Catalogue No. 98-16

ECONOMETRIC ANALYSIS OF CANADIAN SELF-EMPLOYMENT USING SLID

Product Registration Number 75F0002M

November 1998

Wayne Simpson[1], Department of Economics, University of Manitoba

Robert Sproule[2], Department of Economics, Bishop's University

The Income and Labour Dynamics Working Paper Series is intended to document detailed studies and important decisions for the Income and Labour Dynamics program. It is a continuation of the SLID Research Paper Series. These working papers are available in English and French. To obtain a summary description of available documents or to obtain a copy of any, please contact the Client Services Unit, at 7-B5 Jean Talon Building, Statistics Canada, Ottawa, Ontario, CANADA K1A 0T6, by INTERNET (DYNAMICS@STATCAN.CA), by telephone (613) 951-7355 or toll-free 1-888-297-7355, or by fax (613) 951-3012.

---

[1] Department of Economics, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada, simpson@cc.umanitoba.ca

[2] Department of Economics, Bishop's University, Lennoxville, Québec, J1M 1Z7, Canada

EXECUTIVE SUMMARY

Self-employment has grown in importance in recent years. This paper estimates a structural model of self-employment using recently released data from Statistics Canada's *Survey of Labour and Income Dynamics*. In addition to providing a first set of results for Canada, the paper provides a more comprehensive analysis of the self-employment decision than earlier studies. We find that the self-employment behavior of men and women is quite distinct and merits separate assessment. In particular, while both men and women appear to be quite sensitive to the relative earnings opportunities for wage- and self-employment, they respond very differently to local unemployment conditions. Adjustment for sample selection bias arising from both participation and self-employment is found to be significant in earnings equations for both men and women.

TABLE OF CONTENTS

1. INTRODUCTION

Self-employment, often ignored by researchers for reasons of analytical convenience or data limitations, is an important component of the modern labour market. Aronson (1991), for example, reports that in the United States self-employment declined during the early part of this century, but that since 1970 it has "re-emerged," reaching 8.8% of total American nonfarm employment. Aronson also reports comparable, or larger, proportions of nonfarm workers in self-employment across Europe, ranging from 7.5% in the United Kingdom and Denmark to 26.5% in Greece.[1]

In Canada, self-employment has risen at a robust annual rate of 3.6% since 1976, and the proportion of workers who are self-employed has risen from 11% to 16% [Human Resources Development Canada (1998, p. 14)]. What is curious about this recent growth of self-employment is that it has occurred during a period in which paid-employment growth has been sluggish, averaging only 1.4% per annum, and in which the gap between Canadian and U.S. unemployment rates has risen to an historic high [Riddell and Sharpe (1998)].

This recent upswing in the ranks of the self-employed in Canada and elsewhere raises numerous questions. What forces have given rise to this growth? What kinds of workers become and remain self-employed? Is the growth in self-employment the result of inadequate opportunities for paid employment? If not, what factors can account for it? Are the self-employed pushed into this labour-force activity, or were they drawn towards it? What are the consequences of self-employment decisions for earnings?

---

[1] Some descriptive studies of self employment by country include: (a) Canada: Cohen (1988), Tepper (1988), and Human Resources Development Canada (1998); (b) the United States: Becker (1983), and Bregger (1996); and (c) the United Kingdom: Creigh et al. (1986), Daly (1991), and Hakim (1988). For a descriptive study of self employment in Europe, see Loufi (1991).

Although analysis of aggregate time-series data may provide an interesting perspective on some of the questions regarding self-employment posed above, we would argue that much more can be learned about self-employment behaviour from the analysis of household microdata. The objective of this paper is to present an analysis of the determinants of self-employment in Canada using recently released cross-sectional data from Statistics Canada's *Survey of Labour and Income Dynamics* (hereafter SLID) and a three-step procedure to estimate a structural model of the self-employment decision. The paper contributes to the literature on the self-employment decision and the earnings of workers in self- and wage-employment in several ways. First, we provide what we believe to be the first set of results for a structural econometric model of self-employment decisions for Canada. In this respect, the paper adds to a small literature for a few countries which we discuss below in the context of our results. Secondly, we provide a more comprehensive assessment of the self-employment decision than earlier studies. Previous studies have generally ignored sample selection bias arising from labour force participation in the determination of earnings opportunities for self- and wage-employment. Previous studies have also concentrated on men for the most part, and no previous study has estimated models for men and women separately. Our paper addresses both these issues. Finally, our paper examines the effect of local unemployment conditions on self-employment activity, providing new results for men and the first results to our knowledge for women. Our results suggest that these effects are quite different for men and women and require separate consideration.

The paper is organized as follows. In Section 2, we explain the framework for the analysis of self-employment behavior, including the structural model and the estimation procedure. In Section 3, we discuss our data source and present our empirical results. We then compare our results to the literature in this area. In Section 4 we present our conclusions.

2. THE ANALYTICAL FRAMEWORK

At the core of the literature on labour force participation and self-employment, and of the present study, is the notion that the typical individual chooses between labour market alternatives on the basis of an evaluation of the expected utility of these alternatives. While this economic approach is well established in the analysis of labour force participation [e.g., the review by Killingsworth (1983)], it has only recently been applied to the choice between self-employment and paid employment [e.g., Dolton and Makepeace (1990), Pfeiffer and Pohlmeier (1992), Rees and Shah (1986), and Taylor (1996)].

In this study, we assume that individual choice follows in a two-step procedure. In the first step, the individual chooses to participate in the labour force or not. If a decision in favour of labour force participation is made in the first stage then, in the second stage, we assume that the individual must choose between wage-employment and self-employment on the basis of his/her expected utility valuations. In particular, after the expected-utility-valuation models of Rees and Shah (1986) and Pfeiffer and Pohlmeier (1992)), we postulate that the choice for individual i between two alternative activities (j=1,2) will depend on an equation of the form:

$$\begin{cases} I_i = 1 (j = 1) \; if \;\; I_i^* = g + a\,(E[Y_{1i}] - E[Y_{2i}]) + dX_i + n_i > 0 \\ \quad = 0 (j = 2) \; if \;\; I_i^* \le 0 \end{cases} \qquad (1)$$

where I_i^* is a decision variable or, in econometric terms, a latent variable which depends on the difference in the expected returns to each activity, $E[Y_{1i}] - E[Y_{2i}]$, plus other socioeconomic characteristics, X_i , which affect utility comparisons across individuals and thus the choices individuals make.

Since only one alternative, and hence only one return, is observed, the model also needs to specify how the returns available from each activity will vary across individuals. A common starting point is a lognormal earnings equation for each activity of the form

$$Y_j = E[Y_{ji}] + e_{ji} = b_j' Z_{ji} + e_{ji}, \quad j = 1,2 \qquad (2)$$

where Z_{ji} is a vector of characteristics for individual i in state j that determine the expected logarithm of earnings, Y_{ji} , and ε_{ji} is a normally distributed disturbance term with mean zero and constant variance σ_j^2. Given the prospective returns from Equation (2), individual i decides, on the basis of the latent index in Equation (1), (a) between labour-force participation and non-participation and, if the individual decides to participate, (b) between self-employment and wage-employment.

The econometric model based on this sequential discrete choice framework can be stated as follows. First, we specify an equation set to describe the earnings from participation, either self-employed or wage-employed, and the shadow earnings from nonparticipation in accordance with Equation (2):

$$\begin{cases} Y_{Wi} = Z_i^W b^W + x_i^W \\ Y_{Si} = Z_i^S b^S + x_i^S \\ Y_{Ni} = Z_i^N b^N + x_i^N \end{cases} \qquad (3)$$

where Y_{ji} is the logarithm of earnings from wage-employment (j=W) or self-employment (j=S) or the shadow earnings from nonparticipation (j=N) for individual i and Z_i^j represents the determinants of wage-employed earnings, self-employed earnings or shadow earnings derived from not working (j=W, S, and N, respectively). Errors are again assumed to be normally distributed with mean

zero and constant variance σ_j^2. We then specify an equation set to explain the choices between non-work and work and between wage-employment and self-employment in accordance with Equation (1):

$$\begin{cases} P_i^* = g_P + a_P\{\max[E(Y_{Wi}), E(Y_{Si})] - E(Y_{Ni})\} + d_P X_i^P + x_i^P = V_i^P p_P + x_i^P \\ S_i^* = g_S + a_S\{E(Y_{Si}) - E(Y_{Wi})\} + d_S X_i^S + x_i^S = V_i^S p_S + x_i^S \end{cases} \quad (4)$$

where P_i^* and S_i^* are the latent variables which determine the choices between non-work and work and between wage- and self-employment, respectively, for individual i given individual characteristics that affect expected earnings in alternative activities and the other characteristics that affect these choices (X_i^P and X_i^S, respectively). Finally, we specify that we observe earnings Y_i only if an individual works ($P_i^* > 0$), and that these earnings are derived for ind either from self-employment (when $S_i^* > 0$) or paid employment (when $S_i^* \# 0$):

$$Y_i = \begin{cases} Y_{Si} \ \textit{if} \ \ P_i^* > 0 \ \textit{and} \ S_i^* > 0 \\ Y_{Wi} \ \textit{if} \ \ P_i^* > 0 \ \textit{and} \ S_i^* \le 0 \end{cases} \quad (5)$$

This econometric model is a fairly standard mover-stayer model [e.g., Borjas and Rosen (1980)] which can be estimated by maximum likelihood or, more readily, in three stages in the following sequence. The first step provides consistent estimates of the reduced-form participation and self-employment models based on Equation set (4) using maximum likelihood (probit) estimation. The coefficient estimates for γ_P, α_P, and δ_P and γ_S, α_S, and δ_S (combined as π_P and π_S, respectively, in Equation (4) to simplify notation) can be used to estimate the inverse Mills ratio terms for participation and self-employment, respectively, in standard fashion [Heckman (1979)]. In the second step, the earnings equations for wage-employment and self-employment can be estimated consistently by including

the two inverse Mills ratio terms in the earnings equations (Equation set (3) for Y_{Wi} and Y_{Si}) to correct for selection bias arising from the nonrandom decisions associated with participation and self-employment.[2] Using (5), these earnings equations become

$$Y_{Si} = Z_i^S b^S + l_{Pi}(P_i^* > 0) + l_{Si}(S_i^* > 0) + \mathcal{l}_i^S$$
$$Y_{Wi} = Z_i^W b^W + l_{Pi}(P_i^* > 0) + l_{Si}(S_i^* \le 0) + \mathcal{l}_i^W,$$

$$\textit{where} \quad l_{Pi}(P_i^* > 0) = \frac{f(V_i^P \hat{p}_P / \hat{s}_P)}{\Phi(V_i^P \hat{p}_P / \hat{s}_P)},$$

$$l_{Si}(S_i^* > 0) = \frac{f(V_i^P \hat{p}_S / \hat{s}_S)}{\Phi(V_i^P \hat{p}_S / \hat{s}_S)},$$

$$\textit{and} \quad l_{Si}(S_i^* \le 0) = \frac{-f(V_i^S \hat{p}_S / \hat{s}_S)}{1 - \Phi(V_i^S \hat{p}_S / \hat{s}_S)} \qquad (6)$$

where the coefficients used to calculate λ_{Pi} and l_{Si} are estimated in the first step.

In the third step, a structural model of self-employment [the equation for S_i^* in (4)] can be estimated consistently using the predicted earnings difference between wage-employed and self-employed workers obtained from the earnings equation estimates [Equation set (6)].

While the econometric model provides useful information on various aspects of labour market behavior, our primary focus is the structural model obtained in the final estimation step and how economic incentives affect the self-employment decision. This leads us to focus on the effect of earnings

[2] We make a common simplifying assumption that the selection disturbance terms ξ_i^P and ξ_i^S are uncorrelated (e.g., Catsiapis and Robinson, 1982; Hubler, 1989; Macpherson, 1988).

opportunities. The model provides a direct estimate of the influence of (estimated) individual earnings opportunities on the choice between wage- and self-employment. Economic theory contains the strong prediction that workers are drawn to self-employment when it offers superior earnings opportunities; i.e., that $a_S > 0$ in Equation (4).

In addition, however, we also focus on the effect of the local unemployment rate on the self-employment decision. A variety of recent research has suggested that labour market conditions matter, perhaps by affecting the likelihood of receiving a wage-employment offer.[3] What we term the "unemployment push" hypothesis suggests that high unemployment levels result in fewer wage-employment offers and that many workers in this situation may prefer self-employment to long periods without work [Taylor (1996, p. 253)]. To paraphrase Dennis (1996), self-employment may represent for some the only viable option "when nothing else is available." By implication, the "unemployment push" hypothesis predicts that the self-employment rate rises during economic downturns, or that the self-employment rate is counter-cyclical [Meager (1992, p. 89)]. Examples of research supporting this theory include Acs et al. (1994), Bogenhold and Staber (1991), and Evans and Leighton (1989).

Alternatively, the "prosperity pull" hypothesis argues that when unemployment is low, wage-employment offers are frequent. In this situation, workers may take the self-employment option knowing that if the self-employment venture fails another paid job will not be far in the offing [Taylor (1996, p. 253)]. By implication, the "prosperity pull" hypothesis predicts that the self-employment rate follows the economic cycle, or is pro-cyclical [Meager (1992, p. 89)].

---

[3] In our model, we estimate the expected earnings offers in self- and wage-employment, but workers may be more likely to receive the estimated average wage-employment offer, or more, in good labour market conditions, i.e. when the unemployment rate is low.

Examples of research supporting this theory include Blanchflower and Oswald (1991).[4]

3. DATA AND EMPIRICAL RESULTS

In this section, we present our econometric results and compare them with previous research. We begin with a few comments about the data sample.

3.1. The Data

Unlike its predecessors, particularly the Labour Market Activity Survey from 1986 to 1990, Statistics Canada's *Survey of Labour and Income Dynamics* (SLID) provides detailed labour market activity and income information for self-employed individuals as well as workers in paid employment. Table 1 presents the variables drawn from the SLID that are used in this study. More precisely, it presents their definitions and their mean values for self-employed and wage-employed men and women. Individuals are classified as self-employed or wage-employed on the basis of a direct question related to their current or most recent job. The sample includes all respondents under the age of 65 with complete records for all the variables found in Table 1.

[4] The "prosperity pull" hypothesis may be seen as a special case of the hypothesis that individuals are drawn to self employment because the pecuniary and non-pecuniary net benefits of self-employment exceed those of wage employment, where the non-pecuniary benefits of self-employment are tantamount to the "lifestyle" benefits of self-employment [e.g., increased independence, flexibility, and ability to work at home]. The import of the "lifestyle" benefits of self-employment is echoed in a recent paper by Mallett (1991, p. 16). In an analysis of survey data, he observes that money is not the only motivator for the self-employed in Canada.

Table 1: Sample Statistics for Men and Women, Self-Employed and Wage-Employed

Variable [1]	Definition	Men		Women	
		Self-Employed	Wage-Employed	Self-Employed	Wage-Employed
PTC?	Positive earnings in 1994?	70.2%	69.3%	57.2%	53.6%
AGE	Age in years	42.7	38.6	39.8	39.3
YRSCH	Years of Schooling	12.7	12.5	13	12.5
HS?	High school diploma?	60.2%	61.6%	70.4%	66.6%
PS?	Post-secondary diploma?	37.3%	32.7%	38.7%	35.5%
UNIV?	University degree?	16.6%	13.8%	16.6%	12.1%
EDEXCAN?	Educated outside Canada?	8.6%	6.1%	7.5%	6.2%
EVERFT?	Ever worked full time?	99.3%	94.6%	92.7%	88.4%
PCTFT	% of career in full time employment	91.8%	82.2%	64.5%	65.7%
URATE [2]	Unemployment rate	10.7%	11.5%	11.0%	11.4%
VISMN?	Visible minority?	3.1%	3.5%	2.8%	3.4%
ABRG?	Aboriginal?	1.5%	2.7%	3.1%	3.0%
IMMG?	Immigrant?	12.2%	9.1%	10.1%	8.9%
DISAB?	Disabled?	4.5%	8.3%	6.2%	7.6%
STUDENT?	Student in 1994?	6.2%	17.0%	14.7%	17.1%
MRRD?	Married or common law partnership?	82.0%	66.5%	75.3%	%67.6
PRSCHL?	Children under 6 year present?	17.5%	16.0%	18.3%	17.4%
CHLDN?	Children 6-19 years present?	50.7%	44.8%	52.9%	46.4%
ENG?	English mother tongue?	66.4%	66.0%	69.5%	66.5%
FR?	French mother tongue?	19.1%	25.6%	19.7%	24.5%
HSF?	Father completed high school?	19.6%	24.3%	23.6%	22.6%

Variable [1]	Definition	Men		Women	
		Self-Employed	Wage-Employed	Self-Employed	Wage-Employed
UNIF?	Father completed university degree?	7.5%	6.7%	8.1%	6.8%
HSM?	Mother completed high school?	30.0%	31.2%	33.1%	30.7%
UNIM?	Mother completed university degree?	3.5%	4.0%	5.8%	3.3%
ATL?	Resides in Atlantic Canada?	14.8%	22.5%	18.9%	21.5%
QUE?	Resides in Quebec?	16.3%	21.5%	16.1%	21.0%
ONT?	Resides in Ontario?	25.6%	26.0%	25.0%	27.1%
PR?	Resides in Prairie Provinces?	33.1%	21.9%	29.2%	21.7%
BC?	Resides in British Columbia?	10.2%	8.1%	10.8%	8.6%
EARN [3]	Earnings in 1994?	$27,023	$31,796	$16,519	$19,407
HRWK [3]	Hours worked per week	49.6	40.4	33.9	32.3
WKS [3]	Weeks worked	51.4	46.5	48.6	45.8
YRXFT [3]	Work Experience (Full-time equivalent in years)	22.9	16.8	13	11.2
n	Number of Observations	1087	7675	1218	8221

Notes: (1) "?" denotes dummy variable (yes=1; no=0). (2) The unemployment rate is the provincial unemployment rate taken as of June, 1994. (3) Figures reported are means for those with positive earnings; 324 of 1,087 self-employed men have no earnings, hours or weeks in 1994 and YRXFT is not computed; same applies for 2,357 of 7,675 employed men, for 521 of 1,218 self-employed women, and for 3,811 of 8,221 employed women.

Source: Survey of Labour and Income Dynamics Public File 1994.

Table 1 suggests several patterns. The proportion of self-employed men and women is about the same, just under 15% of the sample, which is consistent

with other estimates [Human Resources Development Canada (1998, p. 14)]. On balance, self-employed females (males) are better educated than the wage-employed females (males). This result appears to be inconsistent with the screening hypothesis, which predicts that an unscreened group (the self-employed) will acquire less education to the extent that it is a screening, rather than a skill enhancing, device [Wolpin, 1977]. Married females (or males) with children present are more likely to be self-employed. Finally, we would note that the average self-employed female (and male) works longer hours per week, works a greater number of weeks during the year, and has more full-time work experience, than her (his) wage-employed counterpart. To assess the effect of these and other patterns in the data on wage opportunities and the self-employment decision, we turn to the econometric model.

3.2. Empirical Results

The estimation procedure used is a three-step process outlined in Section 2. The first step is the estimation of reduced-form probit models [Equation (4)] for the labour-force participation decision (P_i^*) and the self-employment decision (S_i^*) for both men and women. The reduced-form estimates, reported in Table 2, show the effect of individual characteristics on the probabilities of labour-force participation and self-employment but are not easily interpreted in terms of the analytical framework described above. The results are used primarily to construct individual estimates of the inverse Mills-ratio terms, λ_{Pi} and λ_{Si}, to correct for bias in the earnings equations (6) estimated in the second step. One result we note is that the local unemployment rate, whose effect we discuss in more detail later, appears to have quite different effects for men and women. Although the effect is negative in both cases, consistent with the prosperity pull hypothesis, it is much weaker and insignificant for women. Moreover, this result does not control

directly for the effect of differing wage opportunities on the self-employment decision as does the structural model estimated below.

Table 2: Reduced-Form Probit Regression Results for Participation and Self-Employment Decisions

[t-values in parentheses]

Variable	Participation? (1)		Self-employed? (2)	
	Men	Women	Men	Women
Constant	-1.070 (-4.596)	-1.477 (-7.789)	-2.040 (-7.354)	-1.583 (-5.114)
AGE	0.067 (5.891)	0.064 (6.802)	0.079 (5.980)	0.013 (0.845)
AGESQ	-0.001 (-9.356)	-0.001 (-10.873)	-0.001 (-4.464)	-0.000 (0.271)
VISMN	-0.294 (-2.492)	-0.011 (-0.111)	-0.384 (2.846)	-0.289 (-1.785)
ABRG	-0.357 (-3.297)	-0.282 (-3.242)	-0.250 (-1.764)	0.140 (1.027)
IMMG	0.206 (1.598)	0.178 (1.822)	-0.025 (-1.764)	-0.080 (-0.575)
DISAB	-1.243 (-21.393)	-0.844 (-14.695)	-0.074 (-0.755)	0.327 (2.883)
ATL	-0.065 (-1.109)	-0.218 (-4.834)		
QUE	-0.193 (-2.256)	-0.273 (-4.180)		
PR	0.123 (2.051)	0.097 (2.142)		
BC	0.193 (2.254)	0.137 (2.214)		
EVERFT	1.121 (12.702)	1.118 (20.920)		
YRSCH	0.050 (7.816)	0.083 (14.361)		
STUDENT	-0.347 (-4.962)	0.073 (1.327)	-0.285 (-3.609)	-0.083 (-1.075)
EDEXCAN	-0.179 (-1.306)	-0.113 (-1.064)	-0.020 (-0.160)	0.015 (0.098)
URATE			-0.041 (-5.881)	-0.013 (-1.604)
PS			0.010 (0.242)	-0.065 (-1.279)
HS			-0.103 (-2.270)	-0.161 (-2.819)
UNIV			0.017 (0.301)	0.068 (0.978)
MRRD	0.386 (6.742)	-0.0560 (-1.510)	0.039 (0.664)	0.136 (2.207)

Variable	Participation? (1)				Self-employed? (2)			
	Men		Women		Men		Women	
PRSCHL	-0.240	(-3.452)	-0.654	(-14.835)	0.132	(-3.878)	0.136	(2.030)
CHLDN	0.120	(2.556)	-0.229	(-6.514)	0.015	(2.344)	0.222	(4.341)
ENG	0.072	(0.828)	0.032	(0.490)	-0.313	(-3.878)	-0.192	(-1.933)
FR	0.163	(1.470)	0.112	(1.331)	-0.476	(-5.352)	-0.369	(-3.319)
HSF					-0.050	(-0.813)	0.011	(0.824)
UNIVF					0.084	(0.910)	0.097	(0.972)
HSM					0.014	(0.275)	0.083	(1.401)
UNIVM					-0.006	(-0.051)	0.478	(4.078)
$\chi 2$	1903.44		2518.4		426.67		164.63	
n	8762		9439		6194		5289	
% Correctly Predicted	88.9%		79.5%		82.7%		89.1%	

(1) Dependent variable is 1 if respondent participates (has positive earnings), and 0 otherwise.

(2) Dependent variable is 1 if worker is self-employed, and 0 if worker is wage-employed.

The second step estimates the earnings Equation set (6) for self-employed and wage-employed individuals for men and women. The estimates of λ_P and λ_S are included to correct for selection bias, and GLS is used to correct for heteroscedasticity arising from the inclusion of estimated, rather than actual, inverse Mills ratio terms [White (1980)]. These GLS estimates are reported in Table 3. They show the effect of individual characteristics on the earnings available to an individual if (s)he were self-employed or wage-employed.

Table 3: Earnings Equation Regression Results Corrected for Sample Selection Bias Arising from Participation and Self-Employment [t-values in parentheses; dependent variable is the log of pre-tax annual earnings for 1994; the estimation method uses Heckman's (1979) technique to adjust for selectivity bias, and White's (1980) method to produce heteroskedasticity-consistent standard errors.]

Variable	Men		Women	
	Self-Employed	Wage-Employed	Self-Employed	Wage-Employed
Constant	6.727 (7.186)	6.716 (47.810)	6.895 (14.340)	6.288 (53.420)
VISMN	-0.182 (-0.739)	-0.156 (-2.586)	-0.121 (-0.492)	0.116 (1.771)
ABRG	0.051 (0.192)	0.059 (1.043)	-0.020 (-0.065)	-0.069 (-0.841)
IMMG	0.226 (1.097)	0.041 (0.737)	-0.070 (-0.284)	0.079 (0.156)
DISAB	0.142 (0.433)	0.034 (0.384)	-0.243 (-0.985)	-0.194 (-0.242)
STUDENT	-0.105 (-0.468)	-0.141 (-4.553)	-0.131 (-1.068)	-0.209 (-6.488)
EDEXCAN	-0.197 (-0.893)	0.076 (1.306)	-0.091 (-0.300)	-0.133 (-1.937)
ATL	-0.172 (-1.180)	-0.143 (-5.550)	-0.267 (-2..273)	-0.194 (-6.524)
QUE	-0.059 (-0.352)	-0.054 (-1.512)	-0.056 (-0.309)	-0.062 (-1.605)
PR	-0.184 (-1.918)	-0.094 (-3.796)	-0.054 (-0.5120)	-0.124 (-4.564)
BC	-0.042 (-0.302)	0.050 (1.524)	0.149 (0.969)	0.010 (0.234)
EVERFT	1.076 (2.143)	0.500 (5.015)	0.406 (1.359)	0.176 (1.916)
PCTFT	0.701 (2.624)	0.314 (7.126)	0.471 (3.296)	0.299 (6.951)
YRXFT	0.017 (0.731)	0.041 (12.420)	0.036 (2.671)	0.038 (9.417)
YRXSQ	-0.000 (-0.718)	-0.001 (-7.254)	-0.001 (-2.108)	-0.001 (-5.288)
HS	0.211 (2.429)	0.160 (7.256)	0.280 (2.677)	0.196 (7.320)
PS	0.183 (2.282)	0.114 (6.587)	0.159 (1.934)	0.148 (7.052)
UNIV	0.474 (3.655)	0.343 (14.450)	0.519 (4.708)	0.395 (12.610)
HRSWK	-0.010 (4.413)	0.010 (9.277)	0.014 (4.146)	0.017 (6.722)
WKSEM	0.014 (1.976)	0.033 (31.100)	0.024 (5.339)	0.036 (34.280)
ENG	0.152 (0.922)	0.077 (1.782)	-0.052 (-0.299)	0.115 (2.376)
FR	0.027 (0.113)	0.096 (1.880)	-0.234 (-1.076)	0.127 (2.257)
λ_S	-0234 (-0.530)	-0.472 (-14.190)	-0.417 (-11.600)	-0.830 (-3.864)

Variable	Men		Women	
	Self-Employed	Wage-Employed	Self-Employed	Wage-Employed
λ_P	-0.427 (-1.023)	-0.588 (-4.111)	-0.265 (-0.968)	-0.386 (-4.711)
R^2	0.22	0.656	0.386	0.639
F	9.061	438.264	18.418	337.801
n	763	5318	697	4410

The results provide some insight into the earnings opportunities in self-employment and wage-employment available to men and women with differing characteristics. For example, education is just as valuable, or more valuable, to self-employed workers, consistent again with the idea that education enhances productivity rather than serving as a screen for other worker attributes. The estimated returns to high school completion, a postsecondary diploma, or a university degree are higher for self-employed men and women than their wage-employed counterparts. The returns to work experience, on the other hand, are mixed.

There is some evidence that self-employment may be advantageous for women who speak neither official language. Table 3 indicates that women in paid employment receive a premium of about 12% if either English or French is their mother tongue, but this premium disappears (in fact, it is negative but insignificant) for self-employed women. No such pattern is observed for men, perhaps reflecting the different opportunities for wage- and self-employment available to men and women. Other related variables, such as visible minority status and immigration status, show no clear pattern.

The terms which test for selection bias arising from participation and self-employment, λ_{Pi} and λ_{Si}, respectively, are significant in six of the eight cases, implying that OLS estimates would be biased. In particular, we would note that the term associated with participation is significant for wage-employment (but not

for self-employment) for men and women, indicating that studies which ignore participation decisions in assessing earnings, even for men, risk biased estimates. Since the earnings regression results in Table 3 are typically used to estimate the structural model of self-employment decisions, as we do below, any bias arising from the omission of the correction for participation behaviour would be transmitted to the estimates of the structural model.

Finally, we note that the R^2 is much lower for the self-employed earnings equations for men and women. Since mean earnings for self-employment and wage-employment are similar (see Table 1), this implies that the variance of the estimate is much larger for the self-employed than for the wage-employed–0.95 vs. 0.37 for men and 1.05 vs. 0.42 for women. These results are consistent with the idea that the self-employment option offers greater risk than the wage-employment option (Rees and Shah, 1986).

The third and final step involves the estimation of the structural-probit model corresponding to Equation (4). This model uses the estimated difference in expected log earnings between self-employment and wage-employment, ΔEARN in Table 4, derived from the estimates in Table 3. The estimated coefficient on this variable corresponds to α_S in Equation (4), which measures the effect of the earnings difference on the latent index, S_i^*, which in turn guides the self-employment decision, other factors (X_i^S) considered. In Table 4, we present two sets of estimates. First, we present what we term the unrestricted estimates, which include a large list of additional factors–including age to capture life cycle effects, family status, many of the other variables which appeared in the earnings equations, and the local unemployment rate. Secondly, we present what we term the restricted estimates, which include only the local unemployment rate as an additional factor. In this way, we can gauge the robustness of our findings with respect to ΔEARN and the unemployment rate to alternative specifications of X_i^S.

Table 4: Structural Probit Model of the Self-Employment Decision [The dependent variable is 1 if self-employed and 0 if wage employed; t-values are in parentheses.]

Variable	Men		Women	
	Unrestricted	Restricted	Unrestricted	Restricted
Constant	-5.844 (-15.195)	0.060 (0.713)	-12.706 (-9.538)	-4.237 (-18.384)
ΔEARN (1)	4.834 (39.050)	1.974 (26.814)	13.048 (14.859)	6.441 (24.558)
URATE	-0.142 (-14.384)	-0.076 (-10.280)	0.033 (1.297)	0.031 (2.207)
AGE	0.337 (17.771)		0.099 (2.008)	
AGESQ	-0.003 (-15.285)		-0.000 (-0.154)	
VISMN	-0.929 (-4.962)		2.101 (2.181)	
ABRG	-0.518 (-2.685)		-0.588 (-1.214)	
IMMG	-1.004 (-6.752)		2.008 (4.485)	
DISAB	-1.550 (-10.988)		2.997 (8.408)	
STUDENT	-0.948 (-8.912)		-1.262 (-4.887)	
EDEXCAN	1.198 (7.193)		-1.135 (-2.363)	
HS	-0.255 (-4.296)		-1.413 (-7.118)	
PS	-0287 (-5.067)		-0.369 (-2.163)	
UNIV	-0.476 (-6.200)		-0.422 (-1.664)	
MRRD	0.286 (3.688)		0.431 (1.994)	
PRSCHL	0.160 (2.132)		-0.373 (-1.671)	
CHLDN	0.120 (2.107)		0.603 (3.291)	
ENG	-0.769 (-7.485)		1.343 (4.355)	
FR	-0.690 (-5.978)		3.042 (7.532)	
Log-likelihood Function	-1531.6	-2417.7	-175.604	-471.27
n	6081	6081	5107	5107
% Correct Predicted	88.9%	79.7%	98.5%	96.2% .

(1) ΔEARN is the earnings difference between self-employment and paid employment predicted from individual characteristics and the earnings equation results in Table 3.

The effect of ΔEARN on the self-employment decision is positive and significant as predicted by economic theory. The estimated coefficients are quite large, indicating a considerable amount of sensitivity in the self-employment decision to the earnings opportunities in self- and wage-employment. Put differently, the evidence suggests that Canadian men and women with characteristics more generously rewarded by self-employment are those most likely to be drawn into self-employment, other factors considered.

The effect of the local unemployment rate differs dramatically between men and women. For men, lower unemployment rates associated with better labour market conditions result in a higher incidence of self-employment, other factors considered (including relative personal rewards in self- and wage-employment captured by ΔEARN). This result is consistent with the "prosperity pull" hypothesis. For women, on the other hand, lower unemployment rates are associated with a lower incidence of self-employment, which is consistent with the "unemployment push" hypothesis. This effect is only significant in the restricted model, however.

These conclusions, with the exception of the unemployment rate effect for women, hold for both the unrestricted and restricted models. In fact, the estimated effect of ΔEARN is larger for both men and women when additional factors are considered, as is the effect of the unemployment rate for men.

For the unrestricted model, there is a clear life cycle effect represented by AGE and its square, implying that older workers are more likely to engage in self-employment, although the effect is much weaker for women.[5] Moreover, there are some notable differences in the characteristics of self-employed men and women.

[5] This age effect exhibits a concave pattern for men, with a positive effect up to 51 years of age and a negative effect thereafter. For women, however, the effect increases steadily with age.

Visible minority and immigrant men are less likely to be self-employed than their non-minority, native born counterparts; for women these groups are more likely to be self-employed. In households with preschool children, men are more likely to be self-employed whereas women are less likely to be self-employed. And men whose mother tongue is one of Canada's official languages are less likely to be self-employed (or more likely to engage in paid employment) whereas women in these circumstances are more likely to engage in self-employment, other factors considered. In short, there is considerable evidence that the self-employment behaviour of men and women is very different and should be analyzed separately.

3.3. Our Results In The Context of Previous Research:

Only a small number of papers have examined the earnings differentials between self- and wage-employment using the structure described in Equations (6). Of these, an even smaller number have examined the self-employment decision using the structural model described in Equation (4). The studies, in order of publication date, are: Rees and Shah (1986), Macpherson (1988), de Wit and van Winden (1989), Evans and Leighton (1989), Dolton and Makepeace (1990), Fujii and Hawley (1991), Pfeiffer and Pohlmeier (1992), and Taylor (1996). The distinguishing features of these eight studies are collected in Table 5.

Table 5: An Overview To Previous Research

Article	Country	Sample	The Wage Equations				The Structural Probit Equation	
			-1	-2	-3	-4	-5	-6
Rees and Shah (1986)	UK	Males 15-65	+	+,*	na	Self-Employed Equation: Occupation, location Wage-Employed Equation: Education, age, health, marital status, race, occupation	+	Significant Variables: Education, age, health Insignificant Variables: Marital status, race, children
Macpherson (1988)	US	Married females 25-64	+	-,*	+,*	Self-Employed Equation: Education, children Wage-Employed Equation: Education, work experience, race, children	na	
de Wit and van Winden (1989)	Holland	1983 follow up of Males in the sixth grade in 1952	+	+	na	Self-Employed Equation: parent(s) absent?, sex Wage-Employed Equation: scholastic aptitude, education, graduated highest education level, additional training, experience, sex, industry	+	Significant Variables: Employ-ment status of father, self-employment status of father, experience, sex, industry Insignificant Variables: Scholastic aptitude, education, graduated?, highest education level?, religion
Evans and Leighton (1989)	US	Males 14-24 in 1966	ns	ns	na	Self-Employed Equation: Disability, urban, business experience, education, urban Wage-Employed Equation: Marital status, urban, disability, wage experience, education, duration of unemployment, occupation, urban	na	
Dolton and Makepeace (1990)	UK	Recent male and female graduates and diplom-ates in 1980	-	+,*	na	Self-Employed Equation: Class of degree Wage-Employed Equation: Class of degree, school type, A-level score, months unemployed , months working, occupational status, wage, number of jobs, part-time status, university-awarded degree?, MSc?, PhD?, professional qualification?, gender, ethnic group	+	Significant Variables: Age, sex, salary, social class, children, school type, occupational status Insignificant Variables: Marital status, class of degree, A-level score, university-awarded degree?, MSc?, PhD?, professional qualification?, gender, ethnic group
Fujii and Hawley (1991)	US	Males 25-55	+,*	+,*	na	Self-Employed Equation: Education, experience, assets, job tenure, wife's income Wag-Employed Equation: Education, assets experience	+,*	Significant Variables: Education, experience, wife's income, assets Insignificant Variables: Race

Article	Country	Sample	The Wage Equations				The Structural Probit Equation	
			-1	-2	-3	-4	-5	-6
Pfeiffer and Pohlmeier (1992)	West Germany	Males	+	+	na	Self-Employed Equation: Duration of education, duration of vocation training, marital status, disability, self-employment status of father, in agriculture? Wage-Employed Equation: The "employed equation" is not reported.	+	Significant Variables: Age, marital status, size of community, job requires vocational training, vocational or professional degree required?, disability, self-employment status of father, job in agriculture? Insignificant Variables: Unemployment rate, duration of education
Taylor (1996)	UK	Full-time males	+,*	-	na	Self-Employed Equation: Age, occupation, industry, education Wage-Employed Equation: Age, marital status, children , ethnic minority, occupation, industry, education, degree, higher or other qualification, region	+,*	Significant Variables: Marital status, parent self-employed, housing equity, occupation, job attributes Insignificant Variables: Age, children, unemployment-vacancy ratio, education, higher or other qualification(s), region

Legend: Column (1) presents sign ("+" and "-") and significance ("*" if significant) for the term used to correct for bias in wage equation for self-employed arising from selection of self-employment and employment samples. Column (2) presents wage equation for workers in paid employment ["+" , "-", and "*" same as for column (1)]. Column (3) presents bias arising from selection of participant and non-participant samples ["+" , "-", and "*" same as for columns (1) and (2)]. Column (4) shows notable results [viz., the significisant variables] for wage equations. Column (5) shows sign and significance for predicted earnings difference between self-employment and employment in structural probit model used to explain self-employment decision, if estimated ["+" , "-", and "*" same as for columns (1), (2), and (3)]. Column (6) shows notable results from structural-probit equation used to explain the self-employment decision. Finally, "ns" refers to "not significant" (sign not reported) and "na" refers to either "not applicable" or "not available."

We would first note that the present paper adds the first Canadian study of self-employment behaviour to this group. Of the eight previous studies, one originates from Dutch data [de Wit and van Winden (1989)], one from West

German data [Pfeiffer and Pohlmeier (1992)], three from British data [Rees and Shah (1986), Dolton and Makepeace (1990), and Taylor (1996)], and three from American data [Macpherson (1988), Evans and Leighton (1989), and Fujii and Hawley (1991)].

Our paper is also unique in its coverage of the self-employment decisions of men and women. A number of studies have explored differences in the characteristics of self-employed men and women, but they have not used an explicit economic framework [e.g., Cachon and Carter (1989), Cromie (1987), Welsch and Young (1984)]. As Table 5 illustrates, all but two of the studies which have used an explicit economic framework are restricted to men [Rees and Shah (1986), de Wit and van Winden (1989), Evans and Leighton (1989), Fujii and Hawley (1991), Pfeiffer and Pohlmeier (1992), and Taylor (1996)]. One study in Table 5 examines only married women [Macpherson (1988)], while one paper analyzes both men and women [Dolton and Makepeace (1990)], but only for recent graduates. Although Dolton and Makepeace include both men and women in their analysis of the self-employment decision, they use a pooled sample. They find gender to be a significant factor in the self-employment decision but, unlike our paper, they do not analyze self-employment separately for men and women. Yet our results suggest that there are substantive differences in the results for men and women that should be taken into consideration in reaching any conclusions about self-employment behaviour.[6]

[6] Using a likelihood ratio test, we soundly reject the hypothesis that the observations for men and women can be pooled in the structural probit model of the self-employment decision [j^2_3=400 for the restricted model; j^2_{19}=1680 for the unrestricted model]. We also reject the hypothesis that the observations for men and women can be pooled in the earnings equation for the wage-employed [j^2_{24}=129], although the hypothesis cannot be rejected for the earnings equation for the self-employed [j^2_{24}=17]. Thus, separate regressions for men and women are dictated by the data and pooling, with or without a gender dummy, is rejected. The detailed results for these pooled equations are available from the authors upon request.

A third contribution of our paper is that it provides a complete analysis of the structural model set out in section 2. Specifically, we estimate models of the earnings of the self-employed and wage-employed which account for the selection biases arising from both participation and self-employment and we use these results to estimate the structural probit model of the self-employment decision. Previous studies have each accomplished some, but not all, of these tasks. Six of the eight studies in Table 5 estimate the structural model of the self-employment decision but adjust the earnings equations only for selection bias arising from self-employment, not participation [Rees and Shah (1986), de Wit and van Winden (1989), Dolton and Makepeace (1990), Fujii and Hawley (1991), Pfeiffer and Pohlmeier (1992), and Taylor (1996)]. One study [Evans and Leighton (1989)] estimates only the earnings equations and finds no selection bias arising from self-employment. Only Macpherson (1988), in his study of married women in the U.K., tests for selection bias arising from participation and, as in our study, finds it to be significant. While one might be more concerned about participation bias for married women (because of the greater frequency of employment disruption within this group), we find that participation bias is significant for both men and women and should not be ignored.

Finally, our study adds to the small literature concerning the effect of labour market conditions on the self-employment decision. Of the six studies which estimate a structural-probit model of the self-employment decision, only two have included local labour market conditions in the form of either the local unemployment rate or the unemployment-vacancy ratio [Pfeiffer and Pohlmeier (1992) and Taylor (1996)]. Both studies are restricted to men and find that the unemployment rate is insignificant. In contrast we find a significant negative relationship between the local unemployment rate and self-employment consistent with the "prosperity pull" hypothesis discussed earlier.

Dolton and Makepeace (1990) include the local unemployment rate in their reduced form model of self-employment, but not in their structural model. For a combined sample of men and women, they find that the unemployment rate exerts a positive and significant effect on the probability of self employment, which provides modest support for the "unemployment push" hypothesis. Our results for women (Table 4) are consistent with Dolton and Makepeace's (1990) findings for women but not for men. Indeed, our results suggest that men and women behave very differently in making self-employment decisions and that they should not be combined into a single group.

4. CONCLUSION

Although self-employment is an important aspect of the modern labour market, much remains to be learned about the motivation for self-employment and its economic impact. New microdata sets, such as the Survey of Labour and Income Dynamics in Canada, provide important and hitherto unavailable information on the self-employed and an opportunity to examine these issues more carefully. In this paper, we set out a conventional economic model to explain who is self-employed which relies, among other factors, on the earnings opportunities for each individual in wage- and self-employment. These earnings opportunities are not directly observable, but they can be estimated in conventional fashion from an underlying model of labour force participation and self-employment decisions and earnings equations.

The model, which has been partially estimated in a few earlier studies, involves three stages. First, we estimate reduced form probit models for participation and self-employment. Secondly, we estimate earnings equations for wage- and self-employment from observed earnings which correct for sample selection bias arising from participation and self-employment. Finally, we estimate

a structural model of the self-employment decision using the estimated earnings difference between wage- and self-employment obtained in the second step.

Our estimates of this full model yield several interesting results. In particular, while other studies have largely ignored women, or combined women and men in a single sample, we find that the self-employment behavior of men and women is quite distinct and merits separate assessment. For example, while both men and women appear to be quite sensitive to the relative earnings opportunities for wage- and self-employment, they respond very differently to local labour market conditions. Lower local unemployment rates are associated with increased self-employment activity for men, reflecting what we have termed the "prosperity pull" hypothesis. Lower unemployment rates result in reduced self-employment activity for women, however, consistent with the "unemployment push" hypothesis. The effects of age and family circumstances on the self-employment decision also differ quite dramatically for men and women. Adjustment for sample selection bias arising from both participation and self-employment is found to be significant for both men and women in our analysis. Whether these results are unique to Canadian circumstances, or are consistent with results for other countries, awaits comparable analyses of the full structural model of self-employment using microdata elsewhere.

REFERENCES

Acs, Z., D. Audretsch, and D. Evans (1994), "Why does the self-employment rate vary across countries and over time?," Center for Economics Policy Research, Discussion Paper No. 871.

Aronson, R.L. (1991), Self-Employment: A Labour Market Perspective, Cornell Studies in Industrial and Labour Relations, No. 24 (Ithaca, NY: ILR Press).

Becker, E. (1983), "Self-employed workers: An update to 1983," Monthly Labour Review 107, 14-18.

Blanchflower, D., and A. Oswald (1991), "Self-employment and Mrs. Thatcher's enterprise clinic," Center for Economic Performance Discussion Paper, No. 30.

Bogenhold, D., and U. Staber (1991), "The decline and rise of self employment," Work, Employ-ment and Society 5, 223-39.

Borjas, G., and S. Rosen (1980), "Income prospects and job mobility of younger men," in R. Ehrenberg, ed., Research on Labour Economics, Volume 3 (Greenwich, Connecticut: JAI Press).

Bregger, J. (1996), "Measuring self-employment in the United States," Monthly Labour Review, January-February, 3-9.

Cachon, J.-C., and S. Carter (1989), “Self-employed females and the workforce: Some common issues across the Atlantic,” Journal of Small Business & Entrepreneurship 6 (4), 20-31.

Catsiapis, G., and Robinson, C. (1982) “Sample Selection Bias with Multiple Selection Rules,” Journal of Econometrics 18, 351-68.

Cohen, G. L. (1988), Enterprising Canadians: The Self-employed in Canada (Ottawa: Statistics Canada).

Creigh, S., C. Roberts, A. Gorman, and P. Swayer (1986), "Self-employment in Britain: Results from the labour force surveys, 1981-84," Employment Gazette 95, 183-194.

Cromie, S. (1987), “Similarities and differences between women and men who choose business proprietorship,” International Small Business Journal 5 (3), 43-60.

Daly, M. (1991), "The 1980s - A decade of growth in enterprise: Self-employment data from the labour force survey," Employment Gazette 99, 109-34.

Dennis, W. (1996), "Self-employment: When nothing else is available?," Journal of Labour Research 17 (4), 645-61.

De Wit, G., and F.A.A.M. van Winden (1989), "An empirical analysis of self-employment in the Netherlands," Small Business Economics 1, 263-272.

Dolton, P., and G. Makepeace (1990), "Self employment among graduates," Bulletin of Economic Research 42, 35-53.

Evans, D., and L. Leighton (1989), "Some empirical aspects of entrepreneurship," American Economic Review 79, 519-535.

Fujii, E., and C. Hawley (1991), "Empirical aspects of self-employment," Economics Letters 36, 323-329.

Hakim, C. (1988), “Self-employment in Britain: Recent trends and current issues,” Work, Employment and Society 2 (4), 421-450.

Heckman, J. (1979), "Sample selection bias as a specification error," Econometric 47, 153-161.

Hubler, O. (1989) “Individual Overtime Functions with Double Correction for Selectivity Bias,” Economics Letters 29, 87-90

Human Resources Development Canada (1998), "Diverging trends in self-employment in Canada," Applied Research Bulletin 4 (1), 13-18.

Killingsworth, M. (1983), Labor Supply (Cambridge, U.K.: Cambridge University Press).

Loutfi, M. (1991), “Self-employment patterns and policy issues in Europe,” International Labour Review 130, 1-19.

Macpherson, D.A. (1988), "Self-employment and married women," Economic Letters 28, 281-84.

Mallett, T. (1991), "A profile of self-employment earnings," Journal of Small Business & Entrepreneurship 8 (3), 15-23.

Meager, N. (1992), "Does unemployment lead to self-employment?," Small Business Economics 4, 87-104.

Pfeiffer, F., and W. Pohlmeier (1992), "Income, uncertainty and the probability of self-employment," Recherches Economiques de Louvain 58 (3-4), 265-281.

Rees, H., and A. Shah (1986), "An empirical analysis of self-employment," Journal of Applied Econometrics 1, 95-108.

Riddell, W.C., and A, Sharpe (1998), "The Canadian-US unemployment rate gap: An introduction and overview," Canadian Public Policy 24 (Supplement), S1-S37.

Taylor, M. (1996), "Earnings, independence or unemployment: Why become self-employed," Oxford Bulletin of Economics and Statistics 58 (2), 253-265.

Tepper, E. (1988), Self-employment in Canada Among Immigrants of Different Ethno-Cultural Backgrounds (Ottawa: Employment and Immigration Canada).

Welsch, H., and E. Young (1984), "Male and female entrepreneurship characteristics and behaviors: A profile of similarities and differences," International Small Business Journal 2 (4), 11-20.

White, H. (1980), "A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity," Econometrica 48, 817-838.

Wolpin, K. (1977), "Education and Screening," American Economic Review 67 (5), 949-958.