# Self-employment among male immigrants in Denmark

## Long-term unemployment and comparative advantages

Cecilie Dohlmann

*Social integration and maginalisation*
*Working Paper 08:2001*

**Working Paper**
*Socialforskningsinstituttet*
*The Danish National Institute of Social Research*

# *Self-employment among male immigrants in Denmark*

## *Long-term unemployment and comparative advantages*

*Cecilie Dohlmann*

***Social integration and marginalisation***
***Working Paper 08:2001***

**Abstract**

This paper investigates whether or not the self-employment choice is different among ethnic Danish men and male immigrants from less developed countries. The main issue is to determine whether the large incidence of self-employment among immigrants can be explained by past difficulties in finding wage-employment, by comparative advantage in self-employment, or simply by demographic and macroeconomic factors. The literature concerning immigrant wage-earners and self-employment choice is abundant, but there is little empirical economic literature dealing with immigrants' self-employment choice in Denmark. The results clearly indicate that ethnicity and past experience of unemployment increase the likelihood of entering into self-employment.

THEME Occupational choice

KEYWORDS Self-employment, choice of occupation, male immigrants

JEL-Code J24,J44,J62

Continue next page...

# Contents

Continue next page...

# 1 Introduction

This paper examines the connection between experience of unemployment and self-employment decisions among immigrants. To some extent, an immigrant's choice of self-employment is theoretically and empirically different from the ethnic Dane's choice of self-employment. The analysis attempts to pinpoint the reasons behind the high incidence of self-employed immigrants in comparison to the low incidence of self-employed ethnic Danes.

Empirical studies have found that immigrants from Nordic countries, Northern Europe, USA, Canada, Australia and New Zealand behave similarly to ethnic Danes in the Danish labour market (e.g. Statistics Denmark, 1998).[1] The focus in this paper is therefore on first and second generation male immigrants from less developed countries which include Turkey, Pakistan, Iran and "other" Asian countries.[2] In 1997 about 7 pct. of the Danish population between 15 and 67 were immigrants but less than 40 pct. of the immigrants came from less developed countries.

Fig. 1 shows the decrease in the share of self-employed among employed ethnic Danes from 1989 to 1997.[3] In the same period the percentage of self-employed among employed first and second generation immigrants increased strongly. It is also observed that the unemployment rate among first generation immigrants is at least 10 pct. points higher than among ethnic Danes throughout the same period (see Fig. 2). The unemployment rate among second generation immigrants increased from 1989 to 97 and appears to approach convergence with first generation unemployment in 1997. The proportion of self-employed immigrants really began to increase in the beginning of the 90's during the economic recession and the increasing unemployment rate. The relationship between a high unemployment rate and a high self-employment rate at the aggregated level could be explained by a relationship between entering self-employment and unemployment at the micro-level. Individuals that are entering self-employment are therefore the main focus in the study.

[1] For a definition of an immigrant, see appendix 8.1.
[2] For a definition on "Other" Asian countries, see appendix 8.2.
[3] The data section describes the sample and definitions of occupation and unemployment rates.

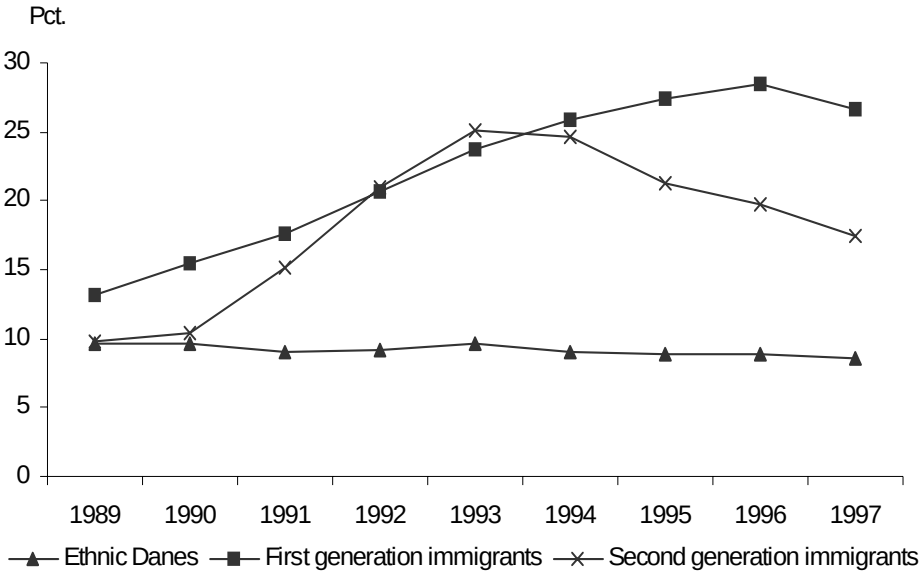


Fig. 1. Proportion of self-employed in different ethnic groups, males aged 15 - 67.

An individual's occupational choice (wage-earner versus self-employed) is explained by a simple choice model, resting on the assumption that a person chooses the occupation that will provide the largest utility. In this instance, utility is a function of the person's expected future labour market situation.

A panel data set based on a sample taken from Danish register data is used. The probability of choosing self-employment in preference to an alternative occupation given that the individual is unemployed is estimated. The transition is estimated by using a simple conditioned probit model and a random effect model. Finally the model is estimated by taking selection into account. Not taking heterogeneity and sample selection into account still seems to give robust estimates.

The empirical results show that controlling for ethnicity is important in modelling transitions into self-employment compared to alternative labour market states. Immigrants from less developed countries have a higher probability of choosing self-employment rather than a job as a wage-earner compared to ethnic Danes. Additionally, it appears that long-term unemployment (measured by historical unemployment) results in a higher probability of transition into self-employment than into wage-employment. Long-term unemployment can of course be voluntary and involuntary. However in this paper, unem-

ployment will be considered as involuntary.

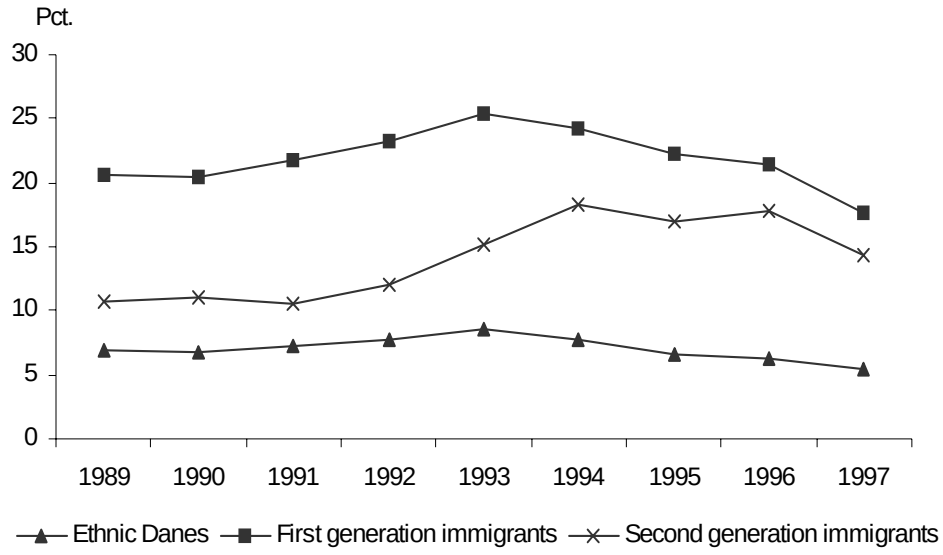


Fig. 2. Unemployment rate among ethnic groups in Denmark, males aged 15 - 67.

It appears from the results that comparative advantage (in the form of having a wife, many children and sector experience) also increases the probability of transition into self-employment over alternative labour market states. The large incidence of self-employed immigrants compared to ethnic Danes is partly explained from the empirical results, which support the hypothesis regarding long-term unemployment and comparative advantage. However, the estimation results also indicate that it is the interaction between many factors that cause an individual to choose self-employment.

The paper is organised as follows. Section 2 explains the conceptual framework and the existing empirical literature. Descriptive statistics are presented in section 3. The empirical model and related problems are described in section 4, while section 5 includes the empirical results. The conclusion is presented in section 6.

Continue next page...

# 2 Background and theory

Economic literature about self-employed immigrants in the Danish labour market has been scarce up to this point. The negative and positive consequences of immigration on the labour market has only been acknowledged and studied within the last 15 years in Denmark. Ejrnæs et al. (2001) find that some immigrant groups choose to be self-employed because they can't find wage-employment and want to escape unemployment. As is the case in Denmark a lot of industrialised countries have a relatively large share of self-employment among immigrants compared to the share of self-employment among the native population. There already exists a considerable amount of international literature about self-employed immigrants.

**The choice of self employment**

The conceptual framework in the present study is inspired by the theoretical approach of Evans and Jovanovic (1989), Evans and Leighton (1989), Taylor (1998) and Carrasco and Ejrnæs (2000). The economy consists of self-employed, wage-earners and unemployed individuals with the following utility functions:

$$U_{it}^{s} = U_{it}^{s}(Y_{it}, X_{it}, \eta_{i}^{s}) \quad (1)$$

$$U_{it}^{w} = U_{it}^{w}(w_{it}, X_{it}, \eta_{i}^{w}) \quad (2)$$

$$U_{it}^{u} = U_{it}^{u}(b_{it}, X_{it}, \eta_{i}^{u}) \quad (3)$$

where $U_{it}^{w}$ is wage-earner $i$ 's utility at time $t$, $U_{it}^{u}$ is the utility derived in unemployment and $U_{it}^{s}$ is the utility of being self-employed. The utility of an occupation is dependent on a pecuniary utility, observed characteristics and a non-pecuniary utility. The pecuniary utility, measured as income, has a positive effect on utility. $Y$ is the self-employed individual's income, $w$ is the wage and $b$ is unemployment benefit. The observed characteristics, $X$, could be age, education, marital status etc. The non-pecuniary utility, $\eta$, is connected with the status that an individual can obtain in a certain occupation. This may include preferences, productivity etc. The reason for including a non-pecuniary utility like

occupational status is that an individual's work is important for individual identity and self-respect. The non-pecuniary utility, $\eta^s$, of being self-employed could include the desire of "being your own master" or personal "freedom". "Safeness", or the comfort provided by having the manager take the last decision, is related to the status of a wage-earner, $\eta^w$. The non-pecuniary utility of being unemployed, $\eta^u$, mainly includes the desire of not fulfilling any demands at all. Often this utility will be associated with a disutility because an unemployed individual may lose self-confidence in a population of employed workers, and secondly, because they may miss out on social contact with co-workers. Individual's different riskaversion is illustrated by the different utility functions.

Assuming an individual's preferences for a certain occupation can be represented by a utility function, the transition into self-employment will occur if the expected utility of becoming self-employed, $E[U^s]$, exceeds the expected utility in an alternative occupation, $E[U^a]$. This relationship can be described by equation (4),

$$E[U^s_{i,t}] > E[U^a_{i,t}] \tag{4}$$

$$where \quad E[U^a_{i,t}] = \max\left[E[U^w_{i,t}], E[U^u_{i,t}]\right] \tag{5}$$

Factors that are considered important to the self-employment decision in earlier literature will now be discussed. Special attention will be paid to the self-employment decision among immigrants.

**Experience of long-term unemployment**

Involuntary unemployment will make self-employment seem attractive because it is a form of employment connected with self-earned income, if the necessary input factors are available.

If the expected utility of self-employment exceeds the expected utility of unemployment but not the expected utility of a wage-employment, then self-employment is chosen to avoid unemployment but not in preference to wage-employment.

In previous literature, the occupational choice of self-employment becoming a second best alternative to employment is mentioned mostly with regard to the disadvantage hypothesis (e.g. Dennis, 1996). This hypothesis has been reinforced by a study among

Canadian men showing the length of unemployment between two jobs to be positively correlated with the probability of becoming self-employed (e.g. Moore and Mueller, 1998).[4] Carrasco and Ejrnæs (2000) and Carrasco (1999) find that unemployed individuals have a high probability of entering self-employment. Furthermore, entry-exit analysis of Canadian men shows, that unfavourable future possibilities of becoming wage-earners make men choose self-employment (e.g. Kuhn and Schuetze, 1998). It has also been found that self-employed individuals often have experienced low wages, high job turnover and numerous short or long unemployment spells. These findings support the disadvantage hypothesis (e.g. Evans and Leighton, 1989).

Several studies on Spanish, British and American data find no support for the disadvantage hypotheses since the aggregated unemployment rate in the country or in a geographical area has a negative or no effect on the self-employment decision (e.g. Clark and Drinkwater, 2000; Taylor, 1996; Alba-Ramirez,1994; Blanchflower, 1998). However, using the aggregated unemployment rate as an approximation for an individual's past unemployment experience cause some problems. First of all the aggregated unemployment rate is likely to indicate the present state of labour market condition in a certain geographical areas. Therefore the approximation seems to indicate the present business cycle, rather than individuals' unemployment probabilities. Second of all, the present aggregated unemployment rate doesn't seem to indicate anything about a person's past unemployment experience.

It is not obvious that immigrant status should result in unemployment. However, if we assume in general that immigrants are less educated than the host population, and that the correlation between low education and unemployment is positive, then immigrants' income will be lower than native population income. Assuming that the distribution of abilities is the same among ethnic groups, it is clear that the percentage of immigrants choosing self-employment is greater than the percentage of the host population if immigrants generally experience more unemployment. Ejrnæs et al. (2001) find results supporting the hypothesis that immigrants in Denmark choose self-employment as a last resort

[4]Moore and Mueller simultaneously find that people on public support have less of a tendency to choose self-employment.

to entering employment.

Immigrants might lack qualifications such as Danish language skills and cultural knowledge, which could create problems for them in the Danish labour market. The missing skills may increase the probability of unemployment or increase the probability of a lower wage. By choosing self-employment, the immigrant can avoid the employers negative evaluation regarding missing language and cultural skills. On the other hand, it could be very difficult to survive in self-employment, since complete independence demands diversified skills. Immigrants with good job possibilities are therefore expected to become self-employed individuals (e.g. Fairlie and Meyer, 1996).

Low education and missing language skills are just some of the factors that can increase the risk of involuntary unemployment. Discrimination and regulations could also increase this risk.

**Discrimination**

The discrimination hypothesis says that the employer selects employees based on imperfect information (skin colour, name etc.). These characteristics are used as a proxy for a persons qualifications despite their irrelevance to the job. The result is that the immigrant is offered a wage that doesn't correspond to his human capital. This wage is lower than the wage offered to an ethnic Dane with exactly the same skills. By becoming self-employed the immigrant can escape discrimination, but only if consumers don't discriminate (e.g. Moore, 1983). Consumers can discriminate, which means immigrants can't always avoid discrimination by becoming self-employed (e.g. Borjas and Bronars 1989).

Clark and Drinkwater's (2000) findings show that wage-differentials between wage-earners and self-employed individuals have a big influence on the self-employment decision among immigrants in the USA. Therefore if the wage-differential increases because of discrimination, more immigrants than Danes are expected to enter self-employment.

**Direct and indirect comparative advantages**

The comparative advantage hypothesis states that an individual will choose self-employment because of a comparative advantage in certain occupations. However, in contrast to the disadvantage hypothesis, the individual doesn't choose to become self-

employed as a last solution to employment. Self-employment is chosen because the comparative advantage has made the expected utility of becoming self-employed exceed the utility in an alternative occupation.

Immigrants comparative advantage could be explained by a "home-country" experience (e.g.Yuengert, 1995). If the emigration country has a large self-employment share, the incidence of immigrants having self-employment experience is expected to be greater than among the host population. The self-employment experience could be a "home-country" advantage. In addition, a "home-country" advantage among immigrants and not among the host population can result in well-functioning immigrant businesses and less profitable ethnic Danish businesses.

A comparative advantage could also include having a self-employed father (e.g. Hout and Rosen, 1999). If it is generally the case that immigrants have better knowledge and information about being self-employed, then the distribution of self-employment ability among immigrants and the host population is different. This will result in a higher proportion of self-employed individuals among immigrants than among the native population.

Another interesting issue is self-employed immigrants' comparative advantages with regard to language and cultural skills in ghettoes (e.g. Borjas, 1986). Yuengert (1995) finds no support for this hypothesis on American data. However, a ghetto definition can be puzzling and cause differences in the results.

The comparative advantages could also be sector specific. In the Danish case, self-employed immigrants are mainly in sectors like retail, restaurants and hotels (e.g. Bager and Rezai 1999).

Individuals can have an indirect comparative advantage as a result of a large and strongly connected family (e.g. Borjas 1986). A big family can make the hiring of reliable workers easier and can save the self-employed individual search costs, such as time and money on hiring and compliance with "union agreement rules". A big family can therefore minimize the cost of self-employment. If immigrants generally have assisting wives and many children , then the proportion of immigrants receiving economical gains by choosing self-employment will be greater than among the native population.

Immigrants compared with ethnic Danes may have clear comparative advantages with respect to family sizes, special language skills and cultural knowledge in ghetto areas. However, at the same time, immigrants may lack comparative advantages with regard to the language, rules and laws of the host country etc. Therefore, immigrant status can result in comparative disadvantages as well.

**Capital restrictions**

A certain amount of capital is necessary to start a business. Previous literature has shown that capital constrained people are less likely to become self-employed (e.g. Evans and Leighton, 1989; Taylor, 1998; Lind and Ohlson, 1996; Blanchflower and Oswald, 1998). Evans and Jovanovic's (1989) structural model found that individual's initial wealth had an effect on choosing self-employment in America, which indicates that capital restrictions may exist.

An alternative explanation concerning the importance of liquidity in self-employment could be that an individual's risk aversion changes when capital becomes less constrained (e.g. Carrasco and Ejrnæs 2000).

Low saving rates and unfavourable loan possibilities because of missing information regarding immigrant credibility, could constrain immigrant capital or cause risk aversion that is higher than among the native population. This would result in fewer immigrants entering self-employment than ethnic Danes.

**Institutional factors**

Past research has shown different results concerning the self-employment decision, when using the same empirical methods and data from different countries. This could be due to country-specific institutional factors that may include self-employment benefits, tax systems and wage-structures.

Self-employment assistance programmes to encourage the growth of self-employment among unemployed and employed individuals have been very popular in industrialised countries. By relaxing the capital constraints through cash flows, the goal of the programmes has been to pull and push individuals into self-employment. In Denmark, both

push and pull effects are found in self-employment assistance programmes (e.g. Høgelund et al., 1992; Plougmann and Buhl, 1998). Wong, Henson and Riddell (1998) find that Canadian self-employment benefits influence the self-employment income and the choice of working full-time. In America, assistant programmes seem to increase the amount of self-employment (e.g. Blau, 1987).

Another interesting institutional factor is the tax system. An incentive to become self-employed exists if the tax laws favour self-employed people. These laws could involve avoiding paying high income taxes. In contrast to a wage-earner, a self-employed individual can earn income in a variety of different ways and hence avoid a high income tax (e.g. Chapman et al., 1998; Blau, 1987; Schuetze, 2000). Bruce (2000) show that reducing an individual's marginal tax rate on self-employment income while holding his marginal wage tax rate constant reduces the probability of becoming self-employed. Conversely, reducing an individual's relative average tax rate in self-employment increases this probability.

Several other institutional factors, such as minimum wage and regulation, can affect the choice of entering self-employment but they will not be explored in this paper.

**Business cycles and assimilation**

An individual's labour market success depends on the relationship between the economic situation in the country and the specific human capital of the individual (e.g. Bradbury, 1994). However, business cycles are not shown to have any effect in the study by Cowling and Mitchell (1997). By contrast, Borjas (1986) finds that due to business cycle conditions, immigrants from the most recent cohorts seem to be self-employed in America.

Time plays an important role in immigrants occupational choice. Assimilation could explain why an immigrant's likelihood of becoming self-employed increases with duration of residence (e.g. Borjas, 1986). Bager and Rezai (1999) find that loan possibilities among immigrants in Denmark improve over time. Therefore an immigrant's probability of becoming self-employed could increase with years of residence because of better financial opportunities.

The time aspect seems to be very important for the occupational choice of temporary

immigrants. A quick way of accumulating capital is by being self-employed. It may also be the only way to find short-term employment. In the end however, it will only be profitable for a temporary immigrant to enter self-employment if laws, rules and financing are conducive.

The self-employment choice is dependent on how factors influence the individual's real labour market situation. In the framework described above, a person with a relatively long unemployment history, a comparative advantage and an unlimited amount of capital would become self-employed.

# 3 Data description

The data consists of a stratified sample taken from the Social Research Register (SFR). The sample includes all first and second generation male immigrants and 7 pct. of ethnic Danish men. The sample years are from 1989 to 97 and includes men between 15 and 67 years of age. Over the nine year period, more than 200.000 individuals are studied. All individuals not connected directly to the labour force (i.e. people receiving pensions, students, people in the agricultural sector, self-employed individuals with many employees, and people on sabbatical) are excluded. [5]

The data set contains valuable information, but because of insufficient income information the idea of estimating a structural model is rejected. In the context of income tax avoidance, it is more difficult to find "true" information about self-employment income than it is to find information about wage-employment income. The limited income information is the reason for choosing a reduced model when investigating the self-employment choice.

The three labour market states mentioned earlier are defined as follows; a **full time self-employed** agent is a person whose main occupation is self-employment, with less than five employees, working in the non-agricultural sector and with a gross annual unemployment rate of less than 25 pct.[6] A **full-time wage-earner** must have a main occupation as a wage-earner (with and without limited income) and a gross annual unemployment rate of less than 25 pct. An **unemployed** individual has a gross annual unemployment rate of at least 25 pct.

The focus of this paper is on people entering self-employment. However, the stock of self-employed ethnic Danes and immigrants seems quite stable in the period 1989 to 97. Annually the vast majority of self-employed individuals do not change occupations (see Fig

[5] Some of the self-employed agents are excluded because of the differences in start up costs and subsidies etc.

[6] Gross unemployment refers to the number of days in a year that a person is net unemployed, sick or in active labour market programmes minus the number of days the person receives subsidies for starting up a new business. Net unemployment refers to insured people who receive unemployment benefits and people who receive cash assistance (from municipalities). Non-insured unemployed individuals and unemployed insured individuals who haven't got the right to receive unemployment insurance for various reasons are also included.

3). The stock of self-employed individuals seem to be less stable among immigrants than among ethnic Danes. In 1997, 85 pct. of self-employed ethnic Danes were also self-employed in 1996. The corresponding share for immigrants was only 75 pct. It is interesting that among ethnic Danes it is mostly wage-earners that enter self-employment, while among immigrants, it is mostly unemployed men that enter self-employment.

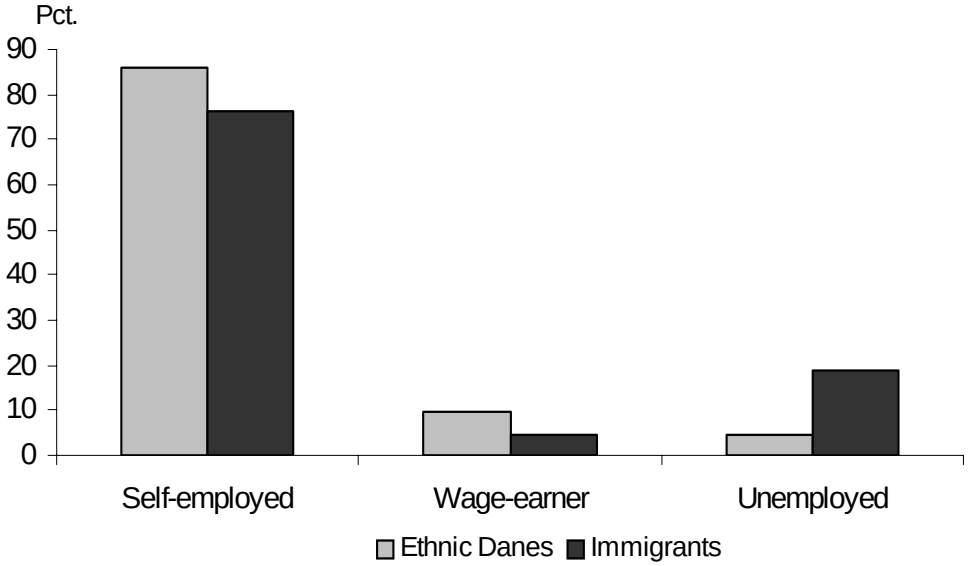


Fig. 3. Self-employed in 1997 and occupation in 1996, males in Denmark aged 15-67.

Fig. 4. shows that entry rates into self-employment from unemployment and from wage-employment are higher among immigrants than among ethnic Danes (unemployed individuals having the highest entry rates). Interestingly self-employed immigrants are also the most likely to exit into unemployment, while ethnic Danes are more likely to exit into wage-employment (see Fig. 5). Entering self-employment does not necessarily seem to give an immigrant steady employment. The exit-rate from self-employment into unemployment declines during economic growth and the exit-rate into wage-employment increases substantially among immigrants. Comparing Fig. 1, 4 and 5, the increase in the self-employment rate among immigrants might be explained by the increase in the entry rate into self-employment among unemployed immigrants from 1994 to 97.[7]

[7]The strong increase in entries among immigrants could be due to an increase in self-employment assistant programmes.

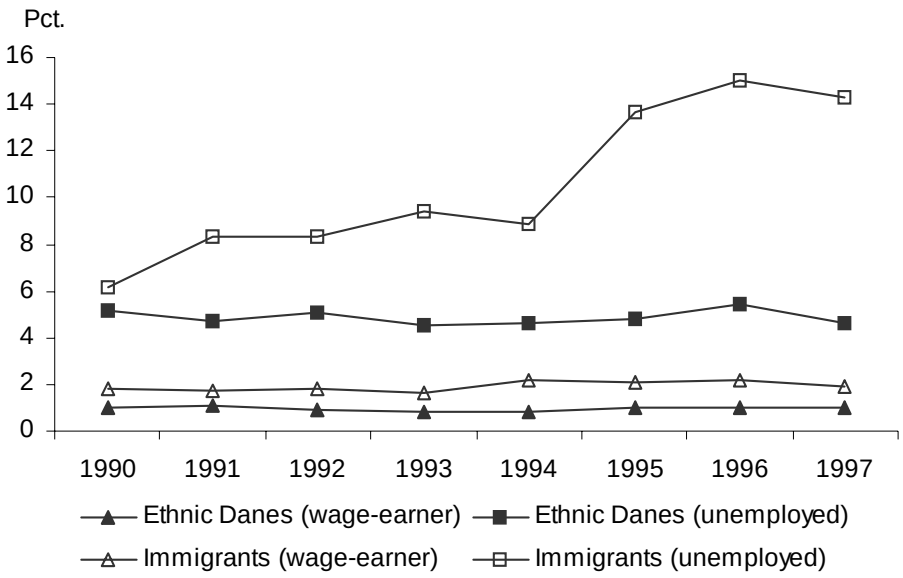


Fig. 4. Self-employment entry rates from wage-earner jobs and unemployment. Males in Denmark, aged 15-67.

Looking at the descriptive statistics in table 1, unemployed immigrants entering into self-employment are on average more likely to have a partner, be well-educated and be wealthier than other unemployed immigrants. They also seem to have experienced high individual unemployment rates within the last year. The unemployed Danes entering into self-employment are characterised by being older, having fewer children, more education, less unemployment and greater wealth.

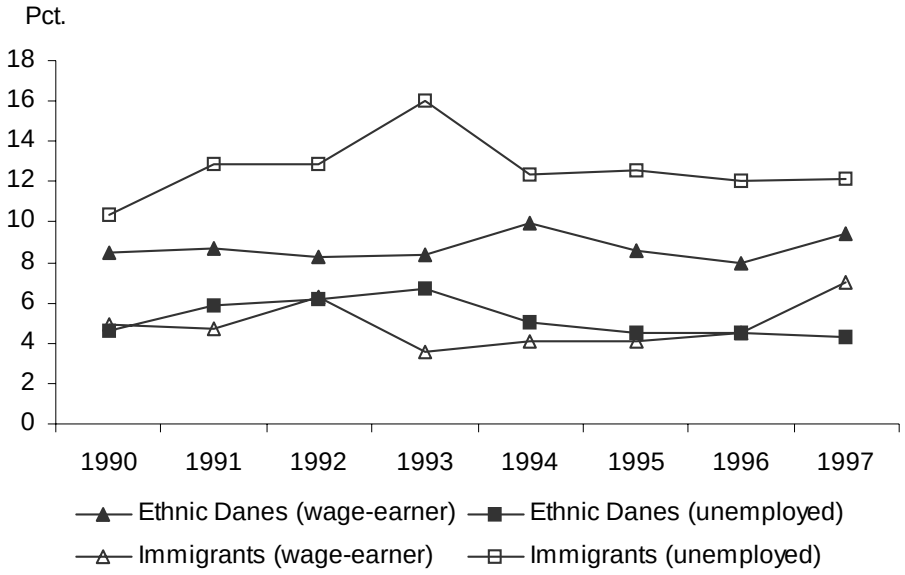


Fig. 5. Self-employment exit rates to wage-earner jobs and unemployment. Males in Denmark, aged 15-67.

**Table 1. Summary statistics for the stratified sample of unemployed males aged 15-67.**

| | | All | U-S | U-W | U-U |
|---|---|---|---|---|---|
| Age (year) | Danes | | 39,33 (10,86) | 34,42 (10,81) | 37,06 (11,54) |
| | Immigrants | | 32,17 (7,99) | 32,17 (8,32) | 32,84 (9,30) |
| | All | 38,64 (11,15) | | | |
| Partner (partner=1) | Danes | | 0,658 (0,474) | 0,540 (0,498) | 0,499 (0,500) |
| | Immigrants | | 0,710 (0,454) | 0,644 (0,479) | 0,683 (0,465) |
| | All | 0,717 (0,451) | | | |
| Children (child=1) | Danes | | 0,969 (4,983) | 0,677 (3,409) | 0,632 (3,784) |
| | Immigrants | | 1,767 (2,831) | 1,520 (2,913) | 1,774 (3,505) |
| | All | 0,957 (2,516) | | | |
| Education (short=0,long=2) | Danes | | 0,793 (0,674) | 0,716 (0,670) | 0,630 (0,655) |
| | Immigrants | | 0,545 (0,579) | 0,391 (0,604) | 0,358 (0,560) |
| | All | 0,822 (0,713) | | | |
| Unemployment (promil) | Danes | | 498,6 (188,7) | 459,4 (171,4) | 544,7 (192,0) |
| | Immigrants | | 555,9 (195,2) | 526,244 (190,0) | 577,1 (193,3) |
| | All | 90,02 (197,9) | | | |
| Wealth (20.000 dkr) | Danes | | 1,753 (24,96) | 1,106 (14,55) | 1,511 (16,05) |
| | Immigrants | | 0,0157 (9,680) | -0,321 (6,553) | -0,189 (5,425) |
| | All | 6,429 (289,0) | | | |

Note: U= unemployed, W=wage-earner, S=self-employed. ( ) indicate standard errors.

Among wage-earners, ethnic Danes entering into self-employment have generally experienced very little unemployment within the last year, have high levels of education and greater wealth than other males in the sample (see Table 2.). Ethnic Danes entering self-employment are older, have fewer children, are more often educated, wealthier and have a shorter unemployment history than immigrant wage-earners.

In general, male immigrants are younger, less educated and poorer than ethnic Danes. They are also more likely to have a partner, children, and longer unemployment spells. The descriptive statistics indicate that self-employed immigrants have different characteristics than self-employed ethnic Danes.

**Table 2. Summary statistics for the stratified sample of male wage-earners aged 15-67.**

| | | All | W-S | W-W | W-U |
|---|---|---|---|---|---|
| Age (year) | Danes | | 38,38 (10,51) | 36,33 (11,89) | 35,89 (12,21) |
| | Immigrants | | 32,91 (8,80) | 31,97 (9,43) | 31,39 (9,32) |
| | All | 38,64 (11,15) | | | |
| Partner (partner=1) | Danes | | 0,738 (0,440) | 0,650 (0,477) | 0,529 (0,499) |
| | Immigrants | | 0,709 (0,454) | 0,644 (0,479) | 0,629 (0,483) |
| | All | 0,717 (0,451) | | | |
| Children (child=1) | Danes | | 0,852 (1,156) | 0,745 (1,760) | 0,635 (2,769) |
| | Immigrants | | 2,057 (5,583) | 1,481 (3,155) | 1,557 (3,457) |
| | All | 0,957 (2,516) | | | |
| Education (short=0,long=2) | Danes | | 0,941 (0,714) | 0,857 (0,719) | 0,654 (0,633) |
| | Immigrants | | 0,401 (0,613) | 0,404 (0,629) | 0,354 (0,542) |
| | All | 0,822 (0,713) | | | |
| Unemployment (promil) | Danes | | 17,12 (45,86) | 23,24 (54,36) | 70,02 (83,21) |
| | Immigrants | | 37,16 (67,65) | 35,35 (65,48) | 68,01 (83,20) |
| | All | 90,02 (197,9) | | | |
| Wealth (20.000 dkr) | Danes | | 13,00 (184,5) | 6,606 (281,4) | 2,395 (17,16) |
| | Immigrants | | 0,777 (14,09) | 0,709 (10,71) | -0,042 (4,206) |
| | All | 6,429 (289,0) | | | |

Note: U= unemployed, W=wage-earner, S=self-employed. ( ) indicate standard errors.

Continue next page...

# 4 Empirical model of entering self-employment

This paper analyses occupational choice by estimating the probability of specific individuals entering a certain occupation.

**Transition into self-employment: a conditioned binomial probit model**

A random utility model capture the idea that self-employment is chosen when utility is higher relative to the satisfaction received in alternative occupations. The estimation measures the probability of individual $i$, choosing self-employment in period $t+1$ given that he was a wage-earner or unemployed in period $t$.

$$\Pr ob\left[(I_{i,t+1}=1 \mid I_{i,t}=0), x_{i,t}\right] = \Pr ob[(u^{s}_{i,t+1} > u^{a}_{i,t+1} \mid u^{s}_{i,t} < u^{a}_{i,t}), x_{i,t}] \tag{6}$$

where $I$ is an observable index. $I$ is 1 if the individual is self-employed and 0 if the individual has an alternative occupation.

The transition probability of entering into self-employment given that the individual wasn't self-employed in the previous period has often been simplified by following a group of wage-earners at time $t$, and then monitoring the transition pattern of these agents at time $t+1$. The probability of choosing self-employment can then be analysed by the simple model described in equation (7),

$$\Pr ob\left[I_{i,t+1}=1 \mid x_{i,t}\right] = \Pr ob\left[\beta' x_{i,t} + \epsilon_i > 0 \mid x_{i,t}\right] \tag{7}$$

An example of the simple binomial probit model is visualized in Table 3. The transition matrix shows who is included and how they are included in the simple model (6). The example shows how the probability of entering self-employment instead of wage-employment, given the person was unemployed, is estimated.

**Table 3. Transition matrix for the simple binomial probit model.**

| | | **t+1** | | |
|---|---|---|---|---|
| | | **Unemployed** | **Wage-earner** | **Self-employed** |
| **t** | **Unemployed** | not included | $I_{i,t+1}=0$ | $I_{i,t+1}=1$ |
| | **Wage-earner** | not included | not included | not included |
| | **Self-employed** | not included | not included | not included |

Note: Same notation as in equation (6).

**Advantages and disadvantages of a simple model**

The simple conditioned binomial model is easy to estimate. It is however, important to be aware of the problems connected to the model. The related problems concern the disadvantages of analysing a reduced form estimation, endogenity, selection bias, ignoring individual specific effects and state dependency.

The parameter estimates from a reduced estimation can indicate a positive or negative correlation between the explanatory variable and the dependent variable. No structural relationships can be examined. The probit model assumes that the explanatory variables are exogenous, but in reality this may be incorrect. The choice of self-employment may determine education and not the reverse, which would cause endogeneity. The endogeneity problem is partially solved by using lagged variables.

Not taking account of individual specific ability (or preferences) can be problematic. This type of mis-specification can also cause biased estimates. The unobserved heterogeneity will be investigated through a random effect model.[8] The problem of ignoring state dependency is similar to not including explanatory variables. Therefore not taking account of occupational choices made before the present choice creates biased estimates.

The assumption about data having a random design is also critical because the employment choice is restricted to two occupations, given that the individual had a certain occupation in the year before (see Table 3). The empirical model is therefore suffering from the initial condition problem. In this paper, the method suggested by Bruce (2000) and Orme (1997) is used to correct for sample selection.[9]

[8] Random effect model, see appendix 8.3.

[9] Maximum-likelihood probit estimation with selection, see appendix 8.4.

# 5 Results and discussion

The characteristics of people entering self-employment in Denmark are based on the results from the conditioned binomial probit model. The study does not concentrate on previously self-employed individuals who continue in self-employment. The characteristics of people who are self-employed and the people who enter self-employment could overlap, but not necessarily.

The empirical results from the estimated probability of entering self-employment in preference to an alternative occupation, given that the individual had a certain occupation in the period before, is shown in Tables 4 and 5. The simple probit model that takes individual specific effects into account and the binomial probit model that takes the initial condition problem into account are displayed for the transition from unemployment into self-employment in Tables 4a and 4b.

The results in Tables 4a and 4b indicate that the impact of initial conditions on the parameter estimates are not always significant. Taking sample selection into account when estimating the probability of unemployed individuals entering self-employment rather than becoming a wage-earner does not change the results significantly. On the other hand, the parameter estimates that measure the probability of entering self-employment for wage-earners changes significantly, when taking the initial condition into account (see Table 5 and Appendix 8.5). Therefore, conclusions on the results will concern the binomial probit estimations that take sample selection into account.

**Table 4a. Probability of entering self-employment instead of wage-employment for unemployed individuals.**

| Variables | Simple Probit | Random effect | Probit with Sample selection |
|---|---|---|---|
| Intercept | -2,558 (0,083)*** | -3,849 (0,160)*** | -2,508 (0,226)*** |
| Ethnic Dane | r.c. | r.c | r.c |
| Immigrant f.g. | 0,478 (0,020)*** | 0,734 (0,037)*** | 0,468 (0,047)*** |
| Immigrant s.g. | 0,603 (0,109)*** | 0,919 (0,167)*** | 0,601 (0,107)*** |
| Age | 0,015 (0,001)*** | 0,022 (0,002)*** | 0,015 (0,002)*** |
| Single children=0 | r.c. | r.c. | r.c |
| Single children $\geq$1 | 0,159 (0,031)*** | 0,232 (0,047)*** | 0,159 (0,032)*** |
| Couple children=0 | 0,193 (0,029)*** | 0,286 (0,047)*** | 0,197 (0,033)*** |
| Couple children $\leq$2 | 0,254 (0,021)*** | 0,381 (0,034)*** | 0,257 (0,026)*** |
| Couple children >2 | 0,340 (0,029)*** | 0,513 (0,047)*** | 0,344 (0,032)*** |
| Basic school | r.c. | r.c. | r.c. |
| Vocational education | 0,140 (0,018)*** | 0,219 (0,028)*** | 0,141 (0,019)*** |
| Short advanced edu. + | 0,065 (0,030)** | 0,100 (0,046)** | 0,071 (0,038)* |
| Construction | r.c. | r.c. | r.c. |
| Wholesale | 0,647 (0,031)*** | 0,940 (0,052)*** | 0,647 (0,031)*** |
| Retail | 1,327 (0,029)*** | 1,932 (0,063)*** | 1,326 (0,028)*** |
| Restaurant & hotel | 1,276 (0,030)*** | 1,871 (0,065)*** | 1,275 (0,030)*** |
| Bank & finance | 0,708 (0,031)*** | 1,031 (0,053)*** | 0,708 (0,032)*** |
| Amusement | 0,186 (0,022)*** | 0,271 (0,034)*** | 0,185 (0,023)*** |
| Unemployment | 0,000487 (0,000269)* | 0,0009 (0,0004)** | 0,000487 (0,000270)* |
| Unemployment in 2. | -2,98e-08 (2,34e-07) | -1,35e-07 (3,48e-07) | -2,99e-08 (2,37e-07) |
| Wealth | -0,0002 (0,0005) | -0,0002 (0,0009) | -0,0002 (0,0007) |
| 1990 | r.c. | r.c. | r.c. |
| 1991 | 0,077 (0,039)** | 0,103 (0,057)* | 0,078 (0,039)** |
| 1992 | 0,124 (0,038)*** | 0,179 (0,055)*** | 0,123 (0,038)*** |
| 1993 | 0,105 (0,038)*** | 0,157 (0,055)*** | 0,104 (0,038)*** |
| 1994 | -0,013 (0,036) | -0,027 (0,053) | -0,019 (0,041) |
| 1995 | -0,023 (0,035) | -0,031 (0,052) | -0,029 (0,042) |
| 1996 | 0,009 (0,035) | 0,004 (0,052) | 0,006 (0,039) |
| 1997 | -0,148 (0,035)*** | -0,223 (0,052)*** | -0,153 (0,042)*** |
| Panel level variance | | 0,234 (0,095)*** | |
| Selection | | | -0,027 (0,123) |
| N | 39093 | 39093 | 39093 |

Note:* Statistically significant at the 10% level, ** 5% level and *** 1% level.( ) indicate standard errors.

**Table 4b. Probit estimation of the initial condition for being unemployed and either entering self-employment or wage-employment.**

| Variables | Simple Probit |
|---|---|
| Intercept | -1,113 (0,012)*** |
| Ethnic Dane | r.c. |
| Immigrant f.g. | 0,384 (0,006)*** |
| Immigrant s.g. | 0,048 (0,041) |
| Age | -0,014 (0,0002)*** |
| No partner | r.c. |
| Partner | -0,180 (0,006)*** |
| Children | -0,001 (0,001) |
| Basic school | r.c |
| Vocational education | -0,065 (0,006)*** |
| Short advanced edu. + | -0,219 (0,008)*** |
| 1990 | r.c. |
| 1991 | -0,020 (0,011)* |
| 1992 | 0,021 (0,011)* |
| 1993 | 0,017 (0,011) |
| 1994 | 0,182 (0,010)*** |
| 1995 | 0,217 (0,010)*** |
| 1996 | 0,139 (0,010) |
| 1997 | 0,218 (0,010)*** |
| N | 716357 |

Note:* Statistically significant at the 10% level, ** 5% level and *** 1% level.( ) indicate standard errors.

**Table 5. Probability of entering self-employment. Estimates of binomial probit models that control for sample selection.**

| | U $\longrightarrow$ S or U | W $\longrightarrow$ S or U | W $\longrightarrow$ S or W |
|---|---|---|---|
| Intercept | -1,869 (0,140)*** | -2,123 (0,024)*** | -2,494 (0,033)*** |
| Ethnic Dane | r.c. | r.c | r.c |
| Immigrant f.g. | 0,229 (0,055)*** | 0,064 (0,013)*** | -0,015 (0,014) |
| Immigrant s.g. | 0,240 (0,095)** | 0,095 (0,075) | 0,148 (0,075)** |
| Age | 0,002 (0,0008)*** | -0,007 (0,0005)*** | -0,0009 (0,0007) |
| Single children=0 | r.c. | r.c | r.c |
| Single children $\geq$1 | 0,132 (0,027)*** | 0,031 (0,019) | 0,018 (0,020) |
| Couple children=0 | 0,138 (0,031)*** | -0,018 (0,016) | 0,005 (0,017) |
| Couple children $\leq$2 | 0,216 (0,027)*** | 0,147 (0,012)*** | 0,154 (0,013)*** |
| Couple children >2 | 0,277 (0,031)*** | 0,196 (0,019)*** | 0,203 (0,020)*** |
| Basic school | r.c | r.c | r.c. |
| Vocational education | 0,160 (0,021)*** | -0,019 (0,011) | 0,019 (0,011)* |
| Short advanced edu. + | 0,186 (0,043)*** | -0,006 (0,017) | 0,073 (0,015)*** |
| Construction | r.c. | r.c | r.c. |
| Wholesale | 0,602 (0,029)*** | 0,185 (0,016)*** | 0,175 (0,016)*** |
| Retail | 1,132 (0,029)*** | 0,200 (0,019)*** | 0,172 (0,021)*** |
| Restaurant & hotel | 0,885 (0,027)*** | 0,254 (0,022)*** | 0,303 (0,025)*** |
| Bank & finance | 0,602 (0,029)*** | 0,241 (0,016)*** | 0,177 (0,017)*** |
| Amusement | -0,020 (0,020) | 0,143 (0,011)*** | 0,046 (0,012)*** |
| Unemployment | -0,0012 (0,0002)*** | -0,006 (0,0003)*** | -0,003 (0,0003)*** |
| Unemployment in 2. | 5,32e-07 (1,99e-07)*** | 0,00002 (1,3e-06)*** | 6,51e-06 (1,48e-06)*** |
| Wealth | -0,00008 (0,0006) | 0,0009 (0,00015)*** | -9,69e-07 (6,57e-06) |
| 1990 | r.c. | r.c | r.c. |
| 1991 | -0,034 (0,033) | 0,015 (0,019) | 0,115 (0,022)*** |
| 1992 | -0,038 (0,034) | -0,036 (0,020)* | 0,0,83 (0,023)*** |
| 1993 | -0,054 (0,035) | -0,013 (0,020) | 0,103 (0,024)*** |
| 1994 | -0,015 (0,034) | -0,031 (0,020) | 0,074 (0,023)*** |
| 1995 | 0,123 (0,031)*** | 0,010 (0,020) | 0,126 (0,023)*** |
| 1996 | 0,120 (0,031)*** | 0,012 (0,019) | 0,122 (0,023)*** |
| 1997 | 0,131 (0,031)*** | -0,006 (0,020) | 0,125 (0,023)*** |
| Selection | 0,214 (0,088)*** | 1,83 (0,078)*** | 1,29 (0,053)*** |
| N | 68707 | 43134 | 112854 |

Note:* Statistically significant at the 10% level, ** 5% level and *** 1% level. ( ) indicate standard errors.

There is clear indication that ethnicity plays an important role when choosing self-employment, given that the person is either unemployed or a wage-earner. First and second generation male immigrants from less developed countries choose self-employment at higher frequencies than ethnic Danes. This behavior among immigrants is also found

in studies by Borjas (1986) and Yuengert (1995) in America.

The results support the hypothesis which states that a person who has experienced long-term unemployment has an incentive to choose self-employment. However, it is important to note that the probability of a wage-earner entering self-employment only increases if the unemployment rate is above 15 pct. within the past year. At this point immigrants should enter self-employment at a higher rate than ethnic Danes, because they seem to experience a higher risk of becoming unemployed.

The significant results on education do not support the hypothesis regarding unskilled individuals entering self-employment because of a high risk of long-term unemployment. Having a high education actually increases the likelihood of entering self-employment among unemployed individuals. This result is to some extent in agreement with the results of Fairlie and Meyer (1996), who find that immigrants with good wage possibilities as wage-earners, have a relatively high probability of becoming self-employed. Among wage-earners, education doesn't seem to play an important role in the transition into self-employment. The results with regard to education and historical unemployment indicate that discrimination might be a better explanation of the occupational choice.

A comparative advantage in the form of having many children and an assisting partner seems to increase the probability of entering self-employment in preference to alternative occupations. The same results are discovered by Borjas (1986). The present study shows that having more than one child and a partner is important to the self-employment transition for wage-earners. At the same time, age has a negative effect on the likelihood of entering into self-employment, but the age effect is positive among unemployed individuals. Immigrant men from less developed countries generally have more children and a higher marital rate. Judging by the parameter estimates, immigrants should enter self-employment at higher frequencies than ethnic Danes.

The results support Yuengert's "home-country" effect, because immigrants - compared to ethnic Danes - have a higher probability of choosing self-employment. However, immigrant wage-earners do not behave significantly different from ethnic Danes regarding all occupational choices. Compared to ethnic Danes, second generation immigrant wage-earners

don't seem to have a significantly higher probability of entering into self-employment. The results do not necessarily indicate that being an immigrant is a comparative advantage. It may actually be a comparative disadvantage or a matter of discrimination, since immigrant wage-earners do not enter into self-employment at a very high frequency relative to ethnic Danes.

The random effects model can indicate whether or not an unmeasured skill or preference has influenced the transition into self-employment. The individual specific effect is only taken into account when examining an unemployed individuals decision between wage-employment and self-employment. The non-significant sample selection and the significant individual specific effect might indicate that an unmeasurable self-employment ability or preference does influence the transition into self-employment. It is important to note, that taking individual specific effects into account does not change parameter estimates.

Having past work experience from sectors outside the construction business, increases the likelihood of entering into self-employment. Experience related to sectors like retail, restaurants and hotels are positively correlated with self-employment.

Among wage-earners especially, wealth is found to be significantly different from zero. This may be due to higher risk aversion among wage-earners than among unemployed individuals, because wage-earner have a job and a steady income. The year dummies may have caught some business cycle effects or assimilation effects, but most of the parameters are insignificant. Changes in institutional factors need to be examined before concluding anything here.

The empirical results indicate that the large proportion of self-employed immigrants can partly be explained by long-term unemployment and most certainly by comparative advantages including family status and ethnic background.

# 6 Conclusion

First and second generation ethnicity seems to increase the likelihood of entering into self-employment. This is especially true for men with partners and children. Having a relatively long unemployment history within the last year makes this transition into self-employment even more likely.

The analysis has many dimensions because the probability of entering into self-employment instead of wage-employment, and the probability of entering into self-employment instead of unemployment is estimated. It is therefore important to pinpoint that the empirical results also indicate that male wage-earners with no historical unemployment and a fortune also have a high probability of entering into self-employment.

In addition, the empirical results show that education, income, initial resources, work experience, sector knowledge, business cycles, institutional factors and individual specific effects influence the occupational choice of self-employment. From the results it seems more reasonable to conclude that an interaction of many factors affect the self-employment decision.

Continue next page...

# 7 References

**Alba-Ramirez, 1994**, Self-employment in the midst of unemployment: the case of Spain and the United States, Applied Economics 26, 189-204.

**Bager, T. and Rezai, S. 1999**, Immigrant business in Denmark: The importance of local, national and international linkages, EGOS Colloquium, July 4th-6th 1999, Birmingham, United Kingdom, Subtheme 3: Business Systems in their International Contexts.

**Baltagi, B., 1995**, Econometric Analysis of Panel Data. New York: John Wiley and Sons.

**Blanchflower, D. G. and A. J. Oswald, 1998**, What Makes an Entrepreneur ? Journal of Labor Economics, 16(1), 26-60.

**Blanchflower, D.G., 2000**, Self-employment in OECD Countries, Labour Economics, 7(5), 471-506.

**Blau, D., 1987**, A Time Series Analysis of Self-Employment in the United States , Journal of Political Economy 95, 445-467.

**Borjas, G. J., 1986**, The self-employment experience of Immigrants, Journal of Human Resources, 21(4), 486-506.

**Borjas, G. J. and S. G. Bronars, 1989**, Consumer Discrimination and Self-Employment, Journal of Political Economy, 97(3), 581-605.

**Bradbury, K.L., 1994**, New England Job Changes During the Recession: The Role of Self-Employment, New England Economic Review 0(0), 45-60.

**Bruce, D., 2000**, Effects of the United States tax system on transitions into self-employment, Labour Economics 7(5), 545-574.

**Carrasco, R., 1999**, Transition to and from Self-Employment in Spain: An Empirical Analysis, Oxford Bulletin of Economics and Statistics 61(3), 315-342.

**Carrasco, R. and M. Ejrnæs, 2000**, Self-Employment in Denmark and Spain: Institutions, Economic Conditions and Gender Differences, Working paper, University of Copenhagen.

**Chapman, B., B. Gregory and J. Klugman, 1998**, Self-Employment in Australia: Description, Analysis and Policy Issues, for OECD International Conference on Self-employment.

**Clark, K and S. Drinkwater, 2000**, Pushed Out or Pulled In ? Self-Employment Among Ethnic Minorities in England and Wales, Labour Economics 7(5), 603-628.

**Cowling, M. and P. Mitchell, 1997**, The Evolution of UK Self-Employment: A Study of Government Policy and the Role of Macroeconomy, The Manchester School of Economic and Social Studies 65(4), 427-42.

**Dennis, W. ,1996**, Self-Employment When Nothing Else is Available , Journal of Labor Research, 17(4), 645-61.

**Ejrnæs, M. ,1999**, Paneldata, Undervisningsnoter til Øvelse i Paneldata, Copenhagen University.

**Ejrnæs, M., K. B. Jensen, H. S. Nielsen and A. Würtz, 2001**, Self-employment for Immigrants: A last resort ? , Work in progress !

**Evans, D.S and B. Jovanovic, 1989**, An Estimated Model of Entrepreneurial Choice under Liquidity Constraints, Journal of Political Economy 97(4), 774-806.

**Evans, D. S. and L. S. Leighton, 1989**, Some Empirical Aspects of Entrepreneurship, The American Economic Review 79(4), 519-535.

**Fairlie, R. W. and B. D. Meyer, 1996**, Ethnic and Racial Self-Employment Differences and Possible Explanations, Journal of Human Resources 31(4), 757-793.

**Green, W.H., 1997**, Econometric Analysis. 3rd ed. Prentice Hall.

**Hout, M. and H.S Rosen, 1999**, Self-employment, Family Background and Race, NBER Working Paper, W7344.

**Høgelund, J. , K. Langager, L. Jakobsen and L. Jagd, 1992**, Iværksætterydelsen, SFI-rapport 92:3.

**Kuhn, P. and H. Schuetze, 1998,** The Self-Employment Dynamics of Men and Women in Canada: 1982-1995, McMaster University Department of Economics working paper 99-05.

**Lindh, T. and H. Ohlson, 1996**, Self-employment and Windfall Gains: Evidence from the Swedish Lottery, Economic Journal, 106, 1515-1526.

**Moore, C.S. and R.E. Mueller, 1998**, The Transition into Self-Employment in Canada: The Importance of Involuntary Separation and Unemployment Duration, OECD CNA Corp.

**Moore, 1983**, Employer Discrimination: Evidence from Self-Employed Workers , The Review of Economics and Statistics, 65(3), 496-501.

**Orme, C., 2000**, Two-Step Inference in Dynamic Non-Linear Panel Data Models, Working paper from School of Economic Studies University of Manchester.

**Plougman, P. and S. Buhl, 1998**, Self-employment in Denmark - Trends and Policy, Danish Technological Institute & Centre for Industrial Analysis for OECD.

**Schuetze, H.J., 2000**, Taxes, economic conditions and recent trends in male self-employment: a Canada-US comparison, Labour Economics 7(5), 507-544.

**Stata Press, 1999,** Stata Reference Manual, Release 7, 2(H-P), College Station, Texas.

**Taylor, M., 1996**, Earnings, Independence or Unemployment: Why Become Self-employed ?, Oxford Bulletin of Economics and Statistics, 58(2), 253-265.

**Taylor, M., 1998**, Self-employment and windfall gains in Britain : Evidence from panel data, Discussion paper from Institute for Labour Research and Institute for Social and Economic Research 98/25.

**Wong, G., G. H. Henson and C. Riddell**, Earnings impact of Self-Employment Assistance for The Canadian Unemployed 1987- 1996, Working paper

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

# 8 Appendix

## 8.1 Definition on immigrants

Immigrants are individuals born in a foreign country, whose mother and/or father were either also born in a foreign country or are citizens of another country. This includes Danish citizens born abroad.

The following rules are applied in deciding the country of origin:

1. If none of the parents are known, the country of origin is defined by the individuals own information. If the individual is an immigrant, then the country of origin is equal to the country of birth. If the individual is a direct descendant, then country of origin is equal to the country of citizenship.

2. If the background of only one of the parents is known, then country of origin is defined by the parent's country of birth. If Denmark is the country of birth, then the country of citizenship is used instead.

3. If both parents are known, then the country of origin is defined by the mother's country of birth or citizenship.

The definition of immigrants also include second generation immigrants. Second generation immigrants include sons and daughters of foreign nationals born in Denmark.

## 8.2 Definition of "other" Asia

Other Asia = Georgia, Armenia, Azerbaijan, Lebanon, Syria, Irak, Israel, West Bank, Jordan, Saudi Arabia, Kuwait, Bahrain, Quatar, United Arabic Republic, Oman, Yemen, Kazakistan, Turkiistan, Uzbekistan, Tadztikestan, Kyrgyzstan, Afghanistan, India, Bangladesh, Maldive Islands, Sri Lanka, Nepal, Bhutan, Myanmar, Thailand, Laos, Vietnam, Cambodia, Indonesia, Malaysia, Brunei, Singapore, Philippines, Mongolia, China, North Korea, South Korea, Taiwan, Hong Kong, Macau.

## 8.3 Random effect model

The problem of individual effects can be explained mathematically by expression (8)[10]

$$\Pr ob\left[I_{it}=1\right]=\Pr ob\left[\beta' x_{it}+\epsilon_{it}>0\right] \tag{8}$$

A problem occurs if the model is estimated without controlling for variations in time and on individuals. Therefore the disturbance term consists of more parts; $\epsilon_{it}=\mu_i+\nu_{it}$ and these parts are not accounted for. The complete likelihood consists of more than just the product of the marginal likelihood-functions because $E(\epsilon_{it},\epsilon_{it+1})=\sigma_\mu^2$. Therefore the estimates will be biased if there is no control for unobserved heterogeneity.

By using a random effects model on panel data, individual specific effects can be taken into account. The model assumes that $\epsilon_{it}=\mu_i+\nu_{it}$, where $\mu_i \backsim IID(0,\sigma_\mu^2)$ and $\nu_{it} \backsim (0,\sigma_\nu^2)$ are stochastic disturbance terms which are independent of each other and $x$ is a vector of explanatory variables.

If the random effect, $\mu_i$, is assumed to be distributed normal, $N(0,\sigma_\mu^2)$, then the probability of occupational choice can be estimated by the following expression [11]

$$\Pr ob(I_i \mid x_i)=\int_{-\infty}^{\infty}\frac{e^{\frac{-\mu_i^2}{2\sigma_\mu^2}}}{\sqrt{2\pi\sigma_\mu}}\left[\prod_{t=1}^{n_i}F(\beta x_{it}+\mu_i)\right]d\mu_i \tag{9}$$

where

$$F(\beta x_{it}+\mu_i)=\begin{array}{ll}\frac{1}{1+\exp(\beta x_{it}+\mu_i)} & \text{if} \quad I_{it}\neq 0\\ 1-\frac{1}{1+\exp(\beta x_{it}+\mu_i)} & \text{else}\end{array}$$

[10] Created from Balthagi (1995) and Ejrnæs (1999).

[11] In stata the intergral is approximated by a M-points Gauss-Hermite quadrature

$$\int_{-\infty}^{\infty}e^{-x^2}f(x)dx\approx\sum_{m=1}^{M}w_m^*f(a_m^*)$$

where $w_m^*$ is quadrature weights and $a_m^*$ is quadrature abscissas. Then the log-likelihood is

$$L=\sum_{i=1}^{n}w_i\log(\Pr ob(I_{it}=1))\approx\sum_{i=1}^{n}w_i\log\frac{1}{\sqrt{\pi}}\sum_{m=1}^{M}w_m^*\prod_{t=1}^{n_i}F(\beta x_{it}+\sqrt{2\frac{\rho}{1-\rho}}a_m^*$$

where $\rho=\frac{\sigma_\mu^2}{\sigma_\mu^2+1}$

## 8.4 Maximum-likelihood probit estimation controlling for initial condition

The problem with sample selection can be explained mathematically by the expression below [12]

$$y_i^* = \beta' x_i + \epsilon_{1i}$$

The probit equation is created by assuming that the above relationship exists. The problem is that the dependent variable is only observed if

$$\begin{aligned} y_i^* &= (\theta' z_i + \epsilon_{2i} > 0) \\ where \quad \epsilon_1 &\sim N(0,1) \\ \epsilon_2 &\sim N(0,1) \\ corr(\epsilon_1, \epsilon_2) &= \rho \end{aligned}$$

when $\rho \neq 0$ a standard probit estimation will yield biased estimates. The following log likelihood takes sample selection into account

$$\begin{aligned} L = & \sum_{\substack{j \in observed \; y_j \\ y_j \neq 0}} w_j \ln \left[ \Phi_2 \left( \beta' x_j + offset_j^{\beta}, \theta' z_j + offset_j^{\theta}, \rho \right) \right] \\ & + \sum_{\substack{j \in observed \; y_j \\ y_j = 0}} w_j \ln \left[ \Phi_2 \left( -\beta' x_j + offset_j^{\beta}, \theta' z_j + offset_j^{\theta}, -\rho \right) \right] \\ & \sum_{\substack{j \notin observed \; y_j \\ y_j \neq 0}} w_j \ln \left[ 1 - \Phi \left( \theta' z_j + offset_j^{\theta} \right) \right] \end{aligned}$$

$$\begin{aligned} where \quad \Phi_2 () &\sim cumulative \; bivariate \; normal \; distribution \; function \\ \Phi () &\sim s \tan dard \; cumulative \; normal \\ w_j &\sim weight \end{aligned}$$

[12] Created from Green (1997) and Stata (1999).

## 8.5 Simple probit estimation for a wage-earners transition into self-employment

**Table 6. Probability of entering self-employment instead of wage-employment for wage-earners.**

| | Simple Probit |
|---|---|
| Intercept | -1,181 (0,040)*** |
| Ethnic Dane | r.c. |
| Immigrant f.g. | -0,312 (0,020)*** |
| Immigrant s.g. | -0,231 (0,110)** |
| Age | 0,00008 (0,0009) |
| Single children=0 | r.c. |
| Single children $\geq$1 | 0,082 (0,031)*** |
| Couple children=0 | 0,151 (0,027)*** |
| Couple children $\leq$2 | 0,463 (0,020)*** |
| Couple children >2 | 0,549 (0,031)*** |
| Basic school | r.c |
| Vocational education | 0,119 (0,018)*** |
| Short advanced edu. + | 0,506 (0,027)*** |
| Construction | r.c. |
| Wholesale | 0,341 (0,028)*** |
| Retail | 0,357 (0,035)*** |
| Restaurant & hotel | 0,456 (0,036)*** |
| Bank & finance | 0,451 (0,030)*** |
| Amusement | 0,252 (0,020)*** |
| Unemployment (1 y) | -0,010 (0,0004)*** |
| Unemployment in 2 (1 y) | 0,00003 (2,06e-06)*** |
| Wealth | 0,002 (0,0003)*** |
| 1990 | r.c. |
| 1991 | -0,081 (0,032)** |
| 1992 | -0,149 (0,033)*** |
| 1993 | -0,205 (0,030)*** |
| 1994 | -0,060 (0,033)* |
| 1995 | 0,106 (0,033)*** |
| 1996 | 0,020 (0,032) |
| 1997 | 0,117 (0,033)*** |
| N | 43134 |

Note:* Statistically significant at the 10% level, ** 5% level and *** 1% level. ( ) indicate standard errors.