for the factors which affect earnings in general. Earnings are positively related to age, gender, years in full-time education and qualifications and negatively related to health. Interestingly an individual who is married tends to earn more than a single person in self-employment and less than a single person in paid-employment. As expected, the coefficient on the ethnicity variable is significant and negative, with an ethnic individual likely to earn 14.2% less in self-employment and 9.3% less in paid-employment than a white individual. This highlights the possible aspect of wage discrimination discussed earlier, but surprisingly the effect is more negative in self-employment[37]. Most of the industry variables are insignificant across both sectors as are most of the region variables for the self-employed. But, the region effect seems to be insignificant only for self-employed people, as most region dummies are significant in the paid-employed equation. An individual in paid-

---

[37] The larger negative effect in self-employment could also be explained by the possibility that self-employed whites have larger businesses, because they cater for a larger market and customer base. Therefore they are likely to earn more relative to ethnic minority self-employed individuals.

22

employment living in London is likely to earn upto 40% more than an individual in paid-employment in the North East.

The resulting predicted values of the earnings differential can now be incorporated in a probit model with several human capital variables, local labour market condition variables, industry dummies and importantly the ethic density variable, to predict their effects on the probability of self-employment. Due to the possibility of colinearity between the ethnic density variable, the ethnic group variables and region variables, the latter two are omitted from the probit model[38]. The probit equation is estimated using equation 3, reproduced below:

$$
\begin{aligned}
Z_i^* = {} & a_0 + a_1(\ln \overline{Y}_i^{SE} - \ln \overline{Y}_i^{PE}) + a_2 Age_i + a_3 Age^2/1000_i + a_4 Male_i + \\
& a_5 Married_i + a_6 Disability_i + a_7 DepChil_i + a_8 UnempRate_i + a_9 GDP_i + \\
& a_{10} YrsEdu_i + a_{11} Edlev\_otherqual_i + a_{12} Edlev\_high\deg_i + \\
& a_{13} Edlev\_1st\deg_i + a_{14} Edlev\_profqual_i + a_{15} Edlev\_Alev_i + a_{16} Edlev\_GCSE_i + \\
& a_{17} Ind\_cons_i + a_{18} Ind\_serv_i + a_{19} Ind\_na_i + a_{20} EthDensity_i + h_i
\end{aligned}
$$

The results of the marginal effects are summarised in table.4, and a full description is given in appendix 4c, with details of diagnostic tests in appendix 4b.

---

[38] This is the case because the ethnic density variable is a combination of both the ethnic group and region variables. Therefore it would not make economic sense to have them together in the same probit regression.

Table.4[39]

PROBIT MODEL: MARGINAL EFFECTS

selfemp	dF/dx	Std. Err.	z	P>z
EarnDiff	0.0104	0.0016	6.4300	0.0000
age	0.0072	0.0011	6.2300	0.0000
Age^2/1000	-0.0399	0.0135	-2.9500	0.0030
male*	0.0698	0.0035	19.7500	0.0000
married*	-0.0045	0.0041	-1.1000	0.2730
disabilityy*	-0.0091	0.0037	-2.4200	0.0150
Depchil	0.0179	0.0018	9.9200	0.0000
UnempRate	-0.0076	0.0015	-5.0400	0.0000
GDP	0.0003	0.0001	4.1500	0.0000
yrsedu	0.0012	0.0005	2.3500	0.0190
Edlev_other*	0.0182	0.0066	2.8600	0.0040
Edlev_high~g*	-0.0154	0.0084	-1.7300	0.0840
Edlev_1stdeg*	0.0101	0.0071	1.4600	0.1440
Edlev_prof~l*	-0.0079	0.0062	-1.2500	0.2110
Edlev_ALev*	-0.0051	0.0059	-0.8400	0.3990
Edlev_GCSE*	0.0055	0.0057	0.9800	0.3280
ind_cons*	0.2646	0.0120	28.3500	0.0000
ind_serv*	0.0405	0.0039	9.6400	0.0000
ind_na*	0.1063	0.0413	3.1600	0.0020
EthDensity	-0.0003	0.0001	-4.0200	0.0000
obs.P	0.1235			
pred.P	0.1036		(at x-bar)	

The propensity to become self-employed is increasing in age, but this effect starts to reverse as an individual gets older, as shown by the negative sign on the second age variable, which accounts for the non-linearity of age. This can be explained by risk. It is believed that older people are less likely to take risks than younger individuals, and the mental and physical demands of self-employment are likely to deter older individuals. The marriage variable has a negative, but insignificant effect, as represented by the low z value of -1.100. The only other insignificant variables in the probit model surprisingly come for the qualification dummies,[40] except for the coefficient on other qualifications. The effect is mixed; therefore in aggregate, formal qualifications have an ambiguous effect on self-employment propensity. The variable for years in full-time education is

---

[39] The table is reproduced in appendix 4c in full detail
[40] Represented by variables beginning with edlev_

24

positive and significant, suggesting that an extra year in education increases the probability of self-employment by 0.12%. This positive relationship arises because a higher level of education is likely to reduce an individual's variation in earnings, because the more educated tend to be better informed, and therefore are more efficient at assessing self-employment opportunities[41].

The longer hours, greater responsibility and more demanding nature of self-employment helps explain the negative coefficient on the disability variable. An extra dependent child increases the probability of self-employment significantly by 1.79%. The argument behind this is that family members provide a convenient source of labour, making self-employment less demanding on the one individual. An individual is 26.46% more likely to be self-employed in the construction industry over the manufacturing industry, and 4.05% more likely in the services over the manufacturing industry. This can be explained by the declining nature of the UK's manufacturing industry, consequently reducing opportunities to enter self-employment. Local labour market conditions also seem to have a significant effect, positively for GDP and negatively for the unemployment rate, both of which make economic sense.

The earnings differential is significant and positive, suggesting that higher earnings in self-employment over paid-employment attract individuals to self-employment. But in contradiction, we have seen in section IV, on average earnings in self-employment are lower than they are in paid-employment. Table.4 shows that a unit increase in the log earnings differential is likely to lead to an increase in the probability of entering self-employment by 1.04% which is much lower than the 9% predicted by Clark &

---

[41] Rees H & Shah A (1986). *An Empirical analysis of SE in the UK*. Journal of Applied Econometrics, Vol.1 No.1

Drinkwater (2000). But, in opposition to my results, they found that on average self-employed individuals actually earn more than paid-employed individuals, which could be a contributing factor to this difference. If earnings discrimination is a factor then other things equal, discrimination in paid-employment will increase the log earnings differential, and therefore increase the probability of self-employment. This is because discrimination lowers the opportunity cost of taking up self-employment.

The focus of the paper is the enclave effect, and whether it helps to explain the high rates of self-employment experienced among the UK's ethnic minorities. The ethnic density variable in table.4 tells us that the effect of living in an enclave is significant and negative. A unit increase in the percentage of co-ethnics residing in a region is likely to reduce the propensity to become self-employed by 0.03%, which does not seem a tremendous amount. Therefore individuals are less likely to be self-employed in regions where there are higher concentrations of co-ethnics. This includes all individuals, including whites. I repeated the probit regression considering only ethnic minorities, but I found that most of the variables had the same sign and marginal effects in the same region as in table 4[42]. The ethnic density variable was still negative, but the negative effect was much greater. Now, a unit increase in the percentage of co-ethnic minorities residing in a region is likely to reduce an ethnic minority individual's probability of self-employment by 0.51%, which is a considerable amount. Therefore I can conclude that, <u>ethnic minorities</u> are <u>LESS LIKELY</u> to be self-employed in regions where there are <u>higher concentrations</u> of <u>co-ethnics.</u>

These findings are similar to what Clark & Drinkwater (2000) found. But, instead of using a continuous variable for ethnic density they characterised their group into, 2-10%,

---

[42] Therefore I have not included the results of this regression in the paper.

10%-25% and over 25% of own ethnic group in ward. They found the coefficient to be more negative as the percentage of co-ethnics in the ward increased; therefore they found minorities are far less likely to be self-employed in areas where there are higher concentrations of co-ethnics[43].

The results can be explained by the hypotheses arguing against the enclave effect outlined in the introduction. First, enclaves involve more competition in the supply of goods and services offered by ethnic businesses due to the fairly small and restricted market. The nature of the market can also lead to difficulties in expanding into the wider market, thus constraining potential for growth, as outlined by Aldrich & Waldinger (1990). Standard economic theory tells us that there is a limit to the number of firms who can operate in a market, especially in one catering for specialist goods, therefore there is a point at which the market becomes saturated and profits are reduced to the normal level, thus reducing the incentives for new entrants to enter the market. Although there may be a greater ratio of businesses with respect to the enclave population[44], this does not rule out the fact that the market can become saturated with no space for duplication of business services already provided.

Also, not all goods and services can be described as having an ethnic dimension, therefore, such goods can be provided by members of any ethnic community[45]. For example taxis, beauty salons etc, and many professional services firms are patronised by white communities who undoubtedly provide their services to all ethnic groups. The larger customer base offered by the wider market means that such firms will be able to

---

[43] They found that over the base category of 'less than 2% of own group in ward', having 2-10% reduces probability of self-employment by 2.2%, 10-25% reduces it by 2.9% and >25% reduces it by 3.0%

[44] As outlined in the positive enclave effect argument in the introduction.

[45] This argument again goes against the positive effect induced by the need for ethnic specific businesses, as outlined in the introduction.

offer goods and services at more competitive rates (due to economies of scale) than an ethnic business aimed at serving the needs of a specific ethnic enclave. This will reduce the incentive for ethnic minorities to enter self-employment. Contrast this to a business selling halal meat, which is aimed at a specific (Muslim) market. In an enclave (most likely Pakistani) there is only going to be scope for a limited number of firms selling halal meat, and because it is an ethnic specific good, the potential for growth into the wider market is restricted. Competition between ethnic groups for the market is likely to deter new entrants.

Secondly, enclaves might offer better opportunities for ethnic minorities in paid-employment if ethnic employers offer non-discriminatory employment opportunities. This overcomes the problem of discrimination from whites in the paid-employment sector. Discrimination in the paid labour market effectively reduces the opportunity cost of going into self-employment[46]. Therefore if an ethnic minority is experiencing wage discrimination in paid-employment from whites, then they can seek other sources of employment. The argument here is that rather than choosing to establish new businesses, ethnic minorities seek employment from co-ethnics from whom they can gain non-discriminatory employment.

As a final point, Clark & Drinkwater (1998) suggest that areas of ethnic minority concentration tend to be low income areas, which reduces entrepreneurial opportunities over paid-employment opportunities. This is a topic which requires more in-depth research.

---

[46] Opportunity cost provides justification for the reallocation of individuals' resources, including time, money and in this case, entrepreneurial talent. This is because no matter what an individual is doing, there is always something else they can do instead.

VI. Conclusion

Using data combined from the last four General Household Surveys, I have investigated the link between ethic enclaves and the high rates of self-employment experienced among ethnic minorities in the UK. The framework used is one in which the self-employment decision is influenced by sectoral earnings differentials and other control characteristics including a variable representing the ethnic density of the region where the individual resides.

This paper finds that the enclave effect is in fact negative in the UK. That is, ethnic minorities are less likely to become self-employed in areas where there is a high concentration of co-ethnics. This conclusion goes against the theory of the protected market hypothesis established by Aldrich & Waldinger (1990), but agrees with the work of Clark & Drinkwater (2000). There are two main factors used to explain this phenomenon. Firstly, the intense competition in enclaves for a small and restricted market discourages individuals from entering self-employment (because the incentives to enter, i.e. profits, are not large enough), and secondly because being in an enclave might offer better opportunities for ethnic minorities in paid-employment if ethnic employers offer non-discriminatory employment opportunities, thus making paid-employment a more attractive opportunity than it previously was.

The results suggest that other than ethnic density, the sectoral earnings differential plays a powerful role in determining an individual's choice of sector, with males also significantly more likely to be in self-employment than females.

The paper provides useful insights into an under-researched area of the UK labour market. The main constraint has been the availability of a large upto date dataset, which over-samples ethnic minorities. If such a dataset was available, the analysis could be conducted for each sub-sample ethnic group, so as to not conceal important inter-group differences. The factors highlighted in this paper do not by any means exhaust the possible explanations for the high rates of self-employment among the UK's ethnic minorities. The aim of the paper was more an attempt to focus on one under-research factor (ethnic enclaves), whilst also accounting for other important variables, such as the sectoral earnings differential.

Another problem constraining the analysis could have been that he regions used may be too large to capture what is meant by an enclave. An enclave may be better thought of as a neighbourhood or a housing estate in which an ethnic minority is concentrated. This factor is again constrained by data limitations.

There are several factors considered to be important, which have not been included simply due to data restrictions. For example access to capital, lending and consumer discrimination as well as parental wealth are all considered to be important in determining the sectoral choice of individuals, but data to proxy such factors is difficult to obtain.

VII. Acknowledgments

I would like to thank my tutor Carlo Perroni, course leader Ian Walker as well as Jeremy Smith for their influential advice and guidance in helping me successfully complete this paper. Their help has been greatly appreciated.

I would also like to thank my friends and colleagues for their support and guidance.

References

Aldrich, H & Waldinger, R (1990). *Ethnicity and entrepreneurship.* Annual Review of Sociology 16, 111-135.

Clark K & Drinkwater S (2000). *Pushed out or pulled in? SE among EM in England and Wales.* Labour Economics, 7, 603-628

Blackaby et al (1998). *The ethnic wage gap and employment differentials in the 1990s: Evidence for Britain.* Economic Letters, 58, 97-103

Blinder, A.S (1973), '*Wage Discrimination: Reduced Form and Structural Variables.*' Journal of Human Resources, 8, 436-455

Oaxaca, R (1973), '*Male-Female Wage Differentials in Urban Labour Markets'* International Economic Review, 14, 693-709

Fairlie R & Meyer B (1996). *Ethnic and Racial SE Differences and possible explanations.* Journal of Human Resources31, 757-795

Rafiq, M (1992). *Ethnicity and enterprise: a comparison of Muslim and non-Muslim owned Asian businesses in Britain.* New Community 19, 43-60.

Borjas G (1986). *The SE experience of Immigrants.* Journal of Human Resources21, 485-506

Dunn, T & Holtz-Eakin, D (1996). *Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links.* Syracuse University, 1996.

Lentz, B & Laband, D (1990). Entrepreneurial Success and Occupational Inheritance among Proprietors. Canadian Journal of Economics 23, 563-579.

Borjas, G & Bronars, S. *Consumer Discrimination and Self-employment.* Journal of Political Economy 97, 581-605.

Coate, S & Tennyson, S. *Labour Market Discrimination, Imperfect Information and Self-Employment.* Oxford Economic Papers 44, 272-288.

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

Rees H & Shah A (1986). *An Empirical analysis of SE in the UK.* Journal of Applied Econometrics, Vol.1 No.1

Fairlie, Robert W. 1999. *The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment*, Journal of Labor Economics, 17(1): 80-108.

Bates, Timothy. 1987, *Self-Employed Minorities: Traits and Trends*, Social Science Quarterly, 68, pp. 539-551.

Yuengert, Andrew M. 1995. *Testing Hypothesis of Immigrant Self-Employment*, Journal of Human Resources 30(1):194-204

Evans, MDR, 1989. *Immigrant Entrepreneurship: effects of ethnic market size and isolated labor market pool.* American Sociological Review 54, 950-962.

Basu, A. (1998). *An Exploration of Entrepreneurial Activity among Asian Small Businesses in Britain.* Small Business Economics, 10 (4), 313-326.

www.data-archive.ac.uk

Gujarati, Damodar (1999). *Essentials of Econometrics,* Second Edition, Irwin McGraw-Hill

Appendix

Appendix 1

Variable	Description
age	Age of the individual
ageSq1000	This accounts for the non-linearity of age
male	Dummy variable, 1 if the individual is male and 0 if female
married	Dummy variable, 1 if the individual is married and 0 if single
disability	Dummy variable, 1 if the individual is considered to have a disability and 0 if otherwise
yrsedu	The number of years the individual has spent in full time education
edlev_othe~l	Dummy variable for the individuals highest attained education level, representing an individual with other qualifications
edlev_high~g	Dummy variable for the individuals highest attained education level, representing an individual with a higher degree
edlev_1stdeg	Dummy variable for the individuals highest attained education level, representing an individual with a normal undergraduate degree
edlev_prof~l	Dummy variable for the individuals highest attained education level, representing an individual with professional qualifications
edlev_ALev	Dummy variable for the individuals highest attained education level, representing an individual with a-levels
edlev_GCSE	Dummy variable for the individuals highest attained education level, representing an individual with GCSE's
ind_cons	Dummy variable representing an individual employed in the construction industry
ind_serv	Dummy variable representing an individual employed in the services industry
ind_na	Dummy variable representing an individual employed in an uncategorised industry
ethnic_All~s	Dummy variable, 1 if the individual belongs to an ethnic minority group and 0 if white
reg_NW	Dummy variable, 1 if the individual lives in the North East region of England and 0 otherwise
reg_YaH	Dummy variable, 1 if the individual lives in the Yorkshire & Humberside region of England and 0 otherwise
reg_EM	Dummy variable, 1 if the individual lives in the East Midlands region of England and 0 otherwise
reg_WM	Dummy variable, 1 if the individual lives in the West Midlands region of England and 0 otherwise
reg_EE	Dummy variable, 1 if the individual lives in the East of England and 0 otherwise
reg_Lon	Dummy variable, 1 if the individual lives in the London region of England and 0 otherwise
reg_SE	Dummy variable, 1 if the individual lives in the South East region of England and 0 otherwise
reg_SW	Dummy variable, 1 if the individual lives in the South West region of England and 0 otherwise
reg_Wales	Dummy variable, 1 if the individual lives in Wales and 0 otherwise
reg_Scot	Dummy variable, 1 if the individual lives in Scotland and 0 otherwise
depchil	Number of dependent children the individual has
UnempR~e	The unemployment rate, relating to each specific region
GDP	The GDP rate relating to each specific region
EthDen~y	Continuous variable representing the ethnic density of the region the individual resides in, depending on the individuals ethnic group and region of residence

Appendix 2

Descriptive statistics for the self-employed and paid-employed				
	Self-employed, n=4294		Paid-employed, n=30484	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
age	44.321	10.806	39.906	11.660
ageSq1000	2.081	0.954	1.728	0.950
male	0.717	0.450	0.505	0.500
married	0.716	0.451	0.618	0.486
disability	0.262	0.440	0.255	0.436
depchil	0.813	1.105	0.669	0.986
selfemp	1.000	0.000	0.000	0.000
ethnic_white	0.935	0.246	0.949	0.220
ethnic_mixed	0.023	0.151	0.016	0.127
ethnic_indian	0.007	0.085	0.008	0.090
ethnic_pak	0.013	0.111	0.004	0.067
ethnic_bang	0.007	0.080	0.006	0.075
ethnic_black	0.005	0.068	0.009	0.096
ethnic_chinese	0.002	0.048	0.001	0.037
ethnic_other	0.008	0.089	0.006	0.074
ethnic_All~x	0.057	0.231	0.045	0.208
yrsedu	17.860	3.733	17.903	3.739
edlev_noqual	0.173	0.378	0.148	0.355
edlev_otherl	0.171	0.376	0.128	0.334
edlev_higherdeg	0.043	0.203	0.050	0.218
edlev_1stdeg	0.143	0.350	0.144	0.351
edlev_prof~l	0.115	0.319	0.128	0.334
edlev_ALev	0.141	0.348	0.175	0.380
edlev_GCSE	0.214	0.410	0.227	0.419
ind_manu	0.095	0.294	0.211	0.408
ind_cons	0.211	0.408	0.053	0.224
ind_serv	0.641	0.480	0.724	0.447
ind_na	0.005	0.068	0.003	0.057
reg_NE	0.030	0.170	0.046	0.209
reg_NW	0.110	0.313	0.114	0.317
reg_YaH	0.079	0.269	0.087	0.282
reg_EM	0.071	0.257	0.075	0.264
reg_WM	0.074	0.262	0.092	0.289
reg_EE	0.103	0.304	0.099	0.298
reg_Lon	0.132	0.339	0.109	0.311
reg_SE	0.171	0.377	0.152	0.359
reg_SW	0.113	0.317	0.092	0.289
reg_Wales	0.046	0.210	0.048	0.214
reg_Scot	0.070	0.255	0.087	0.282
EthDensity	85.278	23.480	87.126	21.048

grmainjb	30679.810	78875.340	31835.620	51610.710
logwage	9.837	1.575	10.160	0.913
EarnDiffb	-0.247	1.368	-0.442	0.848
UnempRate	3.493	1.096	3.592	1.139
GDP	106.024	27.190	103.554	25.442

Appendix 3

Results of the Mincer OLS Equation

SELF-EMPLOYED EARNINGS EQUATION					PAID-EMPLOYED EARNINGS EQUAITON				
Source	SS	df	MS		Source	SS	df	MS	
Model	819.627	26.000	31.524		Model	6304.352	26.000	242.475	
Residual	7103.825	3168.000	2.242		Residual	15646.853	26331.000	0.594	
Total	7923.453	3194.000	2.481		Total	21951.205	26357.000	0.833	
Number of obs	3195.000				Number of obs	26358.000			
F(26, 3168)	14.060				F(26, 26331)	408.040			
Prob > F	0.000				Prob > F	0.000			
R-squared	0.103				R-squared	0.287			
Adj R-squared	0.096				Adj R-squared	0.287			
Root MSE	1.498				Root MSE	0.771			
logwage	Coef.	Std. Err.	t	P>t	logwage	Coef.	Std. Err.	t	P>t
age	0.107	0.020	5.360	0.000	age	0.076	0.003	25.120	0.000
ageSq1000	-1.175	0.224	-5.260	0.000	ageSq1000	-0.862	0.036	-23.730	0.000
male	0.799	0.062	12.930	0.000	male	0.616	0.010	61.480	0.000
married	0.134	0.062	2.160	0.031	married	-0.035	0.011	-3.190	0.001
disability	-0.094	0.061	-1.540	0.123	disability	-0.032	0.011	-2.940	0.003
yrsedu	0.018	0.009	2.030	0.042	yrsedu	0.001	0.002	0.740	0.459
edlev_othe~l	0.166	0.121	1.370	0.172	edlev_othe~l	0.161	0.023	6.840	0.000
edlev_high~g	0.751	0.149	5.020	0.000	edlev_high~g	0.995	0.026	37.700	0.000
edlev_1stdeg	0.678	0.105	6.460	0.000	edlev_1stdeg	0.846	0.020	42.750	0.000
edlev_prof~l	0.396	0.100	3.960	0.000	edlev_prof~l	0.566	0.018	30.790	0.000
edlev_ALev	0.388	0.094	4.140	0.000	edlev_ALev	0.417	0.017	24.350	0.000
edlev_GCSE	0.308	0.084	3.650	0.000	edlev_GCSE	0.274	0.016	17.130	0.000
ind_cons	0.212	0.095	2.220	0.026	ind_cons	-0.001	0.024	-0.040	0.970
ind_serv	-0.003	0.081	-0.040	0.968	ind_serv	-0.177	0.012	-14.440	0.000
ind_na	-0.100	0.397	-0.250	0.800	ind_na	-0.032	0.087	-0.036	0.719
ethnic_All~s	-0.142	0.158	-0.900	0.369	ethnic_All~s	-0.093	0.030	-3.160	0.002
reg_NW	0.119	0.178	0.670	0.505	reg_NW	0.088	0.026	3.330	0.001
reg_YaH	0.045	0.186	0.240	0.810	reg_YaH	0.076	0.027	2.790	0.005
reg_EM	-0.158	0.189	-0.084	0.403	reg_EM	0.108	0.028	3.830	0.000
reg_WM	-0.170	0.188	-0.900	0.366	reg_WM	0.085	0.027	3.120	0.002
reg_EE	0.090	0.179	0.510	0.614	reg_EE	0.198	0.027	7.360	0.000
reg_Lon	0.236	0.175	1.350	0.178	reg_Lon	0.400	0.027	14.940	0.000
reg_SE	0.162	0.170	0.950	0.342	reg_SE	0.258	0.025	10.230	0.000
reg_SW	0.047	0.177	0.260	0.792	reg_SW	0.042	0.027	1.550	0.121
reg_Wales	-0.305	0.202	-1.510	0.132	reg_Wales	0.064	0.031	2.070	0.039
reg_Scot	0.010	0.189	0.050	0.957	reg_Scot	0.144	0.027	5.240	0.000
_cons	6.150	0.481	12.800	0.000	_cons	7.898	0.067	117.880	0.000

Appendix 3b: DISCUSSION OF DIAGNOSTIC TESTS

Test of individual significance
As can be seen from appendix 3a above, STATA already provides t values, which are basically compared to the t critical value to determine whether or not each variable is significant. An example for any variable:

Null hypothesis (H0): $\beta k=0$ k=1,…..,27
Alternative Hypothesis (H1): $\beta k\neq0$
Significance level (α) = 5%
t-critical = $t_{a/2}^{n-k} = t_{0.025}^{\infty} = +/-1.960$ (n=number of observations, k=number of estimated coefficients. n-k >100, so we assume its equal to infinity).

Therefore any t-value greater than 1.960 or less than -1.960 is insignificant. As the table shows in appendix 3a, there are many more insignificant values in the self-employed earnings equation than the paid-employed equation.

Test for Joint Significance
More important is the test for joint significance to ensure that all the variables are jointly significant. We do this through an F-test separately for the self-employed and paid-employed equations. Again STATA provides us with an F-value, which needs to be compared to a critical F value. The F-value is calculated using the following formula.

$$F = \frac{(R^2/k)}{(1-R^2)/(n-k-1)}$$

Null hypothesis (H0): $\beta0= \beta1=……= \beta27=0$
Alternative Hypothesis (H1): Any $\beta k\neq0$
Significance level (α) = 5%
Critical F value: $F_a^{k,n-k-1} = F_{0.05}^{27,\infty} = 1.460$

Paid Employment Equation
F-Value = 408.04
Critical F Value = 1.460
REJECT H0

Self Employment Equation
F-Value = 14.060
Critical F Value = 1.460
REJECT H0

Therefore both Mincer earnings equations (paid-employed and self-employed) reject the null that all coefficients are jointly equal to zero. This result adds validity to the model.

<u>Test for omitted relevant variables and incorrect functional form: RESET TEST</u>

Paid Employment Equation
      Ramsey RESET test using powers of the fitted values of logwage
          Ho: model has no omitted variables
              $F(3, 26328) =$ 30.47
                $Prob > F =$ 0.0000

Self Employment Equation
      Ramsey RESET test using powers of the fitted values of logwage
          Ho: model has no omitted variables
              $F(3, 3165) =$ 1.70
                $Prob > F =$ 0.1655

Both equations pass the reset test.

Appendix 4a

This table simply displays the results from the probit model, which in terms of interpretation use are pretty useless. They are simply presented here for reference.

Probit estimates				Number of obs =	34778	
				LR chi2(20) =	2424.36	
				Prob > chi2 =	0	
Log likelihood = -11787.137				Pseudo R2 =	0.0932	

selfemp	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
EarnDiffb	0.0579	0.0090	6.4300	0.000	.0402728	0.0756
age	0.0398	0.0064	6.2300	0.000	.0272539	0.0523
ageSq1000	-0.2216	0.0752	-2.9500	0.003	-.3689774	-0.0742
male	0.3921	0.0198	19.7500	0.000	.3532088	0.4310
married	-0.0248	0.0226	-1.1000	0.273	-.0691391	0.0195
disability	-0.0511	0.0211	-2.4200	0.015	-.0924085	-0.0098
depchil	0.0996	0.0100	9.9200	0.000	.0799342	0.1193
UnempRate	-0.0421	0.0084	-5.0400	0.000	-.0585039	-0.0258
GDP	0.0016	0.0004	4.1500	0.000	.0008672	0.0024
yrsedu	0.0065	0.0028	2.3500	0.019	.0010798	0.0119
edlev_othe~l	0.0965	0.0338	2.8600	0.004	.0303046	0.1628
edlev_high~g	-0.0899	0.0519	-1.7300	0.084	-.1916329	0.0119
edlev_1stdeg	0.0547	0.0375	1.4600	0.144	-.018712	0.1281
edlev_prof~l	-0.0450	0.0360	-1.2500	0.211	-.1156493	0.0256
edlev_ALev	-0.0284	0.0337	-0.8400	0.399	-.0945862	0.0377
edlev_GCSE	0.0302	0.0308	0.9800	0.328	-.0302886	0.0906
ind_cons	0.9623	0.0339	28.3500	0.000	.8957852	1.0289
ind_serv	0.2399	0.0249	9.6400	0.000	.1911636	0.2887
ind_na	0.4550	0.1438	3.1600	0.002	.1730426	0.7369
EthDensity	-0.0019	0.0005	-4.0200	0.000	-.002777	-0.0010
_cons	-2.9307	0.1544	-18.9800	0.000	-3.233391	-2.6281

Appendix 4b

Testing for individual significance

It is important to test for individual significance. I have mentioned some of the significant and insignificant variables in section V above. The z value has the same interpretation as the t-values in appendix 3a.

We can compare the z value with the same critical value of +/-1.960 (at 5% significance level) to see which variables are significant and insignificant. Most importantly the coefficient on the EthDensity variable is significant.

Testing for joint significance

I test for joint significance using a Chow test.

Null hypothesis:	$\alpha 1 = \alpha 2 = \alpha 3 = \ldots = \alpha 20 = 0$	
Alternative Hypothesis:	Any $\alpha_k \neq 0$	k=1,2,....,20
Signifificnace level:	5%	
Chow Critical:	$c_{20}^2 = 10.85$	
Calculated Chow:	$c_{20}^2 = 2424.36$	

The calculated chow value (taken from table in appendix 4a) is much greater than the critical chow value; therefore the probit model passes the joint significance test. That is, it rejects the null hypothesis that all coefficients are jointly equal to zero.

Appendix 4c

This table shows the marginal effects of each variable. (The z values are exactly the same as those in the table in appendix 4a)

Probit estimates: Marginal Effects				Number of obs	34778		
				LR chi2(20)	2424.36		
				Prob > chi2	0		
Log likelihood = -11787.137				Pseudo R2	0.0932		
selfemp	dF/dx	Std. Err.	z	P>z	x-bar	[95%	C.I.]
EarnDi~b	0.0104	0.0016	6.4300	0.0000	-0.4182	0.0073	0.0136
age	0.0072	0.0011	6.2300	0.0000	40.4513	0.0049	0.0094
age~1000	-0.0399	0.0135	-2.9500	0.0030	1.7720	-0.0664	-0.0134
male*	0.0698	0.0035	19.7500	0.0000	0.5315	0.0630	0.0766
married*	-0.0045	0.0041	-1.1000	0.2730	0.6297	-0.0125	0.0036
disabi~y*	-0.0091	0.0037	-2.4200	0.0150	0.2562	-0.0163	-0.0018
depchil	0.0179	0.0018	9.9200	0.0000	0.6866	0.0144	0.0215
UnempR~e	-0.0076	0.0015	-5.0400	0.0000	3.5797	-0.0105	-0.0046
GDP	0.0003	0.0001	4.1500	0.0000	103.8590	0.0002	0.0004
yrsedu	0.0012	0.0005	2.3500	0.0190	17.8978	0.0002	0.0021
ed~rqual*	0.0182	0.0066	2.8600	0.0040	0.1336	0.0052	0.0312
edl~hdeg*	-0.0154	0.0084	-1.7300	0.0840	0.0491	-0.0319	0.0011
edl~tdeg*	0.0101	0.0071	1.4600	0.1440	0.1435	-0.0038	0.0240
ed~fqual*	-0.0079	0.0062	-1.2500	0.2110	0.1266	-0.0201	0.0042
edlev_~v*	-0.0051	0.0059	-0.8400	0.3990	0.1713	-0.0167	0.0066
edlev_~E*	0.0055	0.0057	0.9800	0.3280	0.2252	-0.0056	0.0166
ind_cons*	0.2646	0.0120	28.3500	0.0000	0.0727	0.2410	0.2881
ind_serv*	0.0405	0.0039	9.6400	0.0000	0.7140	0.0328	0.0482
ind_na*	0.1063	0.0413	3.1600	0.0020	0.0035	0.0252	0.1873
EthDen~y	-0.0003	0.0001	-4.0200	0.0000	86.8977	-0.0005	-0.0002
obs.P	0.1235						
pred.P	0.1036		(at x-bar)				

Appendix 5

TABULAR REPRESENTATION OF THE ETHNIC BREAKDOWN OF EACH
REGION

Region	White	Mixed	Indian	Pakistani	Bangladeshi	Black	Chinese	Other
east midlands	92.63	1.04	2.93	0.67	0.17	0.95	0.31	0.18
east of england	93.98	1.07	0.95	0.72	0.34	0.90	0.38	0.27
london	68.08	3.16	6.09	1.99	2.15	10.91	1.12	1.58
north east	97.27	0.49	0.40	0.56	0.25	0.16	0.24	0.17
north west	93.28	0.94	1.07	1.74	0.39	0.62	0.40	0.20
scotland	97.01	0.25	0.30	0.63	0.04	0.16	0.32	0.19
south east	94.07	1.07	1.12	0.73	0.19	0.71	0.41	0.37
south west	97.05	0.76	0.33	0.14	0.10	0.43	0.26	0.19
wales	97.88	0.60	0.28	0.29	0.19	0.24	0.22	0.18
west midlands	87.35	1.39	3.39	2.93	0.60	1.98	0.31	0.27
yorks and humber	92.82	0.91	1.04	2.95	0.25	0.69	0.25	0.19