INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

A CHANGING LABOUR MARKET:

EMPIRICAL STUDIES OF SELF-EMPLOYMENT

By

HERBERT J. SCHUETZE, B.A.(HON.), M.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

EMPIRICAL STUDIES OF SELF-EMPLOYMENT

DOCTOR OF PHILOSOPHY (1999)
(Economics)

McMaster University
Hamilton, Ontario

TITLE: A Changing Labour Market:
Empirical Studies of Self-Employment

AUTHOR: Herbert J. Schuetze, Hons. B.A. (McMaster University)
M.A. (McMaster University)

SUPERVISORY COMMITTEE: Professors Peter J. Kuhn (Supervisor)
Lonnie Magee
A. Leslie Robb

NUMBER OF PAGES: xi, 143

Abstract

Self-employment has risen dramatically in Canada, accounting for a disproportionate share of job growth since the early 1970's. This trend makes understanding this sector of the labour market imperative to understanding the overall labour market. This dissertation comprises three essays. The goal of the first two essays is to examine possible reasons for the trends in self-employment while the third essay quantifies and characterizes income tax non-compliance among the self-employed.

In the first essay I assess the importance of macroeconomic conditions and the tax environment in explaining the trends in male self-employment in North America. I use microdata for the period 1983-1994 from Canada and the United States, which are perhaps more similar in overall institutional structure than any other two countries, but which differ substantially in their income tax policy, macroeconomic conditions, and self-employment trends. My findings suggest that higher income tax and unemployment rates are associated with an increase in the rate of male self-employment in the two countries. Changes in the tax environment account for a considerable amount of the secular trends in male self-employment over this period, while changing economic conditions played a smaller role in determining these trends.

The second essay, which focuses on Canadian self-employment trends, examines the dynamics of self-employment among men and women using information on labour force transitions from 13 waves of the Survey of Consumer Finances between 1982 and 1995. The results show that the changes in transition patterns underlying the secular rise in self-

employment in Canada were very different for women and men. For women, most of the increase in self-employment is attributable to an increase in their retention rates in self-employment. For men, most is attributable to a decrease in the stability of paid employment.

The final essay provides estimates of the degree of income tax noncompliance by Canadian self-employed households using a series of Canadian Family Expenditure Surveys from 1969 to 1992. These estimates are disaggregated across years, demographic characteristics and occupation to shed some light on the determinants of such activities in Canada and to provide guidance to tax enforcement policy makers. My main findings are as follows. I find that the degree of noncompliance by the self-employed varies significantly by occupation, age and the number of household members self-employed. Further, there is no clear time trend in the estimates of noncompliance and education does not appear to influence these estimates.

Acknowledgements

I thank those whose guidance, patience and fostering aided in the completion of my thesis. Professor Peter Kuhn, my thesis supervisor, provided direction and friendship, all the while allowing me to find my own way. The members of my thesis committee, professors Lonnie Magee and A. Leslie Robb afforded open doors and thoughtful advice on any topic. I also thank the Canadian International Labour Network (CILN) for financial support.

Of course, this would not have been possible without the love and encouragement of my family and friends. I thank my wife, Amy, who endured the years of study and still dispensed unconditional support and love. My parents, Herbert Sr. and Brigitte, and family provided tremendous encouragement and have always been my greatest backers. I am blessed with a second set of parents, Vern and Sandie Gross, whose generosity and adoration were paramount, and an extended family that accepted me as their own.

Thank You All!

Preface

The third chapter, "Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982-1995," was jointly authored with Professor Peter Kuhn for the purpose of publication. My contribution to this work included the empirical work and much of the write-up. Professor Kuhn is responsible for the idea to use this particular data set, which had previously not been used as panel data, and edited the final version of the paper.

A Changing Labour Market: Empirical Studies of Self-Employment

Table of Contents

List of Tables

Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-U.S. Comparison

Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982-1995

List of Figures

Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-U.S. Comparison

Chapter 1

INTRODUCTION

Relative to the wage and salary sector of the labour market in Canada, self-employment has become increasingly important. In Canada between 1979 and 1997 the number of self-employed individuals almost doubled to 2.5 million.[1] Of greater significance, however, is the growth in the self-employment sector of the labour market relative to the wage and salary sector. Picot et. al (1998) show that, over the same period, the self-employment rate (the fraction of employed workers who are self-employed) rose from about 13 percent to 18 percent. These authors also suggest that this relative growth was particularly substantial in the 1990's– accounting for nearly 80 percent of the net increase in employment between 1989 and 1997. Given the increasing importance of the self-employment sector of the labour market over the last two decades, it is imperative to understand these trends and their consequences if we wish to understand the structure of the current labour market.

This dissertation comprises three essays. The goal of the first two essays is to

[1]Picot ,G., Marilyn Manser and Zhengxi Lin, (1998) "The Role of Self-Employment in Job Creation in Canada and the United States", Working Paper US Bureau of Labor Statistics and Statistics Canada.

examine possible reasons for the trends in self-employment while the third essay quantifies and characterizes income tax non-compliance among the self-employed, a possible consequence of such employment.

To date, analysts have primarily focused on changes in technology, industrial restructuring and in the demographic composition of the work force as explanations for the rise in self-employment, with limited success. At the same time, international statistics suggest that country- or region- specific factors, rather than widely-shared trends may play central roles in the evolution of self-employment rates. In the first essay I assess the importance of two less commonly analyzed factors which vary across regions and countries– macroeconomic conditions and the tax environment– in explaining the trends in male self-employment in North America. Using microdata for the period 1983-1994, I study Canada and the United States, which are perhaps more similar in overall institutional structure than any other two countries, but which differ substantially in their income tax policy, macroeconomic conditions, and self-employment trends. Unlike Canada, US males did not experience a dramatic increase in their rates of self-employment.

Because of their focus on data from a single country, and in some cases their lack of regional disaggregation, previous authors who have studied these factors as possible determinants of self-employment have been limited in their ability to isolate their importance: essentially they have been restricted to using a single time series of tax and macroeconomic variables. To overcome this limitation my analysis operates on a pooled sample of all employed individuals in two countries over 12 years. In addition I disaggregate

each country's labor markets and tax jurisdictions to the state/province level, allowing me to use asynchronous variation in these conditions across these 60 provinces and states as a source of identification.

My findings suggest that increases in average income tax rates have large and positive effects on the rate of male self-employment. This effect is larger for average tax rates at higher points in the income distribution, as well as for individuals with higher education levels, and those engaged in industrial activities often involving cash transactions-- who should be more affected by these tax rates. This suggests that under-reporting of self-employment income may be a key motivating factor for becoming self-employed. Increases in the provincial/state unemployment rates are also associated with a rise in the rate of male self-employment. While one might not expect these "macroeconomic" effects to play a large role in explaining secular changes in self-employment rates, I find that, because of the widening Canada-US unemployment rate gap over the period examined here, they do play some role in explaining the widening Canada-US gap in self-employment. This role is however much smaller than that attributed by my model to the tax policy variables.

The second essay examines the dynamics of self-employment among Canadian men and women between 1982 and 1995 to shed some light on the underlying causes for the secular rise in self-employment. Because of a lack of good longitudinal data on labour market activity, Canadian researchers, in particular, have been forced to focus on the stock of self-employed individuals. However, analyses which examine stocks can mask information about the underlying process by which individuals become, or cease to be,

self-employed. This lack of Canadian longitudinal data is overcome using hitherto-unexploited information on labour force transitions from 13 waves of the Survey of Consumer Finances between 1982 and 1995. These data files contain standard labour force data for the week prior to the survey as well as supplemental data on the previous year's work experience and income. As a result, individuals are observed in contiguous years, providing a 2-year panel for each individual.

Changes in the transition patterns underlying the secular rise in self-employment experienced over this period are analyzed in the context of a Markov process among three labour force states: employment in the wage-and-salary sector, self-employment, and non-employment, for men and women, separately. These transition patterns suggest that the reasons for the rise in self-employment were very different for women and men. For women, most of the increase in self-employment is attributable to an increase in their retention rates in self-employment. For men, most is attributable to a decrease in the stability of paid employment, i.e. a rise in transitions from employment to non-employment. This generates an increase in self-employment because non-employed men are much more likely than employed men to enter self-employment. Somewhat paradoxically, self-employment thus rose both in secularly improving (women's) and deteriorating (men's) labour markets, due to different changes in the underlying transition processes.

One likely consequence of the trends in self-employment in Canada arises from the fact that the self-employed are often able to conceal income from the tax authorities because of the absence of a third party reporting income. Public finance economists argue that the

presence of such "underground" activity is economically distortionary and leads to an inequitable distribution of the tax burden (Auerbach and Feldstein 1998 and Slemrod 1992). These economic inefficiencies are likely to increase in importance in Canada in light of the recent trends in self-employment. However, very little is known about the extent to which the self-employed conceal income in Canada and even less is known about the factors which influence the degree of income concealing among the self-employed.

In the third essay, I utilize an expenditure based approach developed by Pissarides and Weber (1989) to estimate the degree of noncompliance by the self-employed using a series of Canadian Family Expenditure Surveys. I obtain estimates of the degree of noncompliance by the self-employed in Canada across years, demographic characteristics and occupations. By disaggregating across these dimensions it is intended that this essay will achieve two goals. The first goal is to identify the determinants of self-employed income tax noncompliance in Canada and the second is to provide guidance to tax enforcement policy makers in designing effective tax policy schemes.

My main findings in this third essay are as follows. First, with regard to the determinants of noncompliance among the self-employed, there appears to be no clear growth in the degree of income concealing by the self-employed over this period. I do, however, find evidence that the degree of noncompliance varies significantly by occupation of the entrepreneur. This suggests that one's "opportunity" to under-report income is likely to be a key determinant of the degree of noncompliance among entrepreneurs. Finally, one's level of education does not appear to be a determining factor in the decision to conceal

income among the self-employed.

Second, with regard to the results looking across characteristics observed on individual income tax forms, there appears to be a great deal of variation in the degree of noncompliance among the self-employed across these dimensions. Households with significant amounts of self-employment income which are headed by younger males tend to conceal income to a greater extent than do those headed by older males. Further, noncompliance among self-employed households is highest among those in which the occupation of the self-employed individual(s) is construction, followed closely by those in all service occupations. Households with members engaged in self-employment activities in the product fabrication industries had the lowest estimates of noncompliance. Finally, there appears to be no difference in the degree of noncompliance among households in which only the husband is self-employed as compared to those in which only the wife is self-employed. However, estimates of the extent of income noncompliance are lower in households in which both members earn self-employment income. These results have important implications for policy makers involved in developing taxation enforcement policies. For instance, a tax audit scheme targeting groups which have been found to conceal income the most (for example, self-employed households headed by younger males or those in the construction and service occupations) is likely to be an effective tool in reducing noncompliance.

Chapter 2

TAXES, ECONOMIC CONDITIONS AND RECENT TRENDS IN MALE SELF-EMPLOYMENT: A CANADA-U.S. COMPARISON

(1) Introduction

The resurgence of self-employment has recently attracted the attention of researchers in a number of countries (e.g. Blau 1987; Evans and Leighton 1989; Devine 1993). Attempts to explain this phenomenon have however met with only limited success, for a number of reasons. First, standard shift-share analyses tend to show that the factors most commonly invoked to explain this trend --industrial restructuring, and shifts in the demographic composition of the workforce-- typically can account for only a small fraction of the observed changes. Second, another commonly-invoked explanation --changes in technology-- remains very difficult to test, and in practice is often simply treated as a label attached to otherwise unattributable changes.

Finally, although this is not typically noted in the literature, there is a third reason to be skeptical of structural and technological explanations of rising self-employment: recent trends in self-employment rates are far from uniform across developed countries (OECD 1992). Indeed, with declines almost as common as increases across OECD countries, international statistics strongly suggest that country- or region- specific factors, rather than

widely-shared trends like cheaper computing power, feminization of the labor force, or the move to a service economy, may play central roles in the evolution of self-employment rates.

The goal of this essay is to examine the role of two less commonly analyzed factors which do vary across regions and countries --macroeconomic conditions and the income tax environment-- in explaining recent self-employment trends. Macroeconomic conditions have often been cited as a potential contributor to self-employment, especially to the extent that self-employment is used by some individuals as a "job" of last resort in poor labor markets (Quinn 1980, Becker 1984, and Bishop 1987). Tax policy, and especially the rate of personal income tax, have also been cited (e.g. Long 1982, Blau 1987 and Devine and Mlakar 1993) as possible determinants of self-employment rates, largely because self-employment offers individuals greater opportunity to shelter, or hide, income from tax authorities, an option which is especially valuable in high-tax jurisdictions. Clearly, if macroeconomic conditions and tax policy, rather than fundamental technological change, are driving recent increases in self-employment in some countries, policy prescriptions may differ.

Because of their focus on data from a single country and, in some cases their lack of regional disaggregation, previous authors have been limited in their ability to isolate the importance of tax and macroeconomic factors: essentially they are restricted to using a single time series of tax and macroeconomic variables. In the current essay I address this problem in two ways. First, I use 10 years of microdata covering the 12 year period 1983-1994[1] from

[1]Data pertaining to the years 1984 and 1985 were not available. There was no 1984 Canadian Survey of Consumer Finances and information on incorporated self-employed was not available for 1985 in Canada.

two countries, Canada and the US, which are perhaps more similar in overall institutional structure than any other two countries, but which differ substantially in their income tax policy, macroeconomic conditions, and self-employment trends. Second, I disaggregate each country's labor markets and tax jurisdictions to the state/province level, allowing me to use asynchronous variation in these conditions across these 60 provinces and states as a source of identification. In effect, my analysis operates on a pooled sample of all employed individuals in two countries over 12 years, assessing the importance of the total tax and macroeconomic environment of their province or state as a determinant of self-employment trends. My focus in this essay is on self-employment among males aged 25-64; unlike women, men were not affected by a massive secular rise in wages, experience and overall labor force participation rates which could obscure the effects of the tax and macroeconomic factors that are the subject of this essay. Men in this age group are also less affected by secular increases in school attendance, a trend which plays an important role for other age groups.

My main findings are as follows. First, perhaps surprisingly, non-primary self-employment rates (hereafter simply referred to as self-employment rates) for males aged 25-64 were higher in Canada than the US during the period 1983-1994, which is the focus of my analysis.[2] Second, while much has been made of an earlier increase in US male self-

[2]The rate of non-primary self-employment is the fraction of individuals employed in non-primary industries who are self-employed in their main job in either incorporated or unincorporated businesses. "Primary" industries consist of agriculture, forestry, fisheries and mining. Although the self-employment rates are calculated based on the individual's main job in the survey week in the Canadian survey and to the individual's longest job in the previous year in the US survey, these definitional differences do not appear to

employment, and of the continuing increase in female self-employment, the self-employment rate of prime-age males actually fell over the period covered by my data, from 13.9 percent in 1983 to 12.4 percent in 1994. Third, in contrast, Canada experienced an increase in male self-employment over this period, from 16.3 to 18.5 percent of the employed labor force. While certainly not conclusive, these trends and levels are strongly suggestive of both macroeconomic and tax explanations, because Canada had increasing income tax rates and a deteriorating macroeconomic environment relative to the US over this period.

Fourth, in a regression context that controls for province/state effects and industry specific time trends, I find that increases in average income tax rates have large and positive effects on the rate of male self-employment, as expected. This effect is larger using average tax rates at higher points in the income distribution, as well as for individuals with higher education levels and among those in industries characterized by "informal suppliers", who should be more affected by these tax rates. Overall, the estimated impact of a 30 percent increase in taxes is a corresponding rise in the rate of male self-employment of between 0.9 to 2 percentage points in Canada and between 0.8 and 1.4 percentage points in the US, over 1994 levels. This implies that under-reporting of self-employment income is one of the motivating factors for becoming self-employed. In fact, a decomposition demonstrates that changes in the average tax rates are the largest contributing factor, of the determinants examined, for the secular trends in male self-employment in North America.

---

account for the differences observed in the rates. See section 4 for a discussion.

Finally, increases in the provincial/state unemployment rates appear to also be associated with a rise in the rate of male self-employment. However, estimates of the elasticity of self-employment with respect to the unemployment rate are considerably smaller than those associated with the tax rates. For instance, a 30 percent rise in the rate of unemployment in Canada (3 percentage points) would result in a 0.6 percentage point increase in the rate of self-employment using 1994 figures. While one might not expect these "macroeconomic" effects to play a large role in explaining *secular* changes in self-employment rates, I find that, because of the widening Canada-US unemployment rate gap over the period examined here, they do play some role in explaining the widening Canada-US gap in self-employment. However, this role is much smaller than that attributed by my model to the tax policy variables.

The remainder of this essay is organized as follows: Section (2) summarizes recent international trends in self-employment rates, and Section (3) reviews previous explanations of trends in self-employment. The data are described in Section (4). Section (5) describes the characteristics of self-employed males and the aggregate trends in male self-employment rates, tax liabilities and unemployment rates. In Section (6) I describe and analyze the results of various regression specifications. Section (7) presents the results of a simple decomposition of the predicted male self-employment rates between 1983 and 1992 and Section 8 concludes.

(2) Recent Self-Employment Trends Across Countries

To place the analysis of the current essay into a broader context, Table 1 gives the change in the rate of self-employment[3] between 1973 and 1990 and the 1990 level for a number of developed countries. The nonagricultural self-employed made up over 10 percent of the employed population in half of the countries in Table 1. The highest rates of self-employment are found in Italy, Portugal, and Spain. In contrast, relatively few of the employed in Austria and Norway were self-employed. Rates of self-employment in these countries were just above the 6 percent mark.

There is also a great deal of variation in the aggregate self-employment/labor market trends in these countries. For instance, unlike Canada and the United States, eight of the sixteen countries included in Table 1 experienced a decrease in the rate of self-employment between 1973 and 1990. Of the eight countries that had a decrease in the rate of self-employment, Austria and Luxembourg witnessed the greatest declines. Self-employment rates fell in Austria and Luxembourg by 5.3 and 4.0 percentage points, respectively. Between the same years, the rate of self-employment increased by 5.8 and 4.3 percentage points in Portugal and the United Kingdom, respectively. The fact that these developed countries' self-employment experiences were so different suggests that no over-riding factor, like technological change or industrial restructuring, is responsible for the trends in self-

---

[3]The data sample and self-employment rate definition in Table 1 differ from those in the main body of the essay. Here, the sample includes both men and women and excludes the incorporated self-employed and the rate of self-employment is defined over the total population rather than the number employed.

employment, and directs my attention toward institutional and other factors that vary across these countries.

(3) Previous Studies of Self-Employment Trends

One common explanation examined in the literature and expressed in the popular press suggests that technological advances have increased opportunities for self-employment. The dissemination of personal computers is often cited as an example of a technological change that facilitates self-employment by decreasing capital costs, thereby reducing barriers to entry. Devine and Mlakar (1993) used the price of computing power to estimate the effects of technological advances on the probability of becoming self-employed both across and within industries using a series of Current Population Surveys from 1975 to 1990. They found that, across industries, the price of computing power had little or no effect on the probability of self-employment and that it had a significant positive effect within only one of the industries. Their analysis is however severely limited by use of a single computing power price series in all industries.

Another common explanation examined by researchers, not unrelated to technological change, suggests that shifts in the composition of industries' employment shares towards industries where self-employment is more prevalent can account for the rise in self-employment in North America. The most common example given is the recent shift towards service producing industries, in which self-employment has always been more prevalent, in North America. Blau (1987) and Devine (1993) tested this hypothesis using US

data on males only in the former case and on both males and females in the latter case. Both papers found interindustry shifts in employment to be a significant factor in explaining the increase in the US self-employment rate. However, as Devine points out, these shifts are not the largest contributor. Devine found that within-industry increases in self-employment produced most of the movement in the self-employment rate.

Perhaps the most frequent explanation for the rise in self-employment given by researchers concerns shifts in the demographic composition of the workforce. The hypothesis generally put forth is that the increase in self-employment is a result of increased representation among the employed of demographic groups that have always been more likely to be self-employed. The demographics examined include age, education, marital status and gender. Typically, researchers have found that in a single cross section older, more highly educated, married male workers are more likely to be self-employed.[4] However, changes in the demographic composition of the employed accounted for only a small fraction of the overall increase in self-employment. Moreover, the researchers found that most of the increase in self-employment occurred within demographic groups.

Most analysts pay little attention to institutional factors in their accounts of the recent trends in self-employment. Besides income tax policy, institutional factors that have been examined include minimum wage legislation (Blau 1987), immigration policy (Borjas and Bronars 1989), and retirement policies (Quinn 1980, Parnes and Less 1985, and Iams 1987).

---

[4]See, for example, Devine (1993), Evans and Leighton (1989) and Crompton (1993).

Studies of the effects of income tax policies on self-employment include Long (1982), who identified a statistically significant positive relation between imputed federal income tax liabilities and male self-employment using a single cross section of data from the 1970 US census. Time series studies include Blau (1987), who found that only the higher of the two assumed marginal tax rates included in his study had a positive effect on the male rate of self-employment in the US between 1948 and 1982. One serious concern with this analysis is due to the lack of variation in the tax rates. Variation in the tax variables is limited to that which is captured in a time series for a single country which amounts to 35 observations--one for each year from 1948-1982. A more recent study by Devine and Mlakar (1993) reported that while aggregate personal income tax rates had a positive and significant effect across industries between 1975 and 1990, within industries the same aggregate tax measure was significant only in the trade industry. Like the previous study, this study also lacks cross-section and cross-country variation in the tax measures.

A number of studies have also examined the cyclical aspects of self-employment. Again, the results are mixed. For instance, Becker (1984) observed that, in raw data, the rate of self-employment moved countercyclically in the US between 1948 and 1982. Using panel data on the US between 1968 and 1987, Evans and Leighton (1989) found that white men who are unemployed are almost twice as likely as employed wage workers to enter self-employment. However, they do not use this rather strong finding to explain secular trends in US self-employment. In contrast, Blanchflower and Oswald (1990) find the rate of self-employment to be procyclical using recent data on the United Kingdom. While this is one

of only two papers, to my knowledge, to use regional variation in macroeconomic conditions, there were relatively few regions from a single country and the sample size in any year and region was small. A second study which also utilizes regional variation in macroeconomic conditions using Canadian data (Lin et. al 1998) finds an empirically small procyclical relationship between aggregate self-employment rates and macroeconomic conditions. One limitation of this paper is the fact that regional variation in unemployment rates comes solely from provinces whose tax rates are linked to a great extent[5]. The focus in Lin et. al is on determining if an aggregate cyclical relationship exists between self-employment and demand conditions unlike the current study which focuses on microdata and the effects of demand conditions on secular changes in self-employment.

(4) Data

As noted above, this essay employs a series of microdata files from Canada and the US for the years 1983 to 1994[6]. The microdata files are taken from the Canadian Surveys of Consumer Finances[7] (SCF) and the US Current Population Surveys (CPS) which are

---

[5] With the exception of Quebec, provincial tax rates are calculated as a percentage of the basic federal tax liability.

[6]Because, unlike the CPS, the SCF data employ questions from the Labor Force Survey which refer to the "reference week" (see below) the data files used for Canada involve 1983 to 1994 labor force behaviour but the 1982 to 1993 income years. As explained earlier, 1984 and 1985 are not available.

[7]Additional information to that provided in the public use SCF data files was provided by Statistics Canada. The added information allows a distinction to be made between wage and salary earners and the incorporated self-employed, who are typically grouped together. Because of a trend toward incorporation among the self-employed the true trends are typically masked.

conducted in April and March of each year and contain information on income as well as personal and labor-related characteristics for approximately 75 and 160 thousand individuals per year, respectively. Both the April SCF and the March CPS provide standard monthly labor force data as well as supplemental data on the previous year's work experience and income. The data extracted from these data files were, in some cases, recoded to make variables as consistent as possible both across years within surveys and across the surveys.[8] All samples are restricted to males aged 25 to 64 who were employed[9] in non-primary industries. The individual year/country data files are pooled to create a single data file containing 487,062 observations or approximately 49,000 observations per year of which about 32,000 observations are from the US CPS and 17,000 are from the Canadian SCF.

The variables used to identify whether or not an individual was self-employed in the two surveys were quite similar. In both surveys, respondents were asked to report whether, in their "main" job, they were i) a paid worker in the private sector, ii) a paid worker in the public sector, iii) self-employed in an incorporated business, iv) self-employed in an unincorporated business or v) an unpaid family worker (the exact questions are included in section A of the appendix). However, information distinguishing between the incorporated and unincorporated self-employed is only available for the survey week in the SCF and for

---

[8]A detailed report of the extraction is available from the author.

[9]The data is restricted to those who were employed in the reference week in Canada and at any time in the reference year in the US.

the calendar year preceding the survey week in the CPS.[10] Therefore, in the SCF, worker classification refers to the individual's main job in the week prior to the survey while in the CPS worker classification refers to the individual's longest job in the previous year. Differences between the self-employment rates calculated using the two methods might be expected to be either positive or negative.[11] To get an idea of the quantitative importance of this issue, I compared the rates of self-employment among *unincorporated* businesses based on the survey week and the previous year for a number of years from the US CPS. While the rates of self-employment based on the survey week tend to be higher than those based on the survey year, these differences are very small (less than half of one percentage point).

Because individual income tax liabilities or tax rates are endogenous in a self-employment equation, I use an alternative measure of taxes as a proxy for the "tax environment".[12] The criteria for choosing such a measure are that it should (1) capture changes in the tax code over time rather than fluctuations in income, but also, (2) be rich enough to encompass variation in the tax code across the income distribution. To this end,

---

[10]While data on the unincorporated self-employed is available in both surveys for the survey week, the decision was made to include both the incorporated and unincorporated self-employed because an increasing fraction of the self-employed over this period owned incorporated businesses.

[11] The definition based on the survey week is at a point in time whereas the one based on the survey year requires that the individual be mostly self-employed in the previous year. Since those who are employed throughout the year but self-employed for a short term are more likely to be included in the definition based on the survey week the rate might be higher using that measure. On the other hand, we are more likely to observe individuals who had short spells of self-employment but were not employed for the remainder of the year in the survey year.

[12]An alternative to "proxying" for the tax environment would be estimate using two-stage least squares including the proxy variable as an instrument for actual tax outcomes. However, data on actual tax outcomes is not available in the CPS, though it is in the SCF.

the tax data are calculated by evaluating the income tax liability of a family[13] with constant real income over the period at different points in the income distribution for each year and province/state. This is done by first calculating real family income[14] at the 50th and 90th percentiles[15] in each year and taking the average over the period for the two points in the income distribution. For Canada this procedure yielded figures of $49,965 and $84,365 real 1992 Canadian dollars for the 50th and 90th percentiles, respectively. Similar figures for the US were $51,222 and $96,105 in real 1992 Canadian dollars. These real income figures, one per country for each of the two points in the income distribution, are used to calculate tax liabilities within countries throughout the period. This ensures that the tax measure captures variation in tax liabilities that is due to changes in the tax code and not to tax increases that are due to fluctuations in average incomes. Separate calculations of the 50th and 90th percentiles of family income were computed for each country because tax codes are linked to a great extent to a country's income level and income distribution. Using the nominal equivalent of the real incomes for the two points in the income distribution, tax liabilities net of deductions for each year and province/state are then calculated using information taken

---

[13]The term "family" refers to a couple (male and female) with no children. Children were not considered here for the sake of simplicity.

[14]Family income is calculated by ranking family income (as defined in the survey) for males and taking the male's income associated with family income at the 50th and 90th percentiles. A similar procedure yields female incomes associated with family incomes at the 50th and 90th percentiles. The male income and female incomes are added together to create a "family" (couple) for the 50th and 90th percentiles.

[15]Initially, the 10th percentile was also used. However, because of data limitations that prevent me from calculating tax credits and deductions available to low-income families in Canada the 10th percentile was dropped.

from a series of publications from the Canadian Tax Foundation, for Canada, and using the computer program TAXSIM[16] from the National Bureau of Economic Research for the US (for details on the deductions etc. included in the tax calculations see section B of the appendix). For estimation purposes the combined real federal and provincial/state tax liabilities for each province/state and year are used. US figures are converted to real 1992 dollars using the US consumer price index and then to real Canadian 1992 dollars using the purchasing power parity figure[17] for that year. All tax liabilities are converted to an average[18] tax rate by dividing by the relevant real income used in calculating the liabilities.

Average unemployment rates are calculated for males and females together by province or state using the SCF and CPS microdata files. Aggregate average annual unemployment rates are used as regressors since they are less likely to be endogenous than male unemployment rates alone.

(5) Sample Characteristics and Trends

Sample Characteristics

Table 2 presents raw self-employment rates and the employment shares by

---

[16]For Canada see the Canadian Tax Foundation publications (1982-1993) and for information Regarding Taxsim see: Feenberg et al. (1993).

[17]Purchasing power parity figures are taken from the PENN World Tables (version 5.6).

[18]The average rate is used rather than the marginal rate in part because I am considering discrete changes rather than marginal changes in labor market activities (i.e. whether or not an individual is self-employed in his main job). The individual takes into account the "overall" tax burden. In practice, average and marginal tax rates are likely to be highly correlated across time and space.

demographic group and industry category in Canada and the US, separately, for 1983 and 1992 (two years at similar points in the business cycle[19]). The employment shares are the percentages of the sample (the employed) in each group or cell and sum to one hundred over the demographic and industry groupings in any year and country.

Two opposing shifts in the age structure of employed males in Canada and the US occurred between 1983 and 1992. The first, which is doubtlessly the result of the "baby boom", is an increase in the age of the male working population. Employment shares among the lowest age category fell while these same shares rose or remained constant for those age 36-55. Since male self-employment rates tend to increase with age this shift favours increased self-employment. However, a second shift-- likely caused by a trend towards early retirement in North America-- acted at the same time to decrease employment shares among the oldest workers in the sample (those aged 55-64). The expected effect of the two shifts on self-employment, therefore, is indeterminate.

One interesting characteristic of the self-employment rates across education categories found in Canada but not in the US is the "U-shaped" pattern of self-employment rates. In Canada, the raw self-employment rates are highest in the lowest education category (those with 8 or less years of education).[20] Male self-employment rates fall with education beyond 8 years in Canada but increase slightly for those with any post secondary education

---

[19]Both 1983 and 1992 were relative troughs in the business cycle (i.e. years in which the unemployment rate reached a local maximum) in Canada and the US

[20]Interestingly, in almost all industry and occupation categories the rate of self-employment among low-educated workers in Canada is higher than in the US.

relative to those with 11-13 years of education. In the US, unlike in Canada, male self-employment rates tend to increase monotonically with the level of education. Between 1983 and 1992 the level of education among employed males in both Canada and the US increased. This shift towards more highly educated workers clearly favours increased self-employment in the US but because of the non-monotonic relationship between self-employment rates and education in Canada it is not clear whether or not such a shift has the same effect in that country.

Shifts in the family structure of employed males between the two years, for the most part, worked against increased self-employment. Fewer employed males were married-- the group with the highest self-employment rates-- and more workers were single and divorced, widowed or separated in 1992 relative to 1983. The share of male workers in Canada and the US with no older children (aged 7-17) increased between 1983 and 1992. This shift should act to decrease self-employment rates because, unlike for the presence of young children, male self-employment rates tend to rise with the number of older children present in the family in both countries.

Employment shares among North American males shifted towards industries in which male self-employment rates were high. Rates of self-employment in Canada and the US were highest in the construction, services and retail trade industries and lowest among the manufacturing industries. Between 1983 and 1992 employment shares fell substantially in the manufacturing industries in both countries and rose in the service industries and, to a certain extent, in the retail trade and construction industries as well.

It appears from the raw data, therefore, that in both countries, some shifts in the demographic and industrial structure of the employed favored self-employment, while others worked against it. This casts some doubt on the potential of explanations for the trends in self-employment based on such shifts. Also incompatible with such an explanation is the fact that, in Canada, where aggregate self-employment rose between 1983 and 1992, the rate of self-employment rose within all demographic groups except for single males and within all industries except manufacturing durables and retail trade.

Self-Employment Rate Trends

Figure 1 shows the trends in the non-primary self-employment rates for males aged 25 to 64 for Canada and the US over the period. The rate of male self-employment in Canada lies everywhere above the US rate throughout the period. Between 1983 and 1989, the rates of male self-employment in both countries were relatively stable. In Canada, this rate fluctuates by less than one percentage point and closes out the period at about the same rate (16.3 percent) that it was in 1983. Similarly, in the US this rate was constant, hovering around the 13.5 percent mark - starting out slightly above this mark at 13.9 percent in 1983 and ending up slightly below at 13.1 percent in 1989.

Following 1989, however, the Canadian and US male self-employment rates diverge. In Canada it rises from 16.2 percent in 1989 to 18.7 percent in 1993 and then falls slightly in 1994. In the US, on the other hand, this rate remains stable up until 1992 and then declines by about one percentage point from 1992 to 1994. The gap between the rates in

Canada and the US, which averaged about 3 percentage points between 1983 and 1989, doubled by 1994 to just over 6 percentage points.

Figure 2 disaggregates the country-specific annual self-employment rates by incorporation status. In both Canada and the United States the number of unincorporated self-employed outnumbers the incorporated self-employed over the entire period. The fractions of workers employed in unincorporated businesses in Canada and the US were about the same until 1990, when the Canadian rate rose relative to the US rate. Prior to 1990, the unincorporated self-employment rate in both countries was about 9 percent. After 1990, this rate rose to almost 11 percent in Canada and fell to 8 percent in the US by 1994. One striking difference between the two countries is the large gap between their rates of self-employment in incorporated businesses,[21] which is roughly constant over the period.

Average Tax Rates

The trends in tax rates, like those in the self-employment rates, diverged significantly between Canada and the US over the period.[22] Figures 3 and 4 show the average annual tax

---

[21]The rates of self-employment in incorporated businesses are higher in Canada than the US in all industries. The rates in Canada are about twice those in the US in all industries over the entire period except for construction where the rate is about four times higher in Canada. One would expect to find rates higher in a single industry if these differences were definitional (e.g. if lawyers were considered self-employed in Canada but not in the US).

[22]Throughout the analysis I examine the effects of lagged taxes on current self-employment rates for reasons which will become obvious. For consistency and because tax data for the US in 1994 are not currently available the trends from 1982-1993 are examined here.

rates for the two countries separately[23] for a family earning the median and 90th percentile of income, respectively. In Canada, at both income levels, tax rates rose throughout the period with the exception of a sharp decline between 1987 and 1988. This was a result of tax reform which resulted in a reduction in tax rates and a change in the rate schedule from 10 tax brackets to only 3 brackets. The greatest rise in average tax rates occurred in Canada between 1983 and 1987 among families earning median income. Over the entire period, average tax rates in Canada rose by 1.5 percentage points for median income families and 2 percentage points for families at the 90th percentile. Between 1982 and 1987 tax rates in the US fell by 4.5 percentage points for median family income earners. The downward trend in tax liabilities continued an extra year for US families at the 90th percentile which resulted in a decline of almost 7 percentage points between 1982 and 1988. Tax rates were stable for median US family income earners following 1987 and for those at the 90th percentile following 1988.

Figures 5 and 6 present the average tax rates for a family earning median income for six provinces and six states, respectively, over the period examined. The provinces or states were chosen to illustrate the variation in the regional tax rates. In all cases, the provinces or states with the highest and lowest average tax rates over the period are included in the figures. Average tax rates for a family at the median in Canada (Figure 5) were highest in Quebec and lowest in Alberta. The provincial trends, which for the most part followed a

[23]Average annual tax rates at the country level are calculated as a weighted average of the provincial or state level rates using the sample weights.

similar time path, contained some cross-province variation. For instance, between 1982 and 1987 average tax rates rose by 1.7 percentage points in Quebec and by almost twice that figure in Alberta: rising by 3.3 percentage points. Further, the rather sharp decline in tax rates, experienced to varying degrees in all the provinces between 1987 and 1988, was more than twice as large in Quebec as in Manitoba. Average tax rates fell by 1.8 and 0.7 percentage points in Quebec and Manitoba, respectively.

In the US, the dispersion of average tax rates for a family at the median in any given year was much larger than in Canada. The difference between the highest (New York) and lowest (Florida) tax rate by state in the US (Figure 6) was on average 10 percentage points over the period. The state trends, like the provincial trends, followed similar time paths, however, there was a great deal more asynchronicity in the variation of tax rates in the US than in Canada. For example, average tax rates in Arkansas fell by almost 5 percentage points between 1990 and 1991 and rose by 4 percentage points in North Dakota between 1991 and 1992. These tax rate variations came at a time when the other state tax rates were stable.

Unemployment Rates

Figure 7 shows the aggregate unemployment rates for individuals aged 15 and over in Canada and 16 and over in the US between 1983 and 1994. A gap which opened up in the early 1980's between the Canada and US unemployment rates persists throughout the period. The Canadian unemployment rate increased from about 2 percentage points above the US

rate in the 1980's to 3.5 percentage points above it in the 1990's. In both countries the unemployment rate falls between 1983 and 1989, rises between 1989 and 1992 and falls again after 1992.

Like the trends in average tax rates by province or state, unemployment rates by province/state follow somewhat similar time trends in Canada and the US. However, as Figures 8 and 9 show there is greater regional variation in the time trends for unemployment rates than for average tax rates. In Canada and particularly in the US there are numerous examples where one province or state is experiencing a decline (rise) in its unemployment rate while the other provinces/states are experiencing a rise (decline) in unemployment rates. Further, provincial/ state unemployment rates, even when rising or falling together, do so to varying degrees.

(6) Estimation and Results

An empirical investigation into the role that average tax rates and economic conditions play in the decision of males to become self-employed proceeds by estimating a linear probability model by OLS[24] using the pooled cross-section time-series data on Canada and the US combined with the province/state level tax and unemployment data described above. The data file contains more than 487,000 observations on ten years between 1983-1994 for an average of 49,000 observations a year.

---

[24]OLS was used rather than a probit because of the difficulties that arise performing decompositions with a non-linear model. In any case, work not reported here showed that the coefficients/results are similar to the OLS results when a probit is used to estimate the "base" model.

Separate regressions are estimated using the average tax rates for a family with the median and 90th percentile incomes for the following model:[25]

$$SE_i=\alpha+\beta_1*ltax(PP)_i+\beta_2*urate_i+\beta_3*ind_i+\beta_4*dem_i$$
$$+\beta_5*dem_i^C+\beta_6*reg_i+\beta_7*indt_i+\beta_8*indtsq_i+\epsilon_i \quad (1)$$

α, β_1 and β_2 are scalars while the remainder of the coefficients are vectors. Subscript "i" indexes each of the individual observations (1 to 487,062) and the tax and unemployment rate apply to the year-province/state in which individual i lives. **SE** is a 0-1 indicator variable for self-employment, equal to 1 if the individual is self-employed. **ltax(PP)** is the lagged average tax rate where PP is the two digit percentile (50 or 90)[26] and **urate** is the current year's unemployment rate.[27] **ind** is a series of industry dummy variables, **dem** and $\mathbf{dem}^C$ are a number of demographic variables and those same demographic variables crossed with a country dummy (equal to 1 if the country is Canada and 0 if the US), respectively, and **reg** is a set of dummy variables for individual province or state. **Indt** and **indtsq** are the

---

[25]For a complete description of the variables see appendix, section C.

[26]Ideally one would like to include both personal and business taxes in equation (1) because both tax codes influence the self-employment decision. However, personal tax codes are essentially the same as those for self-employed individuals in unincorporated businesses with the exception that more tax deductions are available to the self-employed. Also, under certain circumstances the tax structure for the incorporated self-employed is the same as that for personal income tax. Therefore, I feel that the costs of including such data outweigh the benefits.

[27]The current tax rate and the lagged unemployment rate were statistically insignificant when the lagged tax rate and the current unemployment rate are included in the regression (see Table A1). This result seems reasonable given that individuals really might not understand the full impact of a tax change until they do their taxes the following year or might not need to react to these increases immediately, but unemployment will have an immediate impact on an individual's labor market activities.

industry dummies crossed with a linear time trend and a linear time trend squared, respectively. One might expect that differences in industry composition would explain much of the difference between self-employment rates in Canada and the US. For this reason, the controls for industry fixed effects as well as time-varying industry characteristics, such as industry-specific technological change, are included.

The estimation results using both the median and 90th percentile tax rates for equation (1) are presented in Table 3. Column two of the table shows the estimated coefficients and the standard errors[28] using the tax rates at the 50th percentile and column three gives similar results for the tax rates calculated at the 90th percentile. In both regressions the coefficients on the average tax rates are positive and significantly different from zero at standard levels. However, the impact of an increase in the average tax rate at the 90th percentile on the probability of becoming self-employed is larger than the impact from the same tax increase (in percentage terms) at the median. The elasticities of male self-employment with respect to the average income tax rates at the median and 90th percentile using 1994 figures in Canada are 0.16 and 0.30, respectively. These same elasticities using US data are slightly larger at 0.21 and 0.37. These elasticities mean that a 30 percent increase in taxes in 1994 would lead to an increase in the rate of male self-

[28]As Moulton (1990) suggests, when a regression is fit to micro observations using both aggregate data and microdata as explanatory variables there is a possibility that the disturbances are correlated within the aggregate groups and even small correlations of this type can cause a large downward bias of the standard errors. The magnitude of the bias depends upon, among other things, the correlations of the regressors within groups. Within provinces/states in any given year in my data both the tax rates and the unemployment rates are perfectly correlated. The standard errors (in parentheses) are White's estimates with province/state-year cells as the primary sampling unit (i.e. the robust command in stata was used for clustered samples by state/year). These estimates account for this type of "group-wise" autocorrelation.

employment in Canada of 4.8 percent or 0.9 percentage points using average tax rates at the median and 9 percent or 2 percentage points using taxes at the 90th percentile. Similar estimates using US data are 6.3 percent or 0.8 percentage points using taxes at the median and 11.1 percent or 1.4 percentage points using taxes at the 90th percentile. To put such a tax increase into perspective, a 30 percent increase in family income tax is equivalent to a tax increase in 1992 Canadian dollars of $2500 for a family at the median and $7100 for a family at the 90th percentile in Canada or $2250 and $6650 in the US over 1994 levels.

An increase in the unemployment rate also had a positive effect on self-employment in both regressions. However, the magnitude of the increase in self-employment resulting from an increase in the unemployment rate was smaller than that from an increase in taxes. The elasticity of self-employment associated with the unemployment rate is equal to about 0.1 in both countries using 1994 figures. This means that a decrease of 5 percentage points in the unemployment rate (about the same decline that occurred between 1983 and 1989 in the US) leads to about a 1 percentage point decrease in the self-employment rate.

Other results shown in Table 3 are consistent with those of previous researchers.[29] I find that: (1) Age has a positive effect on the probability of self-employment and the effect is similar in both countries. It might be the case that older workers have accumulated entrepreneurial abilities, savings and business links making them more likely to be self-

[29]See Aronson (1991) for a review of this literature.

employed.[30] (2) Increases in educational attainment lead to increases in the probability of being self-employed in the United States. Unlike what is typically found in studies on US males, increases in education had almost no effect on the probability of being self-employed for Canadian males. (3) The more children present in the family the more likely males were to be self-employed. This fact was particularly true if the children were younger-- less than age 7. This might be because self-employment allows for greater flexibility in hours and the ability to work at home which gives workers the chance to take care of younger children. (4) Married men were most likely to be self-employed followed by men who were divorced widowed or separated. This is probably because married men are more likely to be in a family with a second income and fringe benefits that extend coverage to the entire family. A spouse's earnings and fringe benefits provide easier access to capital and allow greater risks to be taken. (5) Though not presented in the table, males in construction trades were most likely to be self-employed followed by retail trade and service industries.

The positive and significant effects of the tax environment and economic conditions on the probability of self-employment for North American males found here are robust for different specifications. Table 4 presents coefficient estimates for the tax rates and unemployment rates using a number of different samples or specifications[31]. Panel A,

---

[30]Another issue dealt with in the self-employment literature related to age is the effects of retirement on self-employment. Researchers have found that retired individuals are more likely to be self-employed than non-retirees (for example, Parnes and Less 1985 and Iams 1987). To control for retirement status and the trend towards early retirement equation (1) was re-estimated for males age 25-54. Excluding older males from the sample had little effect on the parameter estimates (see appendix Table A1)

[31]A complete list of variables and their associated standard errors are included in the appendix Table A2

included for comparison, presents the same coefficients and standard errors as in Table 3 using the average tax rates calculated on a family earning the median and 90th percentile of income for the sample which includes both the incorporated and unincorporated self-employed. Panel B gives similar results where only individuals who are self-employed in incorporated businesses are included in the sample while Panel C includes only those who are self-employed in unincorporated businesses. Panel D redefines which individuals in the sample are considered self-employed in an attempt to capture individuals who may be "part-time" self-employed or who are self-employed in a secondary job. Here individuals who had non-zero self-employment income from unincorporated businesses are considered self-employed.

In almost all cases, the coefficients on the lagged average tax rate and the unemployment rate are positive and significant at the five percent level. The one exception is Panel B where only males self-employed in incorporated businesses are considered. The coefficients on both the tax and unemployment rates in Panel B are small and insignificant. The most likely explanation for this result is the fact that firms tend to start out as unincorporated firms and only incorporate as the firm grows. Thus, the link between self-employment status and the tax environment or aggregate demand is diluted. The results in Panel D suggest that tax exemptions and write offs alone, which are available to both full and part-time self-employed, are not enough of an incentive to lure some of these males into self-employment. Panel D shows that the "part-time" self-employed males are less responsive to changes in the tax environment. While still highly significant, the coefficients on the

lagged average tax rates are dampened when both "full-time" and "part-time" unincorporated self-employment status is included as the dependent variable compared to Panel C where only "full-time" self-employed in an unincorporated business are considered on the left-hand-side. Part-time self-employment, besides allowing a number of tax write offs, does not enable individuals to under-report as large a proportion of income to tax authorities as those engaged in self-employment full-time.

The positive relation found here between the probability of self-employment and the average tax rates is contingent upon the assumption that the industry-specific time trends are the same in both countries. In other words, within industries, factors like technological change are assumed to have the same effect across the two countries, which is likely to be the case given the very strong links between the two economies of Canada and the US. This assumption, while not ideal, compares favourably to previous research on this topic. Only when fairly strict controls are introduced into the model by allowing the industry-specific time trends to vary across the two countries do the coefficients on the average tax rates become small and statistically insignificant.[32] However, relaxing this assumption is likely to eliminate much of the variation in the average tax rates. Controlling for within-industry time trends by country while simultaneously including controls for fixed effects across provinces/states is likely to soak up much of the variation in the tax data.

---

[32]The coefficients and standard errors on lagged average taxes for a family earning the median and the 90th percentile are -.0343 (.0634) and .0695 (.0621) when the industry-specific time trends are crossed with a country indicator. The standard errors correcting for group-wise autocorrelation are in parentheses.

Supporting Evidence

In order to show that the positive correlation found here between the income tax rates and the probability of self-employment is not simply a spurious relationship I present two pieces of supporting evidence.

First, one might expect that individuals with greater potential productivity (higher income) would be affected more by increases in income tax rates associated with income at higher points in the income distribution than individuals with lower potential productivity. Therefore, the fact that the elasticities of self-employment with respect to the average tax rates at the 90th percentile are larger than those with respect to the median tax rates is encouraging. However, because income tax rates here are taken from a single point in the income distribution and applied to males at all points in the income distribution, we can not conclude from this that males at the top of the income distribution are more responsive to changes in the tax rates.

In order to examine more closely the possibility that males with higher productivity are in fact more responsive to changes in the tax code at the 90th percentile, the tax variables are interacted with education category indicators in equation (1). This method links income tax rates to productivity by using education as a proxy for productivity. The results, which are presented in the top half of Table 5, confirm my expectation and provide additional support for the tax finding. Males in the highest education category (those with any post-secondary schooling) appear to be more responsive to changes in the tax rate at the 90th percentile. The coefficient on average income tax rates at the 90th percentile interacted with

the highest education category is significantly different than zero and results in a tax effect that is more than twice the size (0.27 versus 0.12) of those interacting the tax at the 90th with the other education categories. An F-test on the joint significance of the tax-education interaction coefficients reveals that these coefficients are jointly significant at the 5 percent level when the tax is calculated at the 90th percentile. As we might expect, increases in educational attainment had weaker effects on the tax coefficient calculated at the median. The coefficient on the tax-education interaction terms were individually not significantly different than zero and were jointly insignificant at the 5 percent level.

Second, it is undoubtably the case that the self-employed in different industries or occupations have different abilities to avoid taxes. For instance, those proprietors whom the US Internal Revenue Service (IRS) call "informal suppliers" or "individuals who provide products or services through informal arrangements which frequently involve cash-related transactions or 'off the books' accounting practices"[33] have a greater ability to under-report income. One would expect, therefore, that the tax effect in the current study would vary in terms of magnitude across industries-- with industries characterized by these "informal suppliers" having the largest coefficients.

In the bottom half of Table 5 the tax variable is crossed with industry indicator variables. What I find is that the effect of the tax rate on the probability of self-employment varies significantly across industries. F-tests for the null hypothesis that the coefficients are

---

[33]U.S. Internal Revenue Service, 1996, p.43.

the same are rejected at standard levels. Among the industries in which the tax effect is the largest are the retail trade, construction and transportation industries. The industries in which the tax effect was lowest include finance/insurance/real estate, manufacturing and wholesale trade.

Interestingly, the rank ordering by industry of the IRS estimates of the percentage of taxes under-reported among proprietors conform with the findings in Table 5. Under the US Taxpayer Compliance Measurement Program (TCMP) stratified random samples of income tax returns are subjected to intensive audits. A breakdown of the estimates of under-reporting by proprietors from data collected through the TCMP is provided in US General Accounting Office (1990). According to this report the industries in which the percentage of tax under-reporting was the greatest include: Retail Sales (fixed location) 39%, Transportation 36%, Retail Sales (no fixed location) 31% and production (including construction). The lowest estimates of under-reporting were recorded among the wholesale trade 19%, and finance/insurance/real estate 16%. The remarkable similarities in ranking by industry between Table 5 and the IRS estimates provide strong evidence that the correlation found between tax rates and self-employment is not a spurious relationship.

(7) Decompositions

This section describes a simple decomposition used to determine what fraction of the

overall change in predicted male self-employment rates between 1983 and 1992[34] is explained by movements in average tax rates, unemployment rates and the demographic and industrial composition of the employed. Results for the decomposition are presented for the changes in the predicted self-employment rates in Canada and the US, separately, as well as for the change in the gap between the Canadian and US predicted self-employment rates that existed over this period.

We can examine aggregate predicted self-employment rates for each country separately in any given year by averaging each of the variables in equation (1) as follows:

$$\overline{SE}_{tc}=\alpha+\beta_1*\overline{ltax(PP)}_{tc}+\beta_2*\overline{urate}_{tc}+\beta_3*\overline{ind}_{tc}+\beta_4*\overline{dem}_{tc} +\beta_5*\overline{dem}^{C}_{tc}+\beta_6*\overline{reg}_{tc}+\beta_7*\overline{indt}_{tc}+\beta_8*\overline{indtsq}_{tc} \quad (2)$$

Let t index the year and c index the country. The α and β's are the parameter estimates from equation (1). Then, $\overline{SE}_{tc}$ is the predicted self-employment rate in year t and country c given the average characteristics of the individuals in that year and country. Also, we can define groupings of the independent variables by summing the components as follows:

$$\overline{SE}_{tc}=\alpha+\beta^T*\overline{X}^T_{tc}+\beta^U*\overline{X}^U_{tc}+\beta^D*\overline{X}^D_{tc}+\beta^I*\overline{X}^I_{tc}+\beta^{UN}*\overline{X}^{UN}_{tc} \quad (3)$$

Here, the X's replace variable names. The superscripts indicate groupings of variables: T is

---

[34]These years were chosen because both 1983 and 1992 were relative troughs in the business cycle (i.e. years in which the unemployment rate reached a local maximum) in Canada and the US. This allows me to focus on factors which best explain secular, rather than cyclical, changes in self-employment rates.

the tax variable, U is the unemployment rate, D represents the demographic components including region, I represents the industry fixed effects and UN represents the industry specific time trends, which I think of as the unexplained component.[35]

The change in the predicted rate of self-employment in a given country between any two years $\{t,\tau\}$, for $\tau>t$, can, therefore, be written:

$$\overline{SE}_{\tau c}-\overline{SE}_{tc}=(\overline{X}^{T}_{\tau c}-\overline{X}^{T}_{tc})\beta^{T}+(\overline{X}^{U}_{\tau c}-\overline{X}^{U}_{tc})\beta^{U}+(\overline{X}^{D}_{\tau c}-\overline{X}^{D}_{t})\beta^{D} +(\overline{X}^{I}_{\tau c}-\overline{X}^{I}_{tc})\beta^{I}+(\overline{X}^{UN}_{\tau c}-\overline{X}^{UN}_{tc})\beta^{UN} \quad (4)$$

Then, for example, $\frac{(\overline{X}^{T}_{\tau c}-\overline{X}^{T}_{tc})\beta^{T}}{\overline{SE}_{\tau c}-\overline{SE}_{tc}}$ is the fraction of the overall change in the predicted self-employment rate between τ and t that is explained by the change in the provincial/state average tax rates. We can calculate a similar fraction for changes in unemployment rates, demographics, and the industrial composition of the work force as well as for the unexplained portion. We can examine the fraction of the change in the gap between the Canada and the US male self-employment rates explained by each of these components by differencing the average characteristics in equation (3) by country so that the left-hand-side of (3) becomes the Canadian male predicted self-employment rate in year t minus the US male predicted self-employment rate in year(t).

Table 6 shows the results of the decomposition described above for Canada (Panel

[35]To the extent that technological change is industry-specific but the same on either side of the Canada-US border, these can also be thought of as representing technological change.

A), the US (Panel B) and the Canada-US gap (Panel C) for the years 1983-1992. Rather than repeating the exercise for both sets of tax calculations the decomposition results are shown for the average tax rates calculated on a family with median income only. The first column of each of the panels gives the total predicted change attributable to each of the components (e.g. $(X_{tc}^{T}-X_{ic}^{T})\beta^{T}$) while the second column gives the fraction of the change that is due to changes in each of the components. The lower half of the table breaks down the demographic grouping into its components. As one might expect, the model does not predict the changes in the raw self-employment rates very well. For example, the model predicts only about 35 percent of the rise in the self-employment rate that occurred in Canada between 1983 and 1992. The model is a bit better in predicting the decline in the US rate and the increase in Canada-US gap between the two years. Approximately 50 percent of the actual changes in the US rate and 45 percent of the gap are predicted by the model. However, the results still provide a guide to which of the observed factors examined here has the greatest explanatory power.

In Canada (Panel A), changes in the provincial tax rates between 1983 and 1992 explained the largest fraction of the predicted change in male self-employment rates among the factors examined here. Changes in the average tax rates accounted for 192% of the overall change in the self-employment rate. This suggests that, holding all other factors constant, the male self-employment rate in Canada would have increased by almost twice as much as it did between 1983 and 1992 given the changes that occurred in the provincial average tax rates. In Canada, unlike in the US, changes in the industrial composition of the

employed also helps substantially in explaining the increase in the male self-employment rate. Changes in the unemployment rate, the demographic composition of the workforce and the unexplained portion offset the effects of the average tax rates and industry sector shifts. Each of these offsetting components would have led to a decrease in the rate of male self-employment in Canada between the two years had it been the only variable to change.

The overall fraction explained by changes in the demographic composition of the employed was negative in Canada. However, this masks some of the effects of the individual components that make up the demographic grouping. For instance, changes in the age and education structure of the employed in Canada accounted positively for the overall change in male self-employment. The positive effects of age and education were dominated by the negative effects that changes in the number of children, marital status and province of residence had on the overall rate of male self-employment in Canada.

In the US (Panel B), as in Canada, changes in the average tax rates explained the largest fraction of the overall predicted change (in this case a decline) in the male self-employment rate between 1983 and 1992. Changes in tax rates between 1983 and 1992 in the US accounted for 199% of the overall change in the predicted male self-employment rate. Changes in the industrial composition of workers in the US did not explain the decline in the male self-employment rate between 1983 and 1992. In fact, the results suggest that the male rate of self-employment would have increased if the only factor that changed over the period had been the composition of industries' employment shares. Unlike in Canada, changes in the unemployment rate in the US explained some (79%) of the change in the male self-

employment rate over this period. Like in Canada, changes in the demographic composition of the employed accounted for a negative fraction of the overall change in male self-employment. Also as in Canada, changes in age and education of the US workforce suggest that the rate of male self-employment should have risen while changes in the number of children, marital status and state led to a decline in the rate between 1983 and 1992. However, in the US the effects of shifts in age and educational attainment among the employed dominated the effects of the other demographic changes.

Finally, the decomposition is used to explain the gap between the Canada and US male self-employment rates which widened between 1983 and 1992. The raw numbers suggest that the gap between the two countries' male self-employment rates widened by more than 1 percentage point between 1983 and 1992. The model predicts about a 0.6 percentage point increase in the gap between the two years. As with the individual country analyses, changes in the provincial/state average tax rates account for the largest fraction of the increase in the predicted gap. Changes in average tax rates account for 197% of the predicted increase in the gap between the two self-employment rates. Changes in the unemployment rates account for a relatively small fraction (37%) of the increase in the gap. Shifts in both the demographic and industrial employment shares of the two work forces countered the effects of the tax and unemployment rates on the Canada-US male self-employment rate gap between 1983 and 1992.

Overall, the results from the decompositions show that changes in the average tax rates consistently explain a large fraction of the predicted shifts in the Canada and US male

self-employment rates as well as the gap that opened up between the two rates from 1983 to 1992. The role that unemployment rates played in explaining changes in the male self-employment rates is not quite as clear. In Canada economic conditions explained a negative fraction of the rise in predicted male self-employment while in the US economic conditions had at least some explanatory power. In Canada, shifts in the industrial composition of the employed appears to explain a significant fraction of the increase in male self-employment while changes in the demographics of the workers did not help to explain the increase. Both composition effects did not explain any of the decline in the US rate of male self-employment nor the gap between the Canada and US rates.

(8) Conclusions

The literature on self-employment to this point has primarily focused on factors that have global effects for most developed economies. However, the evidence suggests that no single common factor is responsible for the trends. For this reason, this essay has focused on region-specific factors-- namely the tax environment and economic conditions-- as possible causes for the trends in male self-employment in North America. This examination improves upon previous studies which have examined the effects of taxes and economic conditions on self-employment by incorporating province or state as well as cross-country variation in the tax and unemployment data. The results presented here provide evidence that changes in the tax environment explain a considerable amount of the secular trends in male self-employment in North America while economic conditions explain less of these trends.

The empirical analysis shows that even with fairly strict controls for industry characteristics, increases in average income tax rates have positive and large effects on the rate of male self-employment. The estimated effect of increasing taxes by 30 percent is an increase in the rate of male self-employment in incorporated and unincorporated businesses of between 4.8 and 11.1 percent. This suggests that one of the motivations for becoming self-employed is the relative tax advantages associated with self-employment. In fact, the decompositions demonstrate that changes in average tax rates are the largest contributing factor of the possible determinants examined here for the secular trends in self-employment in Canada and the US. The finding that the probability of self-employed in a secondary job is less responsive to increases in income taxes than the same probability in a main job implies that tax sheltering alone is not enough of an incentive to lure some North American males into self-employment. Instead, for some it is the relative ease of under-reporting income in self-employment that is the factor determining self-employment status among these males.

The results also support the notion that North American males turn to self-employment to some extent during spells of high unemployment. It could be that individuals experiencing unemployment find this transition to be a convenient time to become self-employed or that self-employment is simply employment of last resort. It appears, however, that economic conditions had a smaller role in determining self-employment among these males than the tax environment did. Further, the unemployment rates did not explain much in terms of the secular trends in self-employment in Canada and the US over this period, as

was illustrated by the decompositions.

A number of policy implications arise from these findings. First, raising income taxes may result in increased numbers of workers moving into the self-employment sector where their labor income can be taxed at a lower rate. This will leave fewer tax paying workers which, in turn, may require greater-than-expected increases in income taxes. Second, the fact that self-employment appears to provide employment during downturns suggests that policies that provide assistance to fledgling entrepreneurs may assist in alleviating the particularly harmful negative employment effects of recessions. Not surprisingly, however, this policy prescription should be regarded as highly tentative for a number of reasons. First, it is not clear from this analysis whether or not this finding is a result of an increase in the actual number of self-employed individuals. It could be that jobs in the self-employment sector are simply more insulated against demand shocks than wage and salary jobs. Therefore, in recessions the rate of self-employment may rise because the number self-employed holds constant while the total number of individuals employed falls. Second, supposing that new jobs are created in the self-employment sector during recessions, we are unable to discern from this analysis how stable these newly created self-employment jobs actually are. These jobs could be temporary and, therefore, not worthy of assistance. It seems that an analysis that includes a longitudinal component would be effective in providing answers to these questions. In any case further analysis is required to sort these issues out.

Table 1
Non-Agricultural Self-Employment, 1973-1990: Percentage Point Changes

	1973-1983	1983-1990	1973-1990	1990 level
COUNTRY				
Australia[a]	2.6	0.3	2.9	12.4
Austria	-3.6	-1.7	-5.3	6.4
Belgium	1.1	0.6	1.7	12.9
Canada[a]	0.9	0.3	1.2	7.4
Denmark	-0.8	-1.3	-2.1	7.2
Finland	0.6	1.8	2.4	8.8
France[a]	-0.9	-0.2	-1.1	10.3
Germany	-1.7	0.3	-1.4	7.7
Ireland	0.6	2.6	3.2	13.3
Italy	-2.4	1.6	-0.8	22.3
Japan[a]	-0.8	-1.8	-2.6	11.5
Luxembourg	-2.3	-1.7	-4	7.1
Norway[a]	-1	-0.7	-1.7	6.1
Portugal	4.3	1.5	5.8	18.5
Spain	0.7	0.1	0.8	17.1
Sweden	0	2.2	2.2	7
UK[b]	1.3	3	4.3	11.6
US[a]	1	-0.1	0.9	7.6

a) Excluding owner-managers of incorporated businesses
b) Excluding some owner-managers of incorporated businesses
Source: OECD, Labour Force Statistics, 1970-1990, Paris, 1992.

Table 2
Sample Characteristics
Non-Primary Self-Employment Rates And Employment Shares
by Industry/Demographic Group (Males 1983-1992)

	CANADA				UNITED STATES			
	1983		1992		1983		1992	
Age	rate	share	rate	share	rate	share	rate	share
25-35	11.8	38.9	12.9	36.9	9.8	40.1	8	37.2
36-45	18.1	27.6	19.2	32.4	15	27	14.6	31.6
46-55	20.2	21.2	22.4	21	17.6	19.4	18.4	20.3
55-64	19.4	12.3	25.4	9.7	18.3	13.5	19.9	11
Education								
0-8 Years	18.3	16.2	22	7.6	10.8	7.5	12.6	4.3
9-10 Years	16.5	13.9	19.4	11.7	11.9	6.8	14.1	4.2
11-13 Years	15.2	30.5	16.9	39.5	11.6	39.7	11.4	37.3
Any Post Sec.	16.1	39.4	18.3	41.3	16.6	46	15	54.2
Marital Status								
Single	10.7	11.1	10.4	16.5	9.5	14.3	7.8	18.7
Married	17	83.8	19.7	77.6	14.8	74.9	15	69.3
Div/Wid/Sep.	16.4	5.1	19.7	5.9	12.9	10.8	13.6	12
Children <7yrs								
0	17.1	76.5	17.9	76.3	14.4	79	13.9	79.8
1	12.3	15.1	17.7	14.3	11.9	14.1	11.5	13.8
2	14.9	7.4	20.5	8.1	11.2	6.1	12.9	5.7
3 or More	18.4	1	20.3	1.4	16.3	0.9	11.8	0.7

Table 2 Continued

	CANADA				UNITED STATES			
	1983		1992		1983		1992	
Children 7-17	rate	share	rate	share	rate	share	rate	share
0	15	60.1	17	68.1	13.4	64.5	13.1	67.8
1	17.7	18.7	18.6	16	14.2	17.8	14	16.4
2	18.3	15.2	21.9	12.4	15.3	12.7	14.8	11.6
3 or More	18.9	5.9	24.3	3.6	14.6	5	14	4.3
Industry								
Non-Durables	4.5	11	7.5	9.4	3.8	9.3	4.1	8.4
Durables	5.4	13.7	4.5	12.7	3.7	17.6	4	14.4
Construction	39.8	8.3	44.5	9	23.5	11.5	27.2	11.5
Transportation	9.4	12.9	10.9	12.5	7.7	10.9	6.9	10.4
Whol. Trade	19.7	6.1	23.9	6.9	15.3	5.7	15.5	5.3
Retail Trade	31	9.9	26.4	11.1	21.6	11.5	16.2	13.1
Fin/Ins./Real Es	18.3	4.7	23.1	5	20.3	5.3	18.6	5.3
Services	22.8	23.6	23	24.9	21.8	22.3	19.4	24.8
Public Admin.	0	9.8	0	8.4	0	5.9	0	6.9

Note: Data calculated using sample weights from the SCF and CPS micro-data files

Table 3
Regression Results: Base Model
Linear Probability (OLS), Pooled Data

Variable	Incorporated + Unincorporated Self-Employed	
	(50th)	(90th)
Lagged Tax Rate	**0.181**	**0.200**
	(0.054)**	(0.040)**
Regional Unemployment Rate	**0.194**	**0.184**
	(0.042)**	(0.040)**
Age	**0.015**	**0.015**
	(0.001)**	(0.001)**
Age X Country	**-0.001**	**-0.001**
	(0.001)	(0.001)
Age Squared/1000	**-0.125**	**-0.126**
	(0.006)**	(0.006)**
Age Squared/1000 X Country	**0.006**	**0.008**
	(0.013)	(0.013)
Ed. (9-10 Years)	**0.027**	**0.027**
	(0.003)**	(0.003)**
Ed. (9-10 Years) X Country	**-0.020**	**-0.020**
	(0.005)**	(0.005)**
Ed. (11-13 Years)	**0.048**	**0.048**
	(0.003)**	(0.003)**
Ed. (11-13 Years) X Country	**-0.046**	**-0.047**
	(0.004)**	(0.004)**
Ed. (Post Secondary)	**0.082**	**0.082**
	(0.003)**	(0.003)**
Ed. (Post Sec.) X Country	**-0.080**	**-0.081**
	(0.005)**	(0.005)**
# Children (aged 7-17)	**0.001**	**0.001**
	(0.001)	(0.001)
# Children (7-17) X Country	**0.006**	**0.006**
	(0.001)**	(0.001)**
# Children (aged <7)	**0.007**	**0.007**
	(0.001)**	(0.001)**
# Children (<7) X Country	**0.006**	**0.006**
	(0.002)**	(0.002)**

Table 3 Continued		
Marital Status (married)	**0.030**	**0.030**
	(0.002)**	(0.002)**
(married) X Country	**-0.005**	**-0.004**
	(0.004)	(0.004)
Marital Status (div/wid/sep)	**0.010**	**0.010**
	(0.002)**	(0.002)**
(div/wid/sep) X Cntry...	**0.006**	**0.006**
	(0.005)	(0.005)
N	**487062**	**487062**
R - squared	**0.1**	**0.1**

Notes:
(1)Indicator variables for province/state and industry dummies crossed with a time trend were included but not presented here.
(2)Values in parentheses are White's estimators treating province/state-year cells as the primary sampling unit.
(3)**indicates significance at the 5 percent level and *indicates significance at the 10 percent level.

Table 4
Regression Results: Various Specifications
Linear Probability (OLS), Pooled Data

Variable	Panel A Incorp + Unincorp		Panel B Incorporated Only		Panel C Unincorp. Only		Panel D F/T & P/T S-E	
	(50th)	(90th)	(50th)	(90th)	(50th)	(90th)	(50th)	(90th)
Lagged Tax	0.181 (0.054)**	0.200 (0.040)**	0.016 (0.050)	0.047 (0.040)	0.192 (0.040)**	0.188 (0.028)**	0.110 (0.044)**	0.133 (0.037)**
Unemployment Rate	0.194 (0.042)**	0.184 (0.040)**	0.048 (0.033)	0.042 (0.033)	0.184 (0.037)**	0.177 (0.035)**	0.130 (0.044)**	0.130 (0.044)**
N	487062	487062	441154	441154	460090	460090	450867	450867
R - squared	0.1	0.1	0.06	0.06	0.07	0.07	0.06	0.06

Notes: (1)Indicator variables for province/state and industry dummies crossed with a time trend were included but not presented here.
(2)Values in parentheses are White's estimators treating province/state-year cells as the primary sampling unit.
(3)**indicates significance at the 5 percent level and *indicates significance at the 10 percent level.

Table 5 Lagged Tax Rate Crossed with Education and Industry Dummies		
Variable	(50th)	(90th)
Tax Crossed With Education		
Lagged Tax	0.125 (0.103)	0.121 (0.074)*
Lagged Tax X Ed. (9-10 Years)	-0.020 (0.095)	-0.003 (0.076)
Lagged Tax X Ed. (11-13 Years)	-0.003 (0.089)	0.033 (0.065)
Lagged Tax X Ed. (Post Secondary)	0.125 (0.097)	0.145 (0.071)**
Unemployment Rate	0.195 (0.042)**	0.185 (0.040)**
R - squared	0.1	0.1

Table 5 Continued		
Tax Rate Crossed With Industry		
Variable	(50th)	(90th)
Lagged Tax X Manuf. Non-Durables	0.095 (0.057)*	-0.072 (0.037)*
Lagged Tax X Manufacturing Durables	0.050 (0.076)	0.003 (0.062)
Lagged Tax X Construction	0.499 (0.160)**	0.727 (0.123)**
Lagged Tax X Transportation	0.321 (0.070)**	0.284 (0.053)**
Lagged Tax X Wholesale Trade	0.123 (0.109)	0.183 (0.088)**
Lagged Tax X Retail Trade	0.733 (0.143)**	0.907 (0.096)**
Lagged Tax X Fin./Insur./Real Estate	-0.357 (0.127)**	-0.122 (0.107)
Lagged Tax X Services	-0.046 (0.070)	-0.023 (0.052)
R - squared	0.1	0.1
N	487062	487062

Notes: (1)Regression includes the same variables as the base case (eq. 1) but the results are not presented here.

(2)Values in parentheses are White's estimators treating province/state-year cells as the primary sampling unit.

(3)** indicates significance at the 5 percent level and * indicates significance at the 10 percent level.

Table 6
Decomposition Results: 1983-1992, Canada, United States and Gap

Variable	Panel A Canada		Panel B United States		Panel C Gap (Canada-U.S.)	
	Change 1992-1983	Fraction Predicted	Change 1992-1983	Fraction Predicted	Change 1992-1983	Fraction Predicted
Raw Self-Employment Rate	0.0051	----	-0.0071	----	0.0122	----
Predicted Self-Employment Rate	0.0018	1.00	-0.0037	1.00	0.0055	1.00
Tax Rates	0.0035	1.92	-0.0074	1.99	0.0108	1.97
Unemployment Rates	-0.0009	-0.52	-0.0029	0.79	0.002	0.37
Demographics	-0.0010	-0.56	0.0052	-1.41	-0.0062	-1.13
Industry	0.0032	1.79	0.0050	-1.34	-0.0018	-0.32
Unexplained Portion	-0.0029	-1.64	-0.0036	0.97	0.0006	0.12
Breakdown of Demographics						
Age	0.0016	0.88	0.0021	-0.57	-0.0005	-0.10
Education	0.0002	0.12	0.0045	-1.21	-0.0043	-0.78
# Children	-0.0008	-0.46	-0.0001	0.04	-0.0007	-0.12
Marital Status	-0.0007	-0.39	-0.0013	0.35	0.0007	0.12
Province	-0.0013	-0.71	-0.0001	0.03	-0.0014	-0.25

Notes: (1) The first column of each panel equals the total change attributable to each of the components (e.g. $(X^T_{\tau c}-X^T_{tc})\beta^T$) or $((X^T_{\tau C}-X^T_{tC})-(X^T_{\tau U}-X^T_{tU}))\beta^T$)

Figures

Figure (1)

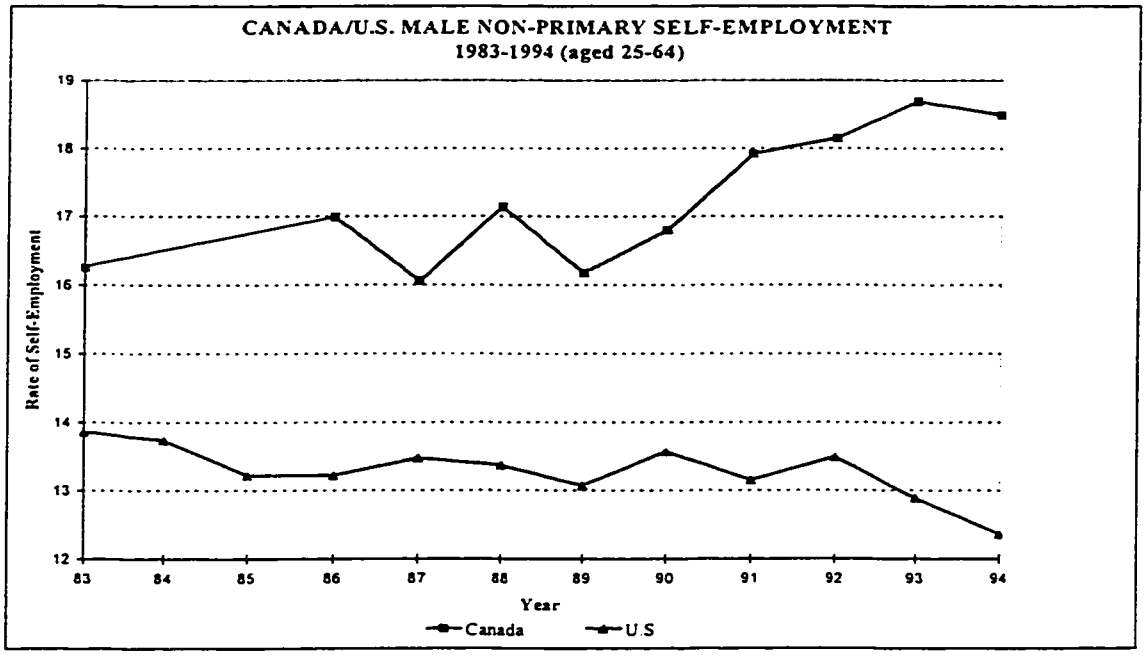


Figure (2)

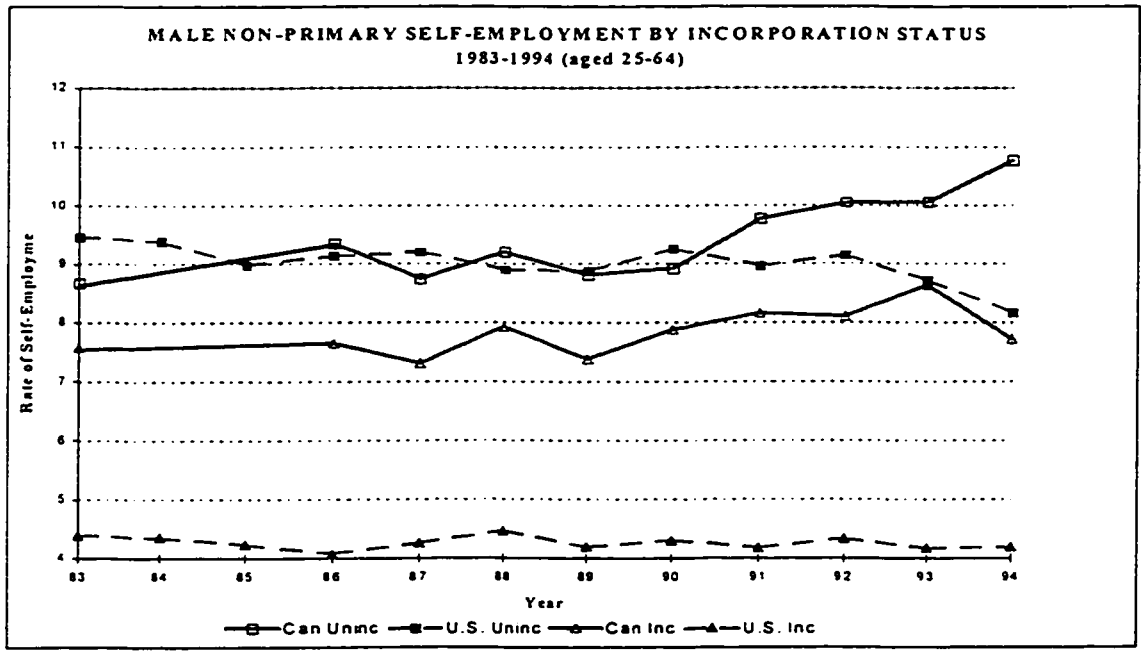


Figure (3)

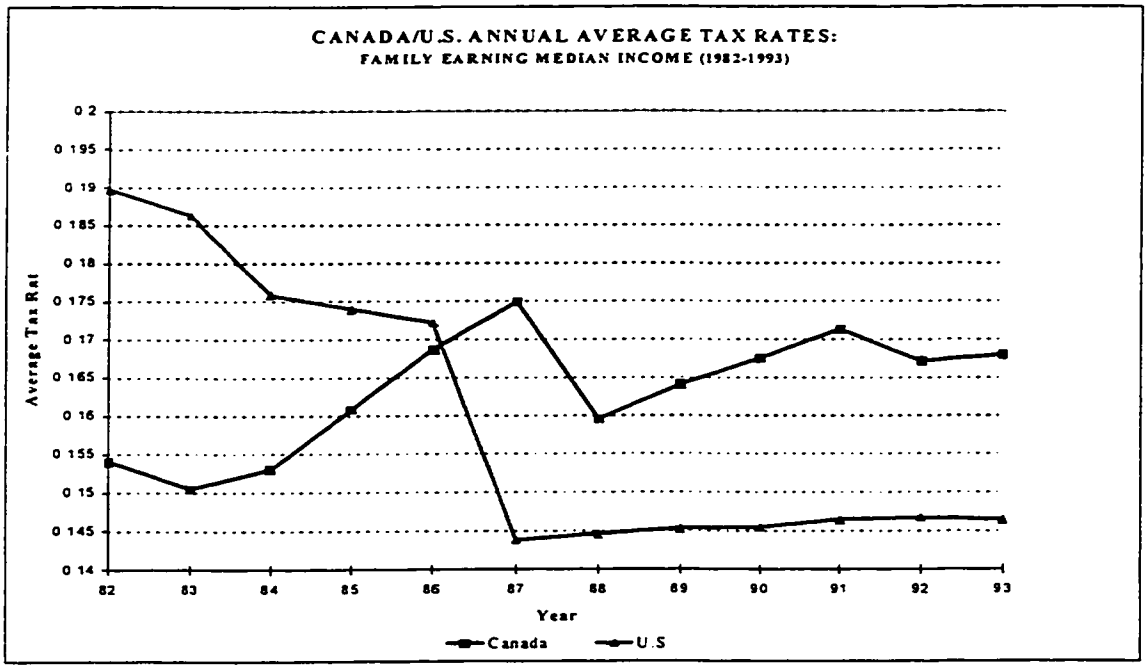


Figure (4)

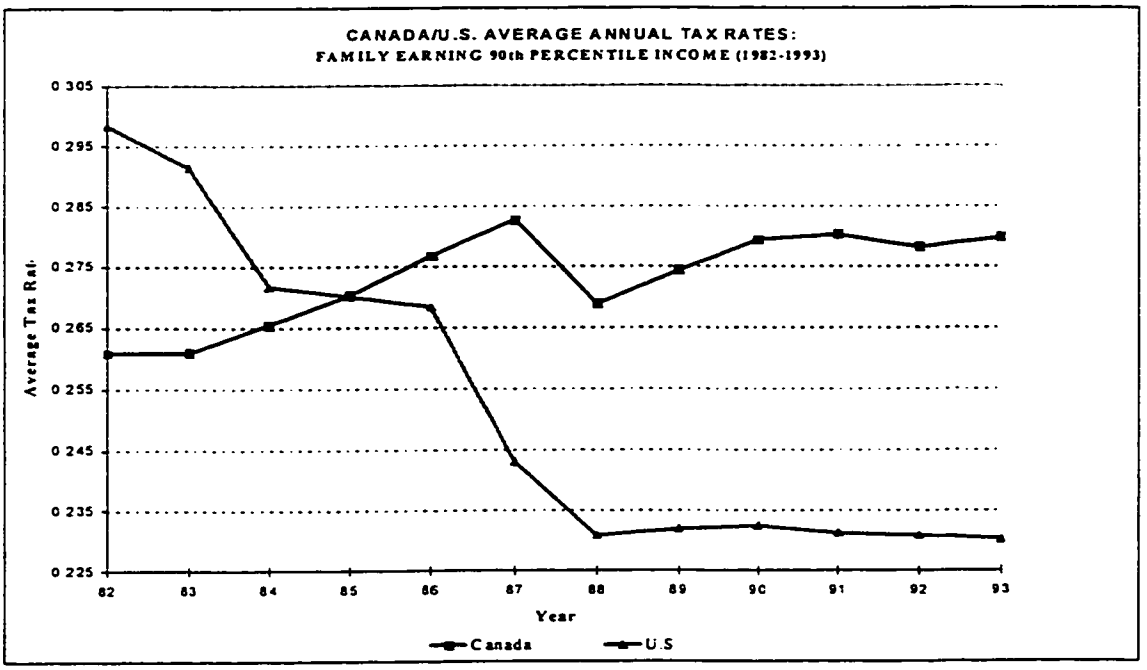


Figure (5)

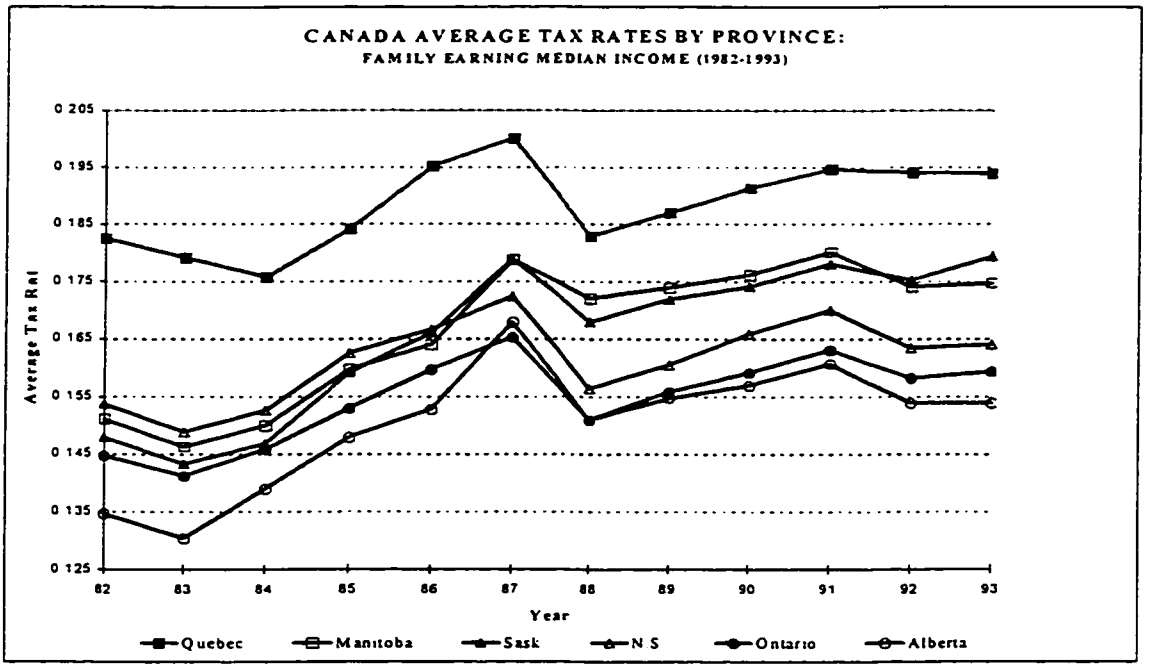


Figure (6)

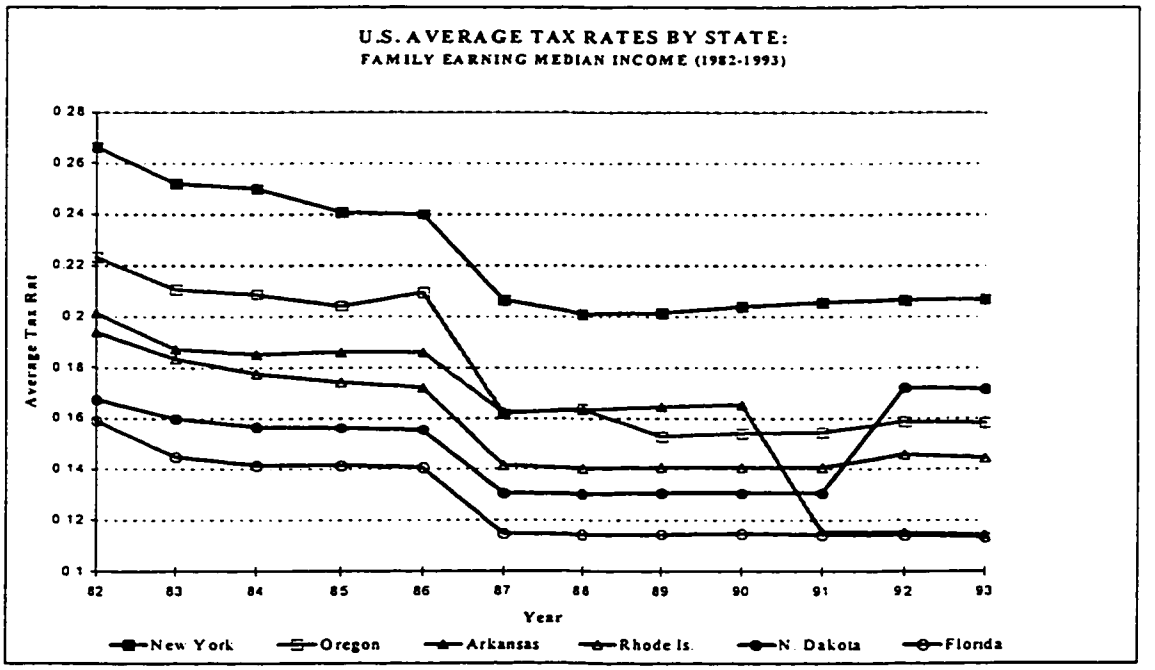


Figure (7)

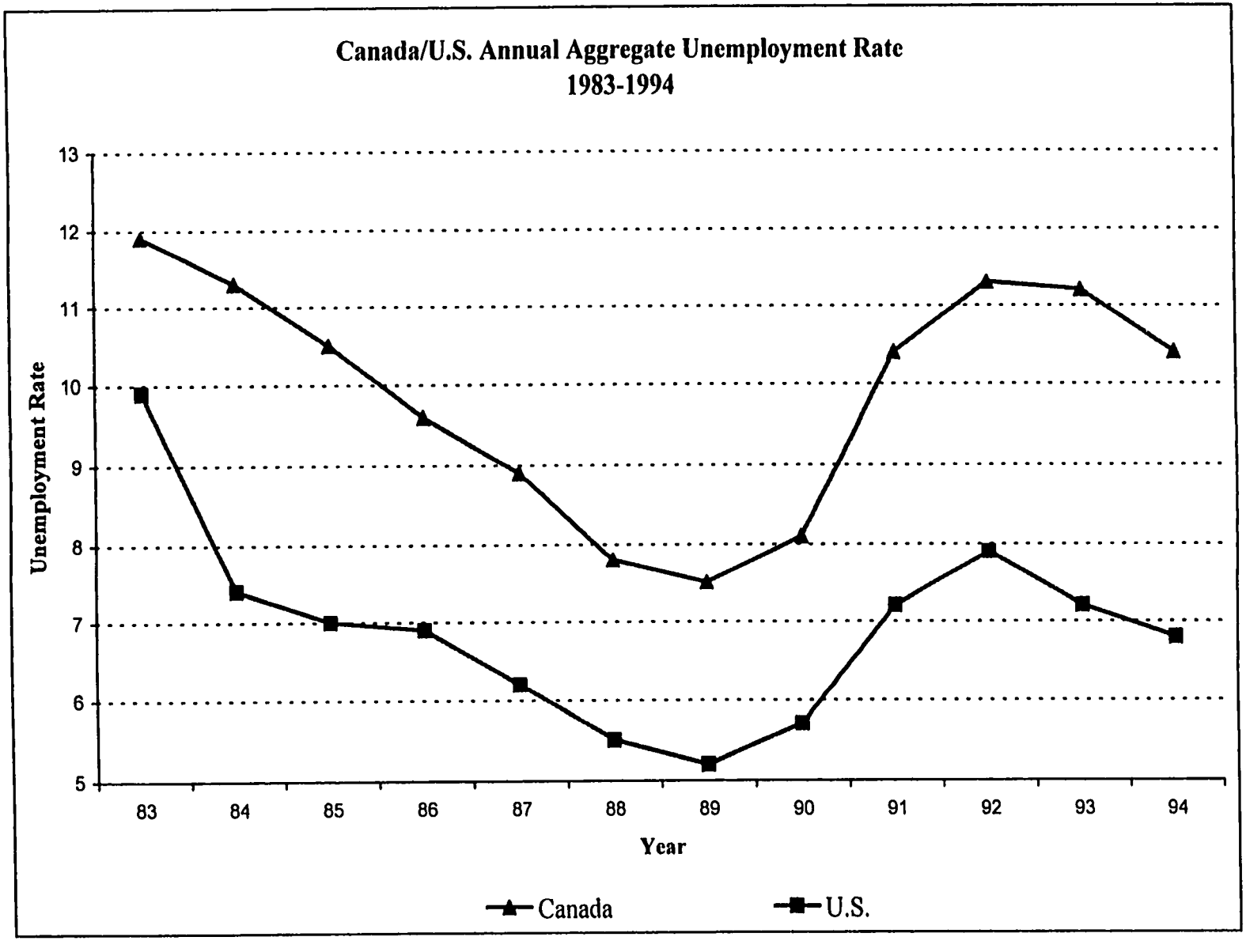

Canada/U.S. Annual Aggregate Unemployment Rate
1983-1994
Unemployment Rate
13
12
11
10
9
8
7
6
5
83
84
85
86
87
88
89
90
91
92
93
94
Year
Canada
U.S.

Figure (8)

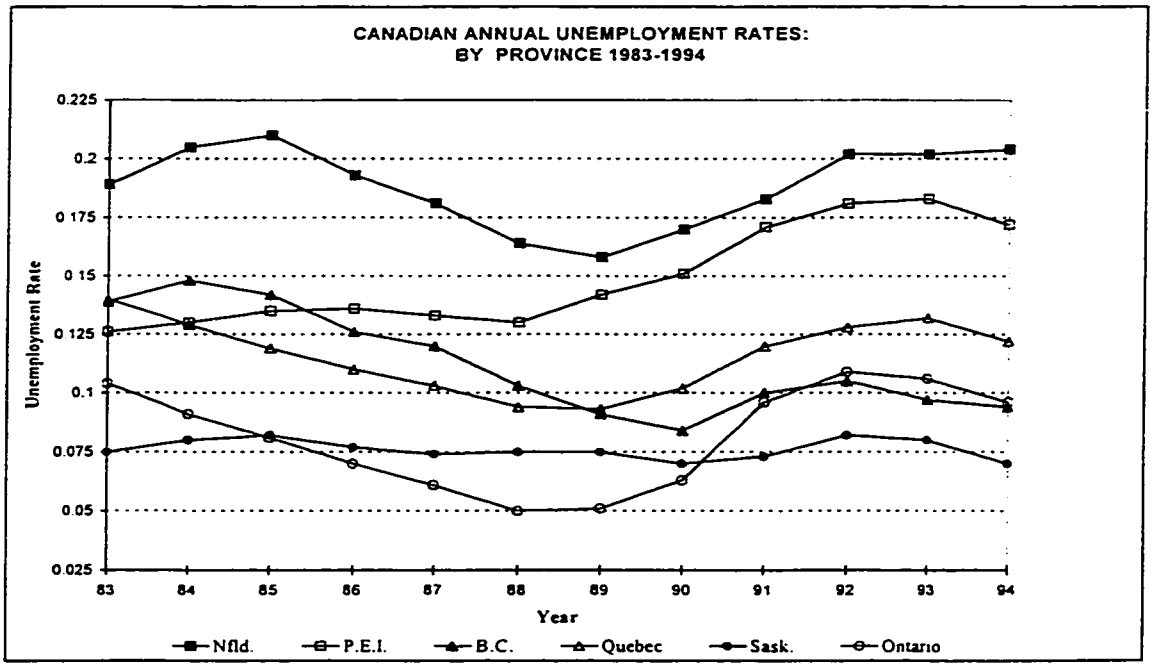
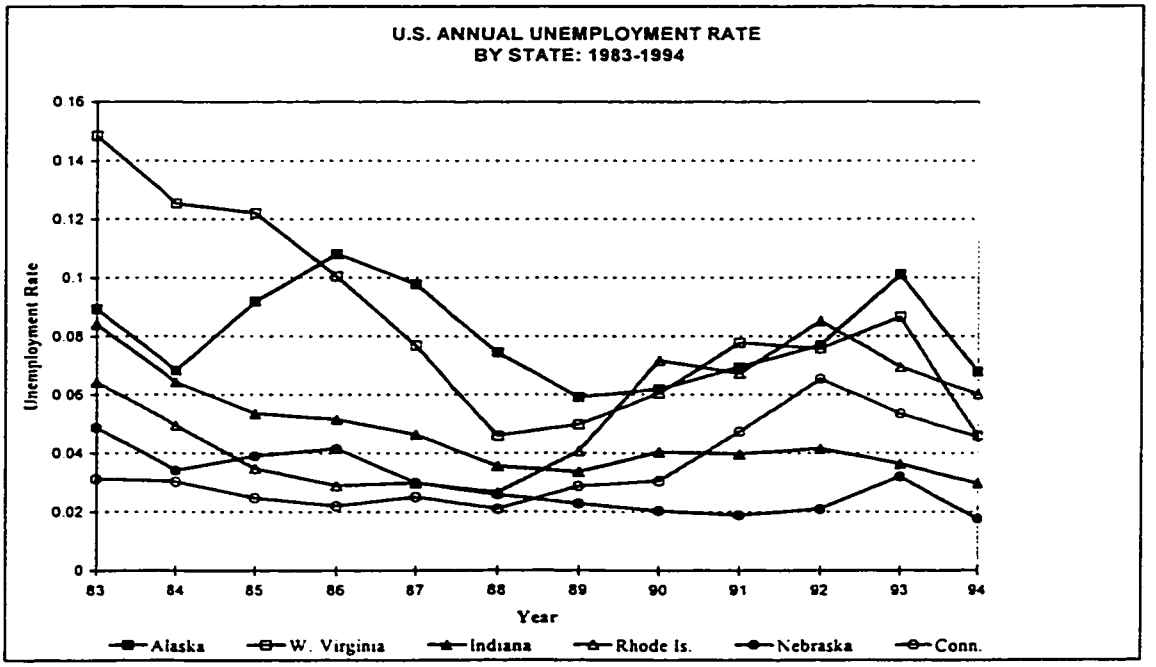


Figure (9)



References

Aronson, Robert L. (1991) Self-Employment ILR Press, Ithaca, New York

Becker, Eugene, (1984) "Self-Employed Workers: An Update to 1983", *Monthly Labor Review* 107 pp. 14-18

Bishop, John H., (1987) "American Job Growth: What Explains It?", *Portfolio: International Economic Perspectives* 12

Blanchflower, D. and Oswald, A., (1990) "Self-Employment and the Enterprise Culture", in British Social Attitudes: the 1990 Report, edited by R. Jowell, S. Witherspoon and L. Brook, Gower.

Blau, David (1987) "A Time Series Analysis of Self-Employment in the United States", *Journal of Political Economy*, Vol 95, no.3, pp. 445-467

Borjas, George J., and Stephen G. Bronars (1989) "Consumer Discrimination and Self-Employment", *Journal of Political Economy* 97 pp. 581-605

Canadian Tax Foundation (1982-1993) The National Finances: An Analysis of the Revenues and Expenditures of the Government of Canada Canadian Tax Foundation Publications, Toronto

Crompton, Susan (1993) "The Renaissance of Self-Employment" *Perspectives on Income and Employment* Statistics Canada Cat. 75-001E, pp. 22-32

Devine, Theresa (1993) "The Recent Rise in U.S. Self-Employment", mimeo, The Pennsylvania State University

Devine, Theresa (1990) "The Recent Rise in Female Self-Employment", mimeo, The Pennsylvania State University

Devine, Theresa and Joyce Mlakar (1993) "Inter-Industry Variation in the Determinants of Self-Employment", mimeo, The Pennsylvania State University

Erard, Brian (1997) "A Critical Review of the Empirical Research on Canadian Tax Compliance" Working Paper 97-6, Business Income Tax Division Department of Finance Canada.

Evans, David S., and Linda S. Leighton (1989) "Some Empirical Aspects of Entrepreneurship", *American Economic Review*, vol. 79, pp. 519-535

Feenberg, Daniel Richard, and Elizabeth Coutts (1993) "An Introduction to the TAXSIM Model", *Journal of Policy Analysis and Management*, vol. 12 no. 1, pp. 189-194.

Iams, Howard M. (1987) "Jobs of Persons Working After Receiving Retired Work Benefits", *Social Security Bulletin* 50 (November) pp. 4-19

Lin, Zhengxi, Janice Yates and Garnett Picot (1998) Rising Self-Employment in the Midst of High Unemployment: An Empirical Analysis of Recent Developments in Canada", Working Paper, Business and Labour Market Analysis Division Statistics Canada.

Long, James E. (1982) "The Income Tax and Self-Employment", *National Tax Journal*. 35 (March): pp. 31-42.

Moulton, Brent R., (1990) "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *The Review of Economics and Statistics*, 72(2), pp. 334-338.

OECD (1992) *Employment Outlook: July 1992*, OECD Publication, Paris

Parnes, Herbert S., and Lawrence J. Less. (1985) "Economic Well-Being in Retirement." in Retirement Among American Men, ed. Herbert S. Parnes, pp. 91-118. Lexington, Mass.: Lexington Books.

Quinn, Joseph F., (1980) "Labor Force Participation Patterns of Older Self-Employed Workers", *Social Security Bulletin* 43 pp. 17-28

Rees, Hedley, and Shah, Anup, (1986) "An Empirical Analysis of Self-Employment in the U.K.", *Journal of Applied Econometrics* 1 pp. 95-108

U.S. Internal Revenue Service (1990) "Income Tax Compliance Research: Net Tax Gap and Remittance Gap Estimates", Publication 1415, Washington, D.C.

White, H., (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 pp. 817-830

Appendices

A. Employment Status Questions:

CANADA (SCF)

"In ...'s job, was he/she a paid worker, self-employed or an unpaid family worker?"

Worked for Others
1.Paid Worker (Private, Public)
2.Unpaid Family Worker

Self-Employment
Incorporated Business - with paid help
Incorporated Business - no paid help
Unincorporated Business - with paid help
Unincorporated Business (include self-employed without business) - no paid help

US (CPS)

"What was ...'s longest job during 19..?"

Class of Worker

Paid	Self-Employment	
1. Private	1. Incorporated?	→ Yes
2. Federal Government	2.	→ No
3. State Government	3. Without Pay	
4. Local Government		

B. Tax Calculations:

Both for Canada and the US, income is assumed to come from employment and the family is assumed to have no dependents. In Canada, the couple is assumed to file independently, deductions are taken for CPP or QPP and UI premiums and, a sales tax credit introduced in 1986 and the Goods and Service Tax credit in 1991 are reflected in the calculations. However, no provincial low income deductions, cost of living credits or sales tax credits are included in the tax calculations for Canada. In the US, the couple is assumed to file jointly and TAXSIM covers ordinary and super tax brackets, earned income credits, secondary earner deductions and other important features of the US tax code.

C. Variable Description:

Industry Indicators: manufacturing non-durables (omitted industry), manufacturing durables, construction, transportation/communication, wholesale trade, retail trade, finance/insurance/real-estate, services and public administration

Demographic Variables: The demographic variables include age, age squared, dummy variables for education (0-8 years (omitted group), 9-10 years, 11-13 years and any post secondary), number of young children (aged less than 7), number of older children (aged 7 to 17), and dummy variables for marital status (single (omitted group), married and divorced/widowed or separated).

Region: The regions include the 10 provinces in Canada and the 50 states and the District of Columbia in the United States. The omitted region is Alabama.

Time and Time Squared: **time** and **timesq** are the time trend and the time trend squared for the omitted industry.

TABLE A1
Regression Results: Includes Both Lagged and Current Tax & Unemployment Rates
Linear Probability (OLS), Pooled Data

Variable	Regression 1 (tax at 50th)	Regression 2 (tax at 90th)
tax(PP)	0.0202 (0.0435)	0.0266 (0.0307)
ltax(PP)	0.1104 (0.0421)	0.1011 (0.0299)
urate	0.1344 (0.0591)	0.1341 (0.0590)
lurate	0.0584 (0.0566)	0.0554 (0.0564)
manufacturing durables	0.0021 (0.0085)	0.0019 (0.0085)
construction	0.2568 (0.0090)	0.2566 (0.0090)
transportation	0.0399 (0.0088)	0.0401 (0.0088)
wholesale trade	0.1209 (0.0105)	0.1210 (0.0105)
retail trade	0.2336 (0.0089)	0.2336 (0.0089)
finance/insurance/real estate	0.1508 (0.0114)	0.1507 (0.0114)
services	0.1672 (0.0078)	0.1672 (0.0078)
public administration	-0.0650 (0.0098)	-0.0647 (0.0098)
age	0.0155 (0.0004)	0.0155 (0.0004)
age^2	-0.0001 (0.0000)	-0.0001 (0.0000)
education: 9-10 years	0.0146 (0.0026)	0.0145 (0.0026)
education: 11-13 years	0.0204 (0.0022)	0.0204 (0.0022)
education: post secondary	0.0446 (0.0022)	0.0447 (0.0022)
# children aged 7-17	0.0027 (0.0006)	0.0027 (0.0006)
# children aged<7	0.0074 (0.0008)	0.0074 (0.0008)
marital status: married	0.0293 (0.0017)	0.0292 (0.0017)
marital status: div/widow/sep	0.0123 (0.0023)	0.0123 (0.0023)
manufacturing durables X time	0.0018 (0.0029)	0.0018 (0.0029)
construction X time	-0.0026 (0.0031)	-0.0026 (0.0031)
transportation X time	0.0023 (0.0030)	0.0023 (0.0030)
wholesale trade X time	0.0021 (0.0036)	0.0021 (0.0036)
retail trade X time	-0.0042 (0.0030)	-0.0042 (0.0030)

Table (A1) Continued

fin/ins/real estate X time	0.0003 (0.0039)	0.0003 (0.0039)
services X time	0.0003 (0.0027)	0.0003 (0.0027)
public administration X time	0.0031 (0.0033)	0.0030 (0.0033)
manuf durables X time^2	-0.0002 (0.0002)	-0.0002 (0.0002)
construction X time^2	0.0005 (0.0002)	0.0005 (0.0002)
transportation X time^2	-0.0002 (0.0002)	-0.0002 (0.0002)
wholesale trade X time^2	-0.0001 (0.0003)	-0.0001 (0.0003)
retail trade X time^2	-0.0001 (0.0002)	-0.0000 (0.0002)
fin/ins/real estate X time^2	-0.0001 (0.0003)	-0.0001 (0.0003)
services X time^2	-0.0002 (0.0002)	-0.0002 (0.0002)
public administration X time^2	-0.0002 (0.0003)	-0.0002 (0.0003)
time	0.0006 (0.0024)	0.0009 (0.0024)
time^2	-0.0000 (0.0002)	-0.0000 (0.0002)
constant	-0.4705 (0.0151)	-0.4818 (0.0155)
R^2	0.0996	0.0996

Notes: (1) Dummy variables for province/state were included but not recorded here
(2) Values in parenthesis are standard errors
(3) X indicates that the variable has been crossed with another variable

TABLE (A2)
Regression Results: Table (3) Continued
Self-Employment Indicator as Dependent Variable Linear Prabability (OLS), Pooled Data

Variable	50th	90th	Variable	50th	90th
Alaska	**0.022**	**0.023**	**Iowa**	**0.001**	**0.000**
	(0.007)	(0.007)		(0.008)	(0.008)
	[0.007]	[0.007]		[0.006]	[0.006]
Arizona	**-0.003**	**-0.004**	**Kansas**	**0.014**	**0.013**
	(0.008)	(0.008)		(0.007)	(0.007)
	[0.007]	[0.007]		[0.005]	[0.005]
Arkansas	**0.040**	**0.038**	**Kentucky**	**0.010**	**0.011**
	(0.008)	(0.008)		(0.008)	(0.008)
	[0.005]	[0.005]		[0.006]	[0.006]
California	**0.015**	**0.009**	**Louisiana**	**0.000**	**0.000**
	(0.006)	(0.006)		(0.008)	(0.008)
	[0.004]	[0.005]		[0.007]	[0.007]
Colorado	**0.006**	**0.005**	**Maine**	**0.038**	**0.034**
	(0.007)	(0.007)		(0.008)	(0.008)
	[0.005]	[0.005]		[0.006]	[0.006]
Connnecticut	**0.014**	**0.013**	**Maryland**	**-0.017**	**-0.018**
	(0.008)	(0.008)		(0.007)	(0.007)
	[0.006]	[0.005]		[0.005]	[0.005]
Delaware	**-0.019**	**-0.022**	**Mass.**	**0.005**	**0.004**
	(0.008)	(0.008)		(0.006)	(0.006)
	[0.007]	[0.007]		[0.004]	[0.005]
D.C.	**-0.033**	**-0.035**	**Michigan**	**-0.010**	**-0.010**
	(0.008)	(0.008)		(0.006)	(0.006)
	[0.007]	[0.006]		[0.004]	[0.004]
Florida	**0.019**	**0.020**	**Minnesota**	**-0.008**	**-0.011**
	(0.006)	(0.006)		(0.008)	(0.008)
	[0.006]	[0.005]		[0.006]	[0.006]
Georgia	**-0.017**	**-0.018**	**Mississippi**	**0.019**	**0.018**
	(0.007)	(0.007)		(0.008)	(0.008)
	[0.006]	[0.006]		[0.007]	[0.007]
Hawaii	**-0.021**	**-0.024**	**Missouri**	**-0.006**	**-0.006**
	(0.008)	(0.008)		(0.007)	(0.007)
	[0.006]	[0.006]		[0.006]	[0.006]
Idaho	**0.022**	**0.019**	**Montana**	**0.031**	**0.028**
	(0.008)	(0.008)		(0.008)	(0.008)
	[0.009]	[0.009]		[0.009]	[0.009]
Illinois	**-0.015**	**-0.014**	**Nebraska**	**0.012**	**0.009**
	(0.006)	(0.006)		(0.008)	(0.008)
	[0.004]	[0.004]		[0.008]	[0.008]
Indiana	**-0.010**	**-0.010**	**Nevada**	**-0.026**	**-0.025**
	(0.007)	(0.007)		(0.008)	(0.008)
	[0.006]	[0.006]		[0.006]	[0.006]

Table (A2) Continued

Variable	50th	90th	Variable	50th	90th
New Hampshire	0.034	0.035	Utah	-0.002	-0.003
	(0.008)	(0.008)		(0.008)	(0.008)
	[0.006]	[0.006]		[0.008]	[0.008]
New Jersey	0.004	0.003	Vermont	0.038	0.034
	(0.006)	(0.006)		(0.009)	(0.009)
	[0.005]	[0.005]		[0.007]	[0.007]
N.M.	0.018	0.015	Virginia	-0.011	-0.013
	(0.008)	(0.008)		(0.007)	(0.007)
	[0.008]	[0.008]		[0.005]	[0.005]
New York	-0.013	-0.015	Washington	0.013	0.014
	(0.006)	(0.006)		(0.008)	(0.008)
	[0.006]	[0.006]		[0.007]	[0.007]
North Carolina	0.011	0.009	West Virginia	-0.029	-0.031
	(0.006)	(0.006)		(0.008)	(0.008)
	[0.005]	[0.005]		[0.006]	[0.006]
North Dakota	0.011	0.009	Wisconsin	-0.006	-0.009
	(0.008)	(0.008)		(0.007)	(0.007)
	[0.006]	[0.006]		[0.007]	[0.007]
Ohio	-0.008	-0.010	Wyoming	0.019	0.020
	(0.006)	(0.006)		(0.009)	(0.009)
	[0.004]	[0.004]		[0.009]	[0.009]
Oklahoma	0.023	0.021	Newfoundland	0.037	0.035
	(0.008)	(0.008)		(0.020)	(0.020)
	[0.007]	[0.007]		[0.023]	[0.023]
Oregon	0.029	0.025	P.E.I.	0.080	0.079
	(0.008)	(0.008)		(0.020)	(0.020)
	[0.006]	[0.006]		[0.024]	[0.024]
Pennsylvania	-0.009	-0.008	Nova Scotia	0.083	0.082
	(0.006)	(0.006)		(0.019)	(0.019)
	[0.004]	[0.004]		[0.023]	[0.023]
Rhode Island	0.007	0.003	New Brunswick	0.072	0.070
	(0.008)	(0.008)		(0.019)	(0.019)
	[0.005]	[0.005]		[0.023]	[0.023]
South Carolina	-0.011	-0.013	Quebec	0.089	0.083
	(0.007)	(0.007)		(0.019)	(0.019)
	[0.007]	[0.007]		[0.023]	[0.023]
South Dakota	0.031	0.032	Ontario	0.099	0.098
	(0.008)	(0.008)		(0.019)	(0.019)
	[0.009]	[0.009]		[0.023]	[0.022]
Tennessee	0.003	0.004	Manitoba	0.111	0.109
	(0.007)	(0.007)		(0.019)	(0.019)
	[0.005]	[0.005]		[0.023]	[0.023]
Texas	0.005	0.006	Saskatchewan	0.124	0.122
	(0.006)	(0.006)		(0.019)	(0.019)
	[0.005]	[0.005]		[0.023]	[0.023]

Table (A2) Continued

Variable	50th	90th	Variable	50th	90th
Alberta	**0.134**	**0.134**	**dindtsq3**	**0.000**	**0.000**
	(0.019)	(0.019)		(0.000)	(0.000)
	[0.023]	[0.023]		[0.000]	[0.000]
B.C.	**0.140**	**0.140**	**dindtsq4**	**0.000**	**0.000**
	(0.019)	(0.019)		(0.000)	(0.000)
	[0.023]	[0.023]		[0.000]	[0.000]
dindt2	**0.002**	**0.002**	**dindtsq5**	**-0.001**	**-0.001**
	(0.002)	(0.002)		(0.000)	(0.000)
	[0.002]	[0.002]		[0.000]	[0.000]
dindt3	**0.001**	**0.001**	**dindtsq6**	**0.000**	**0.000**
	(0.003)	(0.003)		(0.000)	(0.000)
	[0.006]	[0.006]		[0.000]	[0.000]
dindt4	**0.002**	**0.002**	**dindtsq7**	**0.000**	**0.000**
	(0.002)	(0.002)		(0.000)	(0.000)
	[0.003]	[0.003]		[0.000]	[0.000]
dindt5	**0.007**	**0.007**	**dindtsq8**	**0.000**	**0.000**
	(0.003)	(0.003)		(0.000)	(0.000)
	[0.004]	[0.004]		[0.000]	[0.000]
dindt6	**-0.002**	**-0.002**	**dindtsq9**	**0.000**	**0.000**
	(0.002)	(0.002)		(0.000)	(0.000)
	[0.005]	[0.005]		[0.000]	[0.000]
dindt7	**0.001**	**0.001**	**time**	**0.000**	**0.000**
	(0.003)	(0.003)		(0.001)	(0.001)
	[0.005]	[0.005]		[0.002]	[0.002]
dindt8	**0.001**	**0.001**	**timesq**	**0.000**	**0.000**
	(0.002)	(0.002)		(0.000)	(0.000)
	[0.003]	[0.003]		[0.000]	[0.000]
dindt9	**0.004**	**0.003**	**_cons**	**-0.501**	**-0.526**
	(0.001)	(0.001)		(0.014)	(0.015)
	[0.002]	[0.002]		[0.018]	[0.018]
dindtsq2	**0.000**	**0.000**			
	(0.000)	(0.000)	**N**	**487062**	**487062**
	[0.000]	[0.000]	**R - square**	**0.10**	**0.10**

Notes: (2)Values in round parenthesis are White's standard error estimators
(3)Values in square parenthesis are White's estimators treating province/state-year cells as the primary sampling unit.

Chapter 3

SELF-EMPLOYMENT DYNAMICS AND SELF-EMPLOYMENT TRENDS: A STUDY OF CANADIAN MEN AND WOMEN, 1982-1995

(1) Introduction

Between 1982 and 1995 the number of self-employed Canadians between the ages of 25 and 54 increased by 70 percent, compared to a 45 percent increase in paid employment.[1] What explains this increase in self-employment? In this essay we shed some light on this question by examining hitherto-unexploited data on the dynamics of in- and outflows from self-employment over this time period. Data on flows provide clues on the causes of change that are not available by examining stocks alone, because they describe the mechanisms by which individuals become, or cease to be, self-employed, and allow us to examine how these mechanisms have changed over time.

The data used in this essay are the Surveys of Consumer Finances for the years 1982 to 1995. These files contain standard labour force data for the week prior to the survey, as well as supplemental data on the previous year's work experience and income. As a result, we are able to observe individuals in contiguous years, giving us a 2-year panel for each

[1]These figures, like all those in the current essay, exclude individuals working in primary industries. Self-employment refers to an individual's main job, and to unincorporated businesses only. Primary industries consist of agriculture, forestry, fisheries and mining.

individual. We analyze the dynamics of self-employment as a Markov process among three labour force states: employment in the wage-and-salary sector (E), self-employment (S), and not employed (N). The Markov model allows us to easily compute steady-state self-employment rates associated with the transition processes observed at any point in time, and to decompose changes in these steady-state rates into portions attributable to various changes in the transition matrix.

Our main findings are as follows. First, as we might expect, the steady-state rate of self-employment increased between the 1980's and 1990's for both women and men, though by more for women. Second, perhaps surprisingly, the process by which these rates increased was quite different. Between the 1980's and 1990's, prime-age men's transition rates from employment to nonemployment increased substantially. Because non-employed men are more likely than employed men to become self-employed, this can account for most of the increase in the steady-state rate of male self-employment. In contrast, the steady-state rate of self-employment among women rose primarily because of decreased exit from self-employment, and (of somewhat less importance) a rise in entry into self-employment.

The association of men's increased self-employment with a decrease in the permanence of paid jobs, and of women's increased self-employment with higher survival rates in self-employment, suggests, more broadly, that men's behavioural changes were, in part, a response to a secular deterioration in labour market conditions, while women were responding to an improving market. Supporting evidence for this view is provided by two further results in our essay. First, multinomial logit analysis of the key transition

probabilities identified above shows that changes in observable demographic characteristics, such as age, education, and immigration, cannot explain the changes in these probabilities between the 1980's and 1990's. This leaves room for other factors, such as general labour market conditions, to play a central role. Second, we present evidence on secular changes in the *quality* of new self-employment opportunities, measured by three labour market outcomes –earnings, hours, and the presence of employees-- for individuals who have been self-employed for less than a year. All three indicators show a deterioration for men and an improvement for women between the 1980's and 1990's.

The remainder of this essay is organized as follows. Section 2 briefly reviews the existing literature on the determinants of self-employment levels and flows. Section 3 describes the data, and presents simple descriptive statistics on self-employment entrants and leavers. Section 4 presents the basic transition matrices, by period and gender, and computes steady-state rates of self-employment based on these probabilities. Decompositions of the steady-state rates of self-employment are performed in Section 5, identifying *which* changes in transition rates played key roles in explaining steady state changes in self-employment rates. In Section 6 we use a multinomial logit model to assess the effect of demographic changes on these key probabilities. Section 7 presents supporting evidence on trends in the quality of new self-employment opportunities, and Section 8 concludes.

(2) Previous Literature

Most previous studies of self-employment trends focus exclusively on secular

changes in the *stock* of self-employed individuals (e.g. Blau (1987), Fairlie and Meyer (1998) for the US; Lin, Yates and Picot (1998a) and Schuetze (1998) for Canada; Blanchflower (1998) for a variety of OECD countries). These studies focus on a variety of possible causes, including shifting demographic composition of the labour force, tax laws, and general economic conditions, with rather mixed conclusions. A much smaller number of studies focus on flows in addition to stocks (e.g. Evans and Leighton 1989, Meyer 1990 and Alba-Ramirez 1991 for the US; Lin, Yates, and Picot 1998b and Moore and Mueller 1998 for Canada). These studies tend largely to be descriptive of the dynamics of self-employment and generally do not attempt to explain secular trends in self-employment.[2] This is partly due to the lack of panel data on self-employment that is consistent over long periods; it is unfortunate however, because changes in process, or flows, might contain important clues regarding the underlying causes of these secular changes in stocks.

(3) Data and Descriptive Statistics

As mentioned, we draw the 2-year panel data used in this essay from a series of microdata files from Canada for the years 1982 to 1995. The microdata files are taken from the Canadian Surveys of Consumer Finances (SCF's) which are conducted in April of each year and contain standard labour force data for approximately 75 thousand individuals per

---

[2]Lin, Yates and Picot (1998b) present some aggregate provincial statistics on changes in self-employment flows between 1981 and 1995 in Canada. The main focus of their analysis is, however, on the cyclical properties of these flows, i.e. on the effects of short-run fluctuations in various business cycle measures, net of time trends. The only other attempt to use flow data to understand longer-run changes in self-employment stocks of which we are aware is a study of British self-employment, by Blanchflower and Freeman (1994)

year. All extracts we employ are restricted to individuals aged 25 to 54; we focus on this group because it is less likely to be affected by secular increases in school attendance and a tendency to retire earlier. To facilitate the examination of long-run secular changes, and to increase parameter precision (transitions into and out of self-employment are relatively rare events), we pool all the surveys corresponding to the 1980's, and those for the 1990's, and simply compare the two periods to each other. Thus we work with four separate data files--two data files comprising the 1982-1989 surveys, with 104 thousand observations on men and 118 thousand observations on women; and two data files comprising the 1990-1995 surveys with 118 thousand men and 132 thousand women. Both of these periods roughly include a relative "trough" in the business cycle as well as a relative "peak". Overall, however, employment prospects in the 1990's in Canada were somewhat worse than the 1980's. The average unemployment rate among Canadians aged 25 to 54 in the 1980's was about 10 percent versus 10.5 percent in the 1990's.

These data offer a number of advantages over other Canadian data sets that might be used to analyse self-employment dynamics. First, because they contain data for the week prior to the survey as well as supplemental data on the previous year's work experience and income, they enable us to estimate annual gross flows between labour force states. In addition, because the surveys are highly consistent in sampling design and questionnaire structure over many years, we are able to examine changes in these flows over time to evaluate the causes of secular changes in self-employment. Finally, the surveys allow us to observe a number of individual characteristics, such as education and place of birth, which

are not measured in administrative data sets.

One pecularity of using the SCF to study labour force dynamics is the fact that labour force status is not measured the same way in two years for which we observe each individual. For the week prior to the survey, the respondent is asked to report whether or not he or she was employed. If the individual was employed, he or she was asked whether in their "main job" he/she was self-employed or employed in the wage and salary sector. Using this information we assign each individual to one of three states: "employed" (but not self-employed) (E), "self-employed" (S), or "not employed" (N).[3]

For the calendar year prior to the survey, no direct question about self-employment in one's "main job" is asked. Instead, we have data on the number of weeks worked in the year, plus information on the amount and source of income in that year, and use this to impute whether the individual was self-employed or a wage and salary earner in his/her "main job". For the vast majority of individuals in our sample, assigning them a labour market status on the basis of this information is straightforward: most respondents either worked the full year or not at all, and had only one source of labour market income: wage and salary earnings or self-employment income. For part-year workers and those with multiple sources of labour market income, we proceeded as follows. First, we assigned to each individual a probability of working (being either "employed" or "self-employed") equal

---

[3]We did some preliminary work using a four-state model that distinguished unemployment from nonparticipation. After some experimentation, we concluded that the extra insight provided by such a model was not warranted by the very substantial increase in complexity. (Moving to a four-state model yields a sixteen-element, rather than a nine-element transition matrix, with some of its elements –due to small sample sizes– quite imprecisely measured.)

to the number of weeks worked divided by 52. Second, we allocated those weeks of work to self-employment or wage-and-salary employment according to the relative amounts of income earned from each. Thus each part-year worker, and each individual with multiple earned income sources, contributes more than one observation to our data, with the weights assigned to each observation given by our estimate of the probability they were in the corresponding labour market state in a randomly selected week during the previous year.

Clearly, there are some potential problems with the above approach. For example, in some instances self-employment income is negative. While this should still be interpreted as time in self-employment, it is likely to take less time to lose money than to earn it. We address this problem by recasting negative self-employment income as its absolute value times some fraction.[4] Another issue is seasonality: one week in April (the survey week) may not be representative of an entire year. Like the preceding issue however, this will primarily affect the *level* of self-employment transition rates at a point in time. Because the definitions of both previous- and current-year self-employment are the same across all years of the SCF, these issues should not materially affect our estimates of how transition matrices changed over time, which are our main interest in this essay.

A final issue is whether to allocate self- versus regular employment probabilities strictly in proportion to relative incomes from each: wages in the two jobs may not be equal. Further, if strict proportionality is used, an individual who, throughout the entire year,

---

[4]The results reported here use 1/4, but –because negative incomes are relatively rare– the results are highly insensitive to the value used. For instance, in the 1994 survey only 491 of the over 34 thousand individuals reporting earned income had negative self-employment earnings.

worked both part-time in the self-employment sector and full-time in the wage and salary sector would be designated as self-employed in their "main job" for some fraction of the previous year, but would never be so classified according to the "survey week" definition. In the results reported here we correct for these measurement differences by utilizing the fact that our two-year panels overlap. The (year-specific) proportionality measure used to allocate individuals between paid- and self-employment in the results reported here is the one that forces self-employment rates by either of our two measures *for the same year* to be the same. The details of this procedure are provided in Appendix A; it is worth noting however that this correction is made only to refine our main estimates. Indeed, when we replicated our main analysis for the population of individuals with only one source of income in the previous year, the results were virtually unchanged, indicating a lack of sensitivity to the assumptions made to allocate individuals with multiple sources of income between self- and paid-employment.[5]

Table 1 documents the main phenomenon we are attempting to explain-- increasing self-employment rates among both men and women between the 1982 and 1995 SCF surveys. Over this period, the number of self-employed men and women rose by 54 and 108 percent respectively, compared to increases in paid work of 30 and 64 percent respectively.

The first two rows of Table 2 describe the source states of new inflows into self-

---

[5]These results are available from the authors. The insensitivity they demonstrate essentially results from two factors, one of which is the rarity of individuals with both self-employment and other employment income. For instance, in 1994 only 1604 individuals of the over 34 thousand individuals reporting earned income had both forms of income. The other factor, again, is the fact that these measurement issues primarily affect levels at a point in time, not (because they are consistent across surveys) changes over time.

employment, the destinations of those who leave self-employment, and (for comparison) the distribution of the entire population across labour market states. While most entrants to self-employment worked in the wage and salary sector in the previous year, entrants were more likely to be non-employed than the total population. This was particularly true of men: The fraction of men who were not employed prior to entering self-employment was 9 percentage points higher than the population average in the 1980's and grew to 12 percentage points in the 1990's. Also of interest, however, are the trends in the relative sources of men and women entering self-employment. A larger fraction of men entering self-employment in the 1990's came from non-employment relative to the 1980's-- 32 percent in the 1990's versus 25 percent in the 1980's. The opposite was true of women entering self-employment. The fraction of women entering self-employment from non-employment actually fell by 3 percentage points between the 1980's and 1990's. Finally, most men and women leaving self-employment re-enter into wage and salary employment. Likely because of a decline in labour market conditions in the 1990's, however, they were more likely to enter non-employment in the 1990's than in the 1980's.

The remainder of Table 2 simply gives the distributions of three populations –self-employment entrants, self-employment leavers, and the total populaiton, across demographic categories. Thus it appears that most Canadians entering self-employment are married, have no children, are Canadian born and tend to be younger than the population as a whole. There are some interesting gender differences however. For example, relative to the overall population, the self-employment sector attracted more highly educated women, while men

entering self-employment tended to be less educated. Also, women entering self-employment were more likely than the population as a whole to have children while the opposite was true of men.

(4) Transition Matrices

Table 3 reports the probability matrices **P** that summarize the transition rates among self-employment, wage-and-salary employment, and nonemployment in our samples. Elements (p_{ij}) of each 3x3 matrix give the empirical probability that an individual in state i at time t (the year preceding the survey) is in state j at time t+1 (the survey week). We report separate **P** matrices by gender and period (1980's versus 1990's), a total of four in all. Asterisks indicate which of the elements in the 1990's matrix are significantly different at the five percent level from the relevant 1980's elements. Table 3 also reports the ergodic distribution of individuals across these three states under the assumption that the transitions among the labour force states are governed by a Markov process, i.e. that there is no state dependence-- each p_{ij} depends only on the current state and not on history. This ergodic distribution is calculated as the eigenvector $\boldsymbol{q}$ associated with the unit eigenvalue such that:

$$\mathbf{P}\,\boldsymbol{q} = \boldsymbol{q} \qquad (1)$$

By definition, $\boldsymbol{q}$ must sum to one, and is so normalized. Finally, Table 3 also reports a steady-state rate of self-employment for each transition matrix, which is simply the proportion of time a representative individual spends in self-employment over the proportion of time spent in all forms of employment, as implied by the ergodic distribution across states.

Interestingly, the steady-state rates of self-employment in Table 3 approximate the actual self-employment rates quite well-- differing by less than one percentage point in all cases. As one might expect, the steady-state rates of self-employment rose for both men and women between the 1980's and 1990's. Also as one might expect, the increase was more dramatic among women than men. The steady-state rate of self-employment rose by 2.4 percentage points or 56 percent among women between the two periods as compared to 1 percentage point or 10 percent among men.

A number of observations with regard to the estimated transition matrices warrant mention here. In general, men's retention rate in self-employment, i.e. the proportion of self-employed men who remain self-employed one year later, is larger than women's. In addition, women's retention rate in non-employment, at almost 81 percent, is much higher than men's (63 percent in the 1980's and 70 percent in the 1990's). Trends in exit and entry rates to self-employment also varied greatly between men and women. Exit rates (SN and SE) among men either rose slightly or remained stable while these same rates fell for women. Further, the self-employment retention rate (SS) rose significantly for women and remained stable for men. At the same time, entry rates into self-employment (ES and NS) remained stable for both men and women.

There is also evidence that employment opportunities outside of the self-employment sector deteriorated for men and improved for women. For instance, men's retention rate in the wage and salary sector (EE) fell between the 1980's and 1990's, as did the proportion of males who exited non-employment for employment in the wage and salary sector (NE). In

comparison, these proportions actually rose slightly among women. In addition, exit from the wage and salary sector to non-employment (EN) and the retention rate among those not employed (NN) rose substantially among men but fell for women.

(5) Decompositions

In this section we ask which of the changes in the elements of the transition matrices, identified in the last section, can account for the secular rise in self-employment. The elements or groups of elements we examine are as follows. We start with the obvious potential determinants: changes in entry rates to, and exit rates from, self-employment. Next, we examine changes in labour market opportunities outside of the self-employment sector as these can also affect equilibrium self-employment rates. Here, we look at the durations and the transition rate between employment in the wage and salary sector and non-employment.

The decompositions are carried out as follows. We allow the particular elements of interest in the transition matrix pertaining to the 1980's to take on the probability value in the relevant 1990's transition matrix. By definition, each of the columns of the transition matrix must sum to one. Therefore, one cannot simply change one element in any given column and still maintain this requirement. Our approach when a single element in a column was to be changed to the 1990's probability, was to change that element to the 1990's level but maintain

the ratio of the other two probabilities in that column.[6] The ergodic distributions resulting from the newly formed transition matrix and the associated steady-state self-employment rate were then calculated. The fraction of the overall change in the steady-state rate of self-employment attributable to the change in any element(s) of the transition matrix is then estimated as the difference between the steady-state rate associated with the newly formed transition matrix and the rate for the 1980's, divided by the total predicted change between the two periods (i.e. when all the elements of the matrix are changed to the 1990's level). These estimates are presented in Table 4.[7] The first row of this table reports the overall predicted change in the steady-state rates of self-employment, rows two to five give the decomposition results allowing self-employment entry and exit rates to change, and rows six to eight give decomposition results for changes in the transitions between wage and salary employment and nonemployment.

It appears that, for Canadian males, changes in the probabilities of entering and exiting self-employment directly from/to the other two labour force states had little impact on the secular rise in the steady-state rate of self-employment. In fact, rows two to five of Table 4 suggest that if the only changes in the transition matrix between the 1980's and 1990's had been the probabilities of exiting and entering self-employment, the steady-state

---

[6]While not ideal this method seems preferable to changing the entire column to the new (1990's) level. Changing various combinations of elements, in our view, provides more insights into the various possible causal mechanisms at work.

[7]Results are also generated for the sample of individuals with only one of self-employment or wage and salary income in the previous year (available from the authors) and are very similar to those presented in table 4.

rate of self-employment would have declined. Both a decline in the entry rate and an increase in the exit rate contribute to this hypothetical decline. Instead, the factors which appear to have played the largest role in the sectoral rise in male self-employment rates are those elements which determine whether an individual is employed in the wage and salary sector or not employed at all. For instance, the rise in the probability of exiting the wage and salary sector into non-employment and the decline in the probability of entering the wage and salary sector from non-employment (row six of the table) account for 183 percent of the increase in the steady-state rate of self-employment between the two periods. Also, allowing the retention rate in wage and salary work and non-employment to adjust to the 1990's level accounts for 190 percent of the rise.[8]

The explanation for the rise in steady-state self-employment among women is more straightforward. For women, increased entry and decreased exit from self-employment are the two most important causes for the secular rise in self-employment. Together, entry and exit (row 2) account for 112 percent of the rise in the steady-state rate of self-employment between the two periods for women. As row four of the table shows, nearly all of that 112 percent rise can be attributed to a decline in exits from self-employment combined with an

---

[8]It may be worth noting that this predicted increase in the male self-employment *rate* can be further decomposed into a portion due to a changing numerator (i.e the number of self-employed persons) and denominator (the total number of employed persons). The two have somewhat distinct interpretations. For example, in the latter case, self-employment rates can increase in poor labour markets without any increase in the number of self-employed persons, simply because self-employed people are less likely to "lay themselves off". When we performed this decomposition, we found that nearly 60 percent of the increase is due to increase self-employment (the first mechanism) while the remaining 40 percent is due to decreased wage and salary employment. For this analysis we allowed only changes in the ergodic distribution that result from changing the flows between wage and salary employed and not employed to their 1990's level. This allowed us to isolate the two mechanisms described above from the negative effects of direct entry and exit on time spent in self-employment.

increase in the duration of self-employment. Surprisingly, unlike for men, changing transition patterns between the wage-and-salary sector and nonemployment had almost no effect on the steady-state self-employment rate for women. In fact, Table 4 suggests that the steady-state rate of self-employment would have fallen somewhat if only the transition probabilities between E and N had changed. This seems likely to be a result of the increase in labour force participation experienced by women over this period.

(6) The Role of Changing Demographics

One might argue that the changes in transition probabilities, and the resulting changes in steady-state self-employment rates analyzed in the last section, are simply an artifact of the changing demographic mix of the Canadian labour force. After all, the age distribution, education mix, and immigrant share of the population changed substantially over this period; perhaps these shifts account for most of the observed changes we find.

To assess this hypothesis, in this section we estimate a series of multinomial logit models of transition rates, and use these models to assess the affect of changing demographics on both the transition rates themselves and the resulting ergodic self-employment rate. While we present results for all nine elements of each transition matrix, in our discussion we focus mainly on those elements identified in the last section as key elements; namely, increased transitions from wage and salary employment to nonemployment among men, and decreased exit rates from self-employment among women

In more detail, our approach is as follows. First, using the data pertaining to the

1980's we estimate a model of time t+1 choice among the three labour force states on various demographic characteristics by multinomial logit for men and women, separately. The demographic variables include age and age squared, a set of dummy variables for education, marital status, immigration status, the presence of younger children (aged 0-6) and older children (aged 7-17) as well as the number of years since migration. We condition on the observed labour force state at time t; the estimation technique is thus applied separately to each column of the transition matrices for the 1980's.[9] Summary statistics for the different sub-populations used in estimation are included in appendix table B1 and the parameter estimates from multinomial logits are in tables B2, for men, and B3, for women.

Second, we compute predicted transition matrices for the 1990's using the parameter estimates from the 1980's and allowing the demographic variables to change to their 1990's levels. From the predicted transition matrices we compute estimates of the steady-state rates of self-employment. Thus, differences between the actual 1980's transition matrices and steady-state rates of self-employment and those that are predicted for the 1990's arise solely because of changes in demographics between the two periods. These results are presented in Table 5.

Table 5 suggests that differences in demographic composition between the two periods do not explain much of the changes in the transition probabilities, nor in the corresponding ergodic distributions and steady-state rates of self-employment that we

---

[9] Because we observe the time t labour force state for individuals probabilistically (described in section 2) we include all individuals with non-zero probabilities of being in a given state at time t in the estimation and use the probabilities as weights.

observe among Canadian men and women between the 1980's and 1990's. In fact, the predicted changes in the steady-state rates of self-employment that arise from changes in the demographics suggest that the steady-state rates would have fallen, though only slightly, if *only* demographics (and nothing else) had changed between the 1980's and 1990's. Adjusting for changes in demographics predicts a decline equal to 4.1 percent of the actual change for men and a decline of 4.6 percent for women.[10]

We focus now on the predicted changes in those transition probabilities which played a key role in the secular changes in self-employment in order to understand the predicted declines in the steady-state rates of self-employment resulting from demographic shifts in the labour force. For men, recall that the pivotal changes in transition rates were those between employment and nonemployment, and note that the decline in employment retention rates we observed in the actual transition matrices is not predicted by changes in demographics. The model actually predicts the opposite; that wage and salary employment stability should have improved. This is due to increases in the age and education of the sample, together with declines in the fraction who are immigrants, all of which should have reduced the transition rate into nonemployment from wage and salary employment, and raised the transition rate from nonemployment into wage and salary employment. Therefore, it is unlikely that changes in demographic composition explain the secular rise in self-employment among

---

[10]We also check this analysis by "back-casting", i.e. estimating the multinomial logit models using the 1990's data and examining the changes in the probability matrices resulting from allowing the demographics to revert back to their 1980's levels. While this analysis attributes a small rise in the steady-state rates of self-employment to changes in demographics, the fraction attributed to changes in demographics, as in the "forecasting" exercise reported in the essay, was minor. These results are available from the authors on request.

Canadian men.

For women, whose demographic changes (e.g. increasing age and education) were similar to men's, one might expect that these changes should explain at least some of their increases in self-employment duration. In fact, the model does predict that durations in, and exit rates from, self-employment would have improved as a result of changes in demographics. The newly formed ergodic distribution also predicts a slight increase, on average, in time spent in self-employment by women in the 1990's. However, this is offset by relatively larger predicted improvements in the wage and salary sector, which results in a small *decline* in the predicted steady-state rate of self-employment. Therefore, as for men, changes in demographic mix do not provide an obvious explanation for women's increasing self-employment rate. A more likely culprit for the secular rise in self-employment among Canadian women is some unobserved, trended factor, such as a general improvements in women's labour market, reflected in rising experience, qualifications, and earnings.[11]

(7) Job Characteristics of the Newly Self-Employed

A number of researchers have suggested that recent increases in self-employment are, in part, attributable to new opportunities opened up by technology that makes self-

[11] While individual-specific measures of actual labour market experience are not available in the SCF, we were able to compute cohort-specific trends in labour force experience among women over this period in these data. Essentially, they follow a linear time trend, indicating that any unexplained time trend might is empirically indistinguishable from an experience effect

employment more feasible than it once was.[12] In contrast, we have argued that, at least for Canadian men, one of the main causes for the rise in self-employment is a long-term decline in the opportunities available to them in the wage and salary sector. If this is the case, then one might expect the quality of men's new self-employment jobs in the 1990's to have deteriorated. Also, given our interpretation that the increase in self-employment among women is likely caused by improving market conditions and human capital, one would not expect to find such a deterioration in new self-employment jobs among women. In this section we present some evidence on changing labour market conditions for men and women in the 1980's versus 1990's, and examine three measures of the "quality" of self-employment opportunities at our disposal to see whether this is indeed the case.

Turning first to relative labour market conditions for men and women, there does indeed seem to be evidence of a "twist" in conditions against men. Between the 1982-1989 and 1990-1995 periods, the employment-to-population ratio of prime age (25-54) men fell, from 86.3 to 83.4 percent. That of women rose, from 63.4 to 69.2 percent. A similar story holds for unemployment rates: looking at business cycle troughs in each of the two periods to abstract from cyclical effects, men's unemployment rose between the two periods, from 9.7 percent in 1983, to 10.7 percent in 1992. Women's unemployment rate peaked at 10 percent in 1984, but only at 9.7 percent in 1993. Thus, aggregate statistics paint a picture of an improving labour market for women, and a deterioriating one for men.

---

[12]See, for example, Farber (1997) which uses US data and Gauthier and Roy (1997) which uses Canadian data.

Tables 6 and 7 compare three measures of job quality for newly self-employed men and women, respectively, to two reference groups: those with longer job tenures in self-employment, and individuals employed in all sectors, for the two periods. The three measures of self-employment quality are the fraction of new self-employment opportunities with paid help, that are full-time, and that pay high wages. According to Table 6, both men and women who are self-employed are increasingly "own-account", in other words without paid help. This change is especially evident among the newly self-employed (those with tenure under a year), and could reflect both changes in technology favourable to that kind of self-employment, or declines in the quality of new self-employment opportunities.[13] More interestingly, however, the increase in own account self-employment over this period is significantly larger among men than women, especially among the newly self-employed: the fraction of newly self-employed men who are own account increased by more than 12 percentage points between the 1980's and 1990's compared to a 5 percentage point increase for newly self-employed women. Trends in own-account self-employment are therefore consistent with the notion that declining market opportunities played a larger role in men's increasing self-employment than in women's.

A second job "quality" measure, usual hours worked per week, also suggests that the nature of self-employment among men is changing.[14] There was a substantial increase in the fraction of men in both sectors working "part-time" (fewer than 30 hours per week) and a

---

[13]These findings support those found in Gauthier and Roy (1997).

[14]In general, of course, high hours of work do not necessarily indicate a good job. They may however help distinguish genuine active businesses from "stopgap" activities.

decline in the fraction working "full-time" (more than 30 hours) between the 1980's and 1990's. At the same time, the fraction of women who were "full-time" remained stable. There was, however, an increase in the fraction of self-employed women working more than 40 hours and a decline in those working less than 20 hours. This is also consistent with the notion that declining market opportunities played a larger role in men's increasing self-employment than in women's.

Table 7 reports average weekly wages for short and longer tenure men and women in both the self-employment and wage and salary sectors for the 1980's and 1990's. For men, both self-employed and wage and salary earners' wages declined between the two periods, but the decline in self-employment earnings, particularly for the more recently self-employed, was substantially larger than the same decline in the wage and salary sector. This is starkly contrasted by a comparatively large increase in weekly wages among self-employed women relative to women earning a wage and salary. Clearly, women's new self-employment opportunities improved over this period, while men's worsened. Women's increased self-employment thus seems a natural response to improved opportunities. Men's self-employment rates increased despite declining qualities of self-employment opportunities, because wage-and-salary opportunities also became less attractive, and because fewer were employed at all.[15]

---

[15]Our evidence for men is also strongly inconsistent with the notion that increased self-employment consists to a substantial degreee of voluntary "contracting out" of jobs to the same individuals who once did the same job in the wage-and-salary sector. In such situations, one might expect wages to *increase* to compensate for the loss in fringe benefits such as dental and pension plans.

(8) Conclusions

The evidence presented here clearly shows that the changes in the underlying labour force flows which have led to a secular increase in Canadian self-employment between the 1980's and 1990's are significantly different for men and women. Declining opportunities in wage and salary employment between the two periods appear to have had a large impact on male self-employment rates and virtually no impact on women's. In contrast, most of the secular rise in self-employment among women is associated with declining exit rates from self-employment, i.e. increased duration. This interpretation is supported by three indicators of the quality of new self-employment opportunities –hours, wages and the presence of employees–, all of which show an improvement for women (relative to wage and salary opportunities) and a deterioration for men.

If, as our results suggest, Canadian men were indeed "pushed" into self-employment in the 1990's by a secular decline in paid employment opportunities, our results may have some interesting implications for economic policy. For example, self-employment assistance plans offered by the Unemployment (now Employment) Insurance system, rather than creating high-quality new opportunities, may simply be adding to the stock of unemployed men already crowding into the self-employment sector in response to declining opportunities in regular paid employment.

Table 1
CANADIAN EMPLOYMENT TRENDS 1982-1995: Men and Women Aged 25-54

Year	MEN			WOMEN			TOTAL		
	Rate	# Self-Employed	# Wage Employed	Rate	# Self-Employed	# Wage Employed	Rate	# Self-Employed	# Wage Employed
1982	8.08	308,270	3,507,690	5.02	134,110	2,534,950	7.32	442,380	6,042,640
1983	8.42	317,720	3,457,320	5.43	150,350	2,619,750	7.70	468,070	6,077,070
1985	8.37	331,710	3,631,420	6.84	207,340	2,822,040	8.35	539,050	6,453,460
1986	9.01	370,340	3,740,510	4.21	136,650	3,109,610	7.40	506,990	6,850,120
1987	8.61	363,340	3,854,460	4.87	164,400	3,208,020	7.47	527,740	7,062,480
1988	8.82	383,491	3,966,676	4.83	174,331	3,437,111	7.53	557,822	7,403,787
1989	8.50	375,724	4,043,312	4.75	177,923	3,567,259	7.27	553,647	7,610,571
1990	8.59	385,585	4,104,205	5.39	211,199	3,705,920	7.64	596,784	7,810,125
1991	9.27	410,246	4,017,396	5.62	221,921	3,728,414	8.16	632,167	7,745,810
1992	9.47	413,552	3,955,136	5.32	210,019	3,734,388	8.11	623,571	7,689,524
1993	9.41	424,916	4,091,814	5.53	221,696	3,786,373	8.21	646,612	7,878,187
1994	10.17	469,614	4,148,796	6.67	273,345	3,826,264	9.32	742,959	7,975,060
1995	9.43	473,413	4,547,520	6.27	278,633	4,168,168	8.63	752,046	8,715,688
% Change 1982-1995	17%	54%	30%	25%	108%	64%	18%	70%	44%

*Notes: 1) Calculations based on series of SCF data files using sample weights
2) Self-employed excludes owner/operators of incorporated businesses
3) Sample is restricted to those employed in non-primary industries

Table 2
Characteristics of Self-Employment Entrants and Leavers

	Entrants				Leavers				Total Population			
	Men		Women		Men		Women		Men		Women	
	80's	90's	80's	90's	80's	90's	80's	90's	80's	90's	80's	90's
Source/Destination (%)												
Wage & Salary	75.2	68.0	61.4	64.0	74.3	70.6	68.4	62.6	84.5	79.9	62.2	67.4
Not Employed	24.8	32.0	38.6	36.0	25.7	29.4	31.6	37.4	15.5	20.1	37.8	32.6
Age (%)												
25-34	41.6	39.6	43.2	39.0	42.5	37.0	41.3	36.8	42.0	38.6	42.5	38.4
35-44	34.9	39.2	37.3	38.7	35.4	36.7	36.4	38.5	33.9	35.8	33.6	36.1
45-54	23.6	21.3	19.6	22.3	22.2	26.3	22.3	24.7	24.1	25.6	23.9	25.5
Education (%)												
8yrs or less	13.8	5.4	9.0	4.1	11.9	6.5	9.2	6.2	12.8	7.6	12.5	7.4
9-10 yrs	17.0	13.5	14.7	9.2	13.7	13.8	14.5	13.1	13.9	12.4	14.0	11.0
11-13 yrs	29.1	38.9	33.0	37.6	28.4	35.6	35.6	38.3	31.7	39.1	36.2	39.4
some post-sec	8.9	9.2	10.1	9.4	10.3	8.6	9.3	8.7	9.2	8.1	8.4	8.5
post-second	14.1	13.9	15.9	20.3	19.0	16.2	15.2	16.4	14.3	14.3	15.9	18.7
university	17.1	19.1	17.4	19.5	21.6	19.3	16.2	17.3	18.1	18.5	13.0	15.0
Marital Status (%)												
single	17.4	24.9	8.4	10.5	17.1	19.3	9.6	11.8	17.0	21.1	12.1	14.1
married	74.5	65.7	81.3	78.6	74.4	73.2	76.7	77.7	77.3	72.4	76.9	74.6
div/wid/sep	8.0	9.5	10.3	10.9	8.5	7.5	13.7	10.5	5.7	6.5	11.0	11.4
Number of Young Children (aged <7) (%)												
none	74.2	75.4	67.0	67.1	71.9	70.8	72.4	66.7	73.3	75.1	73.5	73.4
one	15.8	13.5	19.0	19.2	16.8	17.9	15.6	22.0	16.7	15.1	16.9	16.5
two	8.6	9.8	11.8	10.6	9.5	9.1	10.1	10.0	8.7	8.3	8.3	8.5
three +	1.4	1.2	2.2	3.2	1.9	2.3	1.9	1.3	1.4	1.5	1.3	1.6

Table 2 (Continued)

	Entrants				Leavers				Total Population			
	Men		Women		Men		Women		Men		Women	
	80's	90's	80's	90's	80's	90's	80's	90's	80's	90's	80's	90's
Number of Older Children (aged 7-17) (%)												
none	61.0	71.9	47.2	59.8	60.8	65.9	50.1	54.6	61.3	67.4	55.2	61.4
one	19.8	13.8	23.8	20.5	19.3	17.1	21.6	20.7	18.4	16.3	22.1	20.0
two	13.5	10.5	20.3	15.0	15.3	13.0	20.1	17.6	15.0	12.6	16.8	14.4
three +	5.7	3.9	8.7	4.7	4.7	4.0	8.3	7.1	5.3	3.7	5.9	4.2
Years Since Immigration (%)												
Canadian brn	78.2	81.5	82.6	82.9	83.9	78.3	80.7	83.7	80.2	81.2	81.2	81.4
0-5 yrs	2.9	5.6	2.1	3.8	1.4	5.7	2.0	3.6	2.8	4.4	2.6	4.4
6-10 yrs	2.9	0.4	2.7	1.8	3.1	1.8	4.0	1.2	2.7	1.3	2.7	1.2
11-20 yrs	5.6	4.6	4.8	3.0	5.4	6.3	5.0	3.8	5.8	4.6	5.6	4.8
21-50 yrs	10.5	8.0	7.8	8.5	6.3	7.9	8.3	7.8	8.5	8.4	7.9	8.1

1. All values calculated using SCF data files
2. Some columns may not sum to 1 because of rounding error.

Table 3
Estimated Transition Matrices, Ergodic Distributions and Steady-State Rates of Self-Employment: 1980's v.s. 1990's

Men

Probability Matrix 1980's

E_t	S_t	N_t	
0.9184	0.1487	0.3121	E_{t+1}
0.0116	0.8180	0.0338	S_{t+1}
0.0701	0.0332	0.6541	N_{t+1}

Ergodic Distribution 1980's

E	0.7602
S	0.0783
N	0.1615

Steady-State Self-Employment Rate = 0.0933

Probability Matrix 1990's

E_t	S_t	N_t	
0.9002*	0.1489	0.2608*	E_{t+1}
0.0103	0.8121	0.0357	S_{t+1}
0.0895*	0.0390	0.7035*	N_{t+1}

Ergodic Distribution 1990's

E	0.6983
S	0.0802
N	0.2215

Steady-State Self-Employment Rate = 0.1031

Change in Steady-State Self-Employment Rate 1990's-1980's = 0.0097

Note: * denotes change in element p_{ij} 1990's-1980's is statistically significant at the five percent level

Table 3 Continued
Estimated Transition Matrices, Ergodic Distributions and Steady-State Rates of Self-Employment: 1980's v.s. 1990's

Women

Probability Matrix 1980's

E_t	S_t	N_t	
0.8984	0.2695	0.1708	E_{t+1}
0.0089	0.6482	0.0127	S_{t+1}
0.0927	0.0823	0.8164	N_{t+1}

Ergodic Distribution 1980's

E	0.6372
S	0.0282
N	0.3346

Steady-State Self-Employment Rate = 0.0424

Probability Matrix 1990's

E_t	S_t	N_t	
0.9043*	0.1589*	0.1778*	E_{t+1}
0.0087	0.7790*	0.0146*	S_{t+1}
0.0869*	0.0621*	0.8076*	N_{t+1}

Ergodic Distribution 1990's

E	0.6470
S	0.0458
N	0.3072

Steady-State Self-Employment Rate = 0.0661

Change in Steady-State Self-Employment Rate 1990's-1980's = 0.0237

Note: * denotes change in element p_{ij} 1990's-1980's is statistically significant at the five percent level

Table 4
Decomposition of Change in the Canadian Steady-State (S-S) Rate of Self-Employment

	Men				Women			
Elements Changed	**S-S Rate 1980's**	**S-S Rate 1990's**	**Change S-S Rate**	**Percent Explained**	**S-S Rate 1980's**	**S-S Rate 1990's**	**Change S-S Rate**	**Percent Explained**
All Elements	0.0933	0.1031	0.0097	100	0.0424	0.0661	0.0237	100
S-E Entry and Exit (ES,SE,NS,SN)	0.0933	0.0868	-0.0065	-66	0.0424	0.0690	0.0266	112
S-E Entry (ES,NS)	0.0933	0.0892	-0.0041	-42	0.0424	0.0443	0.0019	8
S-E Exit (SN,SE,SS*)	0.0933	0.0908	-0.0025	-26	0.0424	0.0661	0.0237	100
S-E Duration (SS)	0.0933	0.0906	-0.0027	-28	0.0424	0.0659	0.0235	99
W&S-Non-Employ Transitions (EN,NE)	0.0933	0.1112	0.0179	183	0.0424	0.0408	-0.0016	-7
W&S Non-Employ Duration (EE,NN)	0.0933	0.1118	0.0185	190	0.0424	0.0400	-0.0024	-10
All W&S Non (EN, NE ,EE,NN,ES*,NS*)	0.0933	0.1058	0.0125	128	0.0424	0.0424	0.0000	0

1. For a description of how the values are calculated refer to the text.
2. * denotes elements that took on their 1990 value because of requirement to sum to one.

Table 5
Adjusted Transition Matrices, Ergodic Distributions and Steady-State Rates of Self-Employment: 1980's v.s. 1990's

Men

Probability Matrix 1980's

E_t	S_t	N_t	
0.9184	0.1487	0.3121	E_{t+1}
0.0116	0.8180	0.0338	S_{t+1}
0.0701	0.0332	0.6541	N_{t+1}

Ergodic Distribution 1980's

E	0.7602
S	0.0783
N	0.1615

Steady-State Self-Employment Rate = 0.0933

Predicted Probability Matrix 1990's

E_t	S_t	N_t	
0.9194	0.1512	0.3257	E_{t+1}
0.0116	0.8168	0.0359	S_{t+1}
0.0689	0.0320	0.6384	N_{t+1}

Ergodic Distribution 1990's

E	0.7679
S	0.0787
N	0.1534

Steady-State Self-Employment Rate = 0.0929

Predicted Change in Steady-State Self-Employment Rate 1990's-1980's = -0.0004

Percent of Actual Change Predicted by Allowing X's to Change to 1990's Levels = -4.1%

Table 5 Continued
Adjusted Transition Matrices, Ergodic Distributions and Steady-State Rates of Self-Employment: 1980's v.s. 1990's

Women

Probability Matrix 1980's

E_t	S_t	N_t	
0.8984	0.2695	0.1708	E_{t+1}
0.0089	0.6482	0.0127	S_{t+1}
0.0927	0.0823	0.8164	N_{t+1}

Ergodic Distribution 1980's

E	0.6372
S	0.0282
N	0.3346

Steady-State Self-Employment Rate = 0.0424

Predicted Probability Matrix 1990's

E_t	S_t	N_t	
0.9038	0.2659	0.1905	E_{t+1}
0.0087	0.6565	0.0139	S_{t+1}
0.0875	0.0776	0.7956	N_{t+1}

Ergodic Distribution 1990's

E	0.6721
S	0.0291
N	0.2988

Steady-State Self-Employment Rate = 0.0415

Predicted Change in Steady-State Self-Employment Rate 1990's-1980's = -0.0010

Percent of Actual Change Predicted by Allowing X's to Change to 1990's Levels = -4.2%

Table 6
Employment Characteristics by Job Tenure: Self-Employed v.s. All Employed

	Self-Employed				All Employed			
	1980's		1990's		1980's		1990's	
Year's Self-Employed	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year
MEN								
Self-Employment Type (%)								
Own Account	70.69	57.11	82.81	64.43	----	----	----	----
With Paid Help	29.31	42.89	17.19	35.57	----	----	----	----
Usual Hours Worked Per Week (%)								
Less than 20	9.32	12.14	12.37	15.56	5.23	14.63	8.50	19.69
20-30 hours	7.76	5.33	10.41	7.05	3.99	1.49	6.04	2.04
31-40 hours	39.10	32.47	35.74	31.55	63.52	59.84	57.75	53.63
41+ hours	43.81	50.06	41.48	45.84	27.26	24.04	27.70	24.64

Table 6 Continued								
	Self-Employed				All Employed			
	1980's		1990's		1980's		1990's	
Year's Self-Employed	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year	Less than 1 year	More than 1 year
WOMEN								
Self-Employment Type (%)								
Own Account	79.25	70.16	84.28	74.81	----	----	----	----
With Paid Help	20.75	29.84	15.72	25.19	----	----	----	----
Usual Hours Worked Per Week (%)								
Less than 20	38.28	42.38	35.04	35.73	25.27	36.75	25.53	35.33
20-30 hours	12.07	9.70	15.30	12.83	14.27	8.08	15.18	8.83
31-40 hours	25.20	21.95	23.55	24.91	50.31	47.65	47.62	49.96
41+ hours	24.45	25.98	26.12	26.53	10.16	7.52	11.66	8.88
* Fractions generated using sample weights.								

Table 7
Average Real Weekly Wages by Sector and Job Tenure
1980's v.s. 1990's, Men and Women, Separately (1994 dollars)

	Men			Women		
	1980's	**1990's**	**Percentage Change**	**1980's**	**1990's**	**Percentage Change**
Self-Employed 1-5 years	561.16	511.28	-9%	249.01	306.34	23%
Self-Employed More than 5 years	660.59	644.19	-2%	275.08	325.82	18%
Wage and Salary 1-5 years	728.90	707.04	-3%	434.97	464.67	7%
Wage and Salary More than 5 years	831.84	824.55	-1%	499.41	526.45	5%

* Self-employment wages were calculated as reported net annual income from non-farm self-employment divided by weeks worked in the previous year. Those with negative net self-employment income were included. Similar results were obtained when those with negative net self-employment income were dropped from the sample.
* Wages for wage-and-salary earners were calculated by dividing annual reported income from wages and salarys by weeks worked.

References

Alba-Ramirez, A. (1991) "Self-Employment in the Midst of Unemployment: The Case of Spain and the United States", Working Paper 91-18 Departmento de Economica, Universidad Carlos III de Madrid

Aronson, Robert L. (1991) Self-Employment ILR Press, Ithaca, New York

Becker, Eugene, (1984) "Self-Employed Workers: An Update to 1983", *Monthly Labor Review* 107 pp. 14-18

Bishop, John H., (1987) "American Job Growth: What Explains It?", *Portfolio: International Economic Perspectives* 12

Blanchflower, D. and Richard B. Freeman, "Did the Thatcher reforms change British labour market performance?", in The UK Labour Market: Comparative Aspects and Institutional Developments, edited by R. Barrell, Cambridge University Press, pp. 51-92.

Blanchflower, D. and Oswald, A., (1990) "Self-Employment and the Enterprise Culture", in British Social Attitudes: the 1990 Report, edited by R. Jowell, S. Witherspoon and L. Brook, Gower.

Blanchflower, D. "Self-Employment Trends in OECD Countries: Evidence from Microdata" paper presented at OECD/CERF/CILN International Self-employment Conference, Burlingon Ontario, September 24,1998.

Blau, David (1987) "A Time Series Analysis of Self-Employment in the United States", *Journal of Political Economy*, Vol 95, no.3, pp. 445-467

Borjas, George J., and Stephen G. Bronars (1989) "Consumer Discrimination and Self-Employment", *Journal of Political Economy* 97 pp. 581-605

Crompton, Susan (1993) "The Renaissance of Self-Employment" *Perspectives on Income and Employment* Statistics Canada Cat. 75-001E, pp. 22-32

Devine, Theresa (1993) "The Recent Rise in U.S. Self-Employment", mimeo, The Pennsylvania State University

Devine, Theresa (1990) "The Recent Rise in Female Self-Employment", mimeo, The Pennsylvania State University

Devine, Theresa, and Joyce Mlakar (1993) "Inter-Industry Variation in the Determinants of Self-Employment", mimeo, The Pennsylvania State University

Evans, David S., and Linda S. Leighton (1989) "Some Empirical Aspects of Entrepreneurship", *American Economic Review*, vol. 79, pp. 519-535

Fairlie, Robert W. and Bruce Meyer (1998) "Trends in Self-Employment Among White and Black Men: 1910-1990", Presented at the International Conference on Self-Employment, sponsored by CILN, September 24-26, 1998, Burlington Ontario.

Farber, Henry S., (1997) "Alternative Employment Arrangements as a Response to Job Loss", Princeton University, Working Paper #391.

Gauthier, James, an Richard Roy (1997) "Diverging Trends in Self-Employment in Canada", Human Resources Development Canada, Working Paper R-97-13E

Lin, Z., J. Yates and G. Picot (1998a) "Rising Self-Employment in the Midst of High Unemployment: An Empirical Analysis of Recent Developments in Canada", Statistics Canada Working Paper.

Lin, Z., J. Yates and G. Picot (1998b) "The Entry and Exit Dynamics of Self-employment in Canada", Statistics Canada Working Paper.

Meyer, B.D. (1990) "Why Are There So Few Black Entrepreneurs?", NBER Working Paper No. 3537

Moore, C.S., and Richard E. Mueller (1998) "The Transition in to Self-Employment in Canada: The Importance of Involuntary Separation and Unemployment Duration", Presented at the International Conference on Self-Employment, sponsored by CILN, September 24-26, Burlington.

Rees, Hedley, and Shah, Anup, (1986) "An Emperical Analysis of Self-Employment in the U.K.", *Journal of Applied Econometrics* 1 pp. 95-108

Schuetze, H.J. (1998) "Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-US Comparison", Canadian International Labour Network (CILN) Working Paper #13.

Appendix A: Defining the "main job" in the pre-survey year

The two-year panels overlap. So, for any given year (say 1994) we observe information for the entire year for one sample of individuals (the 1995 survey) and for the week prior to the survey for a second sample (the 1994 survey). Because both samples are random draws from the entire population the expected average characteristics of the samples should be the same. Therefore, we rank individuals based on the fraction of income from wages and salaries (α) in the first sample. Based on this ranking we partition the sample such that the probability of self-employment in any week of that survey year is equal to the rate in the week prior to the survey in April of that same year. In other words, we select an $\alpha=\alpha^*$ such that: $\frac{\sum_{\alpha=0}^{\alpha=\alpha^*} \beta w}{\sum_{\alpha=0} \beta w}$ equals the rate of self-employment in the second sample (where w is the sample weight, and β is the total number of weeks worked divided by 52). We then assign individuals' time in employment as "self-employed" if their fraction of income from wage and salaries is less than the cutoff value (replace $\alpha=0$ if $\alpha<\alpha^*$) and as "wage-and-salary" time if the fraction is above the cutoff ($\alpha=1$ if $\alpha>\alpha^*$). Then the probability of being self-employed in any week for an individual is $\beta(1-\alpha)$. We lose the first year of data because of the matching process.

Chapter 4

PROFILES OF TAX NONCOMPLIANCE AMONG THE SELF-EMPLOYED IN CANADA: 1969-1992

(1) Introduction

Because of the absence of a third party reporting income, the self-employed are often able to conceal income from the tax authorities. Public finance economists argue that the presence of such "underground" activity is economically distortionary and leads to an inequitable distribution of the tax burden (Auerbach and Feldstein 1998 and Slemrod 1992). The distribution of the tax burden is shifted away from individuals who have the ability and choose not to comply with the tax rules toward those who have little choice but to comply.

These distortions have become increasingly important in light of the fact that self-employment as a share of overall labour market activity has grown in Canada. Between 1982 and 1995 the number of self-employed Canadians between the ages of 25 and 54 increased by 70 percent, compared to a 45 percent increase in paid employment.[1] In the first chapter of this thesis, I presented evidence that suggests that one of the main reasons for the growth in self-employment in Canada between 1983-1994 is the concurrent rise in income taxes which made self-employment more attractive to entrepreneurs. To the extent that these new

[1]Kuhn and Schuetze (1999)

entrepreneurs are attracted to self-employment by the relative ease of concealing income, this result suggests that the *incidence* of tax noncompliance has risen in Canada. However, very little is known about the *extent* to which the self-employed conceal income in Canada. Even less is known about the factors which influence the degree of income concealing among the self-employed.

Previous research has been limited to "aggregate" estimates of income tax noncompliance that do not give much of an indication of the determinants of such activity (Mirus and Smith 1996) and , in some cases, do not allow one to focus on the self-employed (Spiro 1997, Mirus, Smith and Karoleff 1994, and Hill and Kabir 1996). In addition, those studies which do provide estimates over time are subject to large measurement error that make time trend analysis difficult to interpret.

In the current essay, the limitations of previous research are overcome by utilizing an expenditure based approach developed by Pissarides and Weber (1989) (referred to as P&W throughout this essay) to estimate the degree of noncompliance by the self-employed using a series of Canadian Family Expenditure Surveys for six years between 1969 and 1992, which allow estimates across various dimensions. Essentially, this approach uses estimates of the relationship between food consumption and income obtained from a sample of wage and salary earners, who are assumed to report income correctly, in order to impute "true" income levels for the self-employed. When compared to reported incomes these imputed incomes provide an estimate of the degree of noncompliance among the self-employed. Unlike other measures of noncompliance, the estimates obtained using this procedure take

into account both under-reporting of gross income and overstatement of expenses and deductions (see section 3 for a discussion).

I obtain estimates of the degree of noncompliance by the self-employed across years, demographic characteristics and occupations. By disaggregating across these dimensions I hope to achieve two goals. The first goal is to identify the determinants of self-employed income tax noncompliance in Canada and the second is to provide guidance to tax enforcement policy makers in designing effective tax policy schemes. The characteristics of noncompliance among the self-employed might provide clues to the reasons for this activity. For instance, one might expect that the amount of income concealed by entrepreneurs varies with changes in the tax environment and, therefore, also may vary with time. Put simply, an increase in the marginal benefit of concealing self-employment income (the effective marginal tax rate) is likely to lead to an increase in noncompliance. Further, one's opportunity to under-report income from self-employment is likely to vary by occupation. For example, individuals who provide products or services in occupations which frequently involve cash-related transactions have greater opportunity to under-report income.

Tax enforcement agents often face the difficult task of enforcing tax laws because of an imperfect ability to monitor tax compliance. One tool used by tax authorities to enforce compliance is a taxation audit scheme. By identifying the observed patterns of tax noncompliance across characteristics which are commonly included on income tax returns (occupation, age of tax-filer, and number of household adults self-employed) this essay will provide guidelines to aid the tax authority in enforcing income tax compliance. Audits can

be directed towards those who are most likely to be in noncompliance with the income tax code. The results of such audits could then be used to identify the source of noncompliance – overstated expenses or understated income – which would aid in designing tax policy focussed on reducing this activity.

My main findings are as follows. First, the estimates obtained suggest that the mean self-employed Canadian concealed between 11 and 23 percent of their incomes on average between 1969 and 1992. Second, with regard to the determinants of noncompliance among the self-employed, there appears to be no clear time trend in the degree of income concealing by the self-employed over this period. I do, however, find evidence that the degree of noncompliance varies significantly by the occupation of the entrepreneur. This suggests that one's "opportunity" to under-report income may be a determinant of the degree of noncompliance among entrepreneurs. Finally, one's level of education does not appear to be a determining factor in the decision to conceal income among the self-employed.

Third, with regard to the insights provided by looking across characteristics observed by the tax authority, there is a great deal of variation in the degree of noncompliance among the self-employed across these dimensions. Households with significant amounts of self-employment income which are headed by younger males tend to conceal income to a greater extent than those headed by older males. Further, noncompliance among self-employed households is highest among those in which the occupation of the self-employed individual(s) is construction, followed closely by those in all service occupations. Interestingly, these sectors of the labour market are commonly pointed out anecdotally as

occupations with a large degree of "underground" market activity. These occupations often involve cash related transactions, which allow income under-reporting with relative ease. On the other hand, households with members engaged in self-employment activities in the product fabrication industries– an occupation likely involving very few cash transactions– had the lowest estimates of noncompliance. Finally, there appears to be no difference in the degree of noncompliance among households in which only the husband is self-employed compared to those in which only the wife is self-employed. However, estimates of the extent of income concealing are lower in households in which both members earn self-employment income. As suggested above, these results could be used by policy makers involved in developing taxation enforcement policies, which may include tax audit schemes. A tax audit scheme targeting groups which have been found to conceal income the most (for example, self-employed households headed by younger males or those in the construction and service occupations) is likely to be an effective tool in reducing noncompliance. The implications of these results are discussed in greater detail below.

The remainder of this essay is organized as follows. Section 2 briefly reviews the existing literature on income tax compliance– focusing on the various techniques used to derive estimates of income concealing among the self-employed. Section 3 outlines the methodology used to derive estimates of tax noncompliance by the self-employed in Canada. In Section 4 the data used in estimation, the sample restrictions and the empirical approach are described. Section 5 outlines the results and Section 6 concludes.

(2) Previous Research

This section briefly summarizes previous attempts to provide estimates of the extent to which income is concealed, with special attention given to the few studies which provide estimates of noncompliance for the self-employed. For a more thorough review of the literature see, for example, Erard (1997) on Canada and Smith (1986) on the UK. The various sources of information on tax evasion used by researchers to this point are derived from i) national accounts data, ii) tax audit data, iii) tax amnesty data, and iv) survey data.

A common aggregate measure of underground activities employed by researchers is the "GDP gap". The GDP gap measures unrecorded GDP and is calculated as "the difference between total market-based income (whether from legal or illegal sources) in the domestic economy and recorded GDP".[2] Estimates of unrecorded GDP for Canada range, depending upon the year and method used to obtain an estimate, from 2.5% of GDP in the 1970's (Pouftis 1993) to 24% of GDP in 1990 (Mirus, Smith and Karoleff 1994). In addition to the wide variation in estimates, there is also no consensus on the trends in unrecorded GDP in Canada. At least two authors (Ethier 1985 and Pouftis 1993) find evidence that the fraction of GDP unrecorded rose from the 1960's to the early 1990's but fell in the early 1980's. At the same time, estimates provided by Mirus and Smith (1981) and Mirus, Smith and Karoleff (1994) suggest that the percent of GDP that was unrecorded did not change much between 1976 and 1990.

There are several problems with this approach as it relates to the focus of the current

---

[2]Erard (1997) page 2.

essay. Because the GDP gap is measured at the aggregate level one cannot get a breakdown of the estimate of underground activity for individual groups in the economy. For example, no estimate of underground activity by the self-employed can be obtained using this approach. Further, as Erard (1997) points out, measuring unrecorded GDP is not necessarily equivalent to measuring tax non-compliance– which is the focus of this essay. For instance, not all income included in unrecorded GDP is taxable. Some portion of the GDP gap may be attributable to individuals earning income below that which is taxable. Also, the GDP gap captures only the under-reporting of gross income and leaves out overstated tax deductions and credits. Finally, as is suggested by the large variation in estimates of the GDP gap, the potential for measurement error inherent in this approach does not lend itself well to investigations of time trends in under-reporting.

A second source of information on underground activity that is available, though only for the US, is data collected through tax audits. Through the Taxpayer Compliance Measurement Program (TCMP) the US Internal Revenue Service publishes estimates of the difference between federal tax liabilities reported and assessed liabilities owed by individuals and small corporations. Under the TCMP a stratified random sample of individual and corporate income tax returns are subjected to intensive audits. Estimates based on these audits are produced and published for various groups. Most notably, a breakdown by industry for sole proprietors provided by the US General Accounting Office (1990) suggests that the degree of underground activity by sole proprietors varies greatly by industry. Among the industries in which the percentage of noncompliance by sole proprietors was the greatest

include: retail sales (fixed location) 39%, transportation 36%, retail sales (no fixed location) 31% and production (including construction) 24%. The lowest estimates of noncompliance were recorded among those in wholesale trade 19%, and finance/insurance/real estate 16%. These compare to an estimate by wage and salary employees of only about 1%. Although these estimates are based on data from the US, given the similarities between Canada and the US in terms of their labour market characteristics and overall tax structure, they are relevant to the current study.

While audit data provides estimates of both under-reported income and overstated tax deductions and credits, these estimates are not very reliable. Estimates based on audits rely heavily on the auditors' ability to identify under-reporting and interpretation of the tax laws as they pertain to tax deductions and credits. It is highly unlikely that income tax auditors are able to identify all income that is concealed from the tax authorities. Also, no audit data or estimates are available for Canada.

Tax amnesty data also provide a direct source of information on tax noncompliance. Data on self-reported evasion by those choosing to participate in the amnesty program have been used to obtain estimates of noncompliance (see Andreoni, Erard and Feinstein 1998 for a discussion). The biggest problem with such data is that of sample selection bias. Those who participate in such programs are not likely to be representative of the entire population. Further, they might provide information only on sources that were likely to be traced later.

Finally, a unique approach which uses household expenditure micro-survey-data has been developed by P&W (1989). Essentially, this approach uses what is thought to be an

undistorted estimate of the marginal propensity to consume food in order to impute true income levels for the self-employed. Based on the assumption that the wage and salary employed report income correctly, an estimate of the "true" relationship between food expenditure and income may be obtained by restricting the sample to the wage and salary employed. Provided that the self- and wage and salary employed have the same food consumption patterns, inversion of the expenditure function allows one to impute "true" income levels for the self-employed from their reported food expenditure.

This approach has been applied recently by Mirus and Smith (1996) using the 1990 Canadian Survey of Family Expenditures. They estimate that in 1990 the self-employed concealed 12.5% of their incomes, on average, or $6 billion, in aggregate. While informative, the fact that their estimate is aggregated across the economy and based on a single year is limiting. As noted earlier, disaggregating is likely to reveal clues about the determinants of noncompliance among the self-employed and provide guidelines in designing effective tax policies. In addition, the results presented below suggest that estimates based on a single year of data without controlling for hours worked can be misleading[3].

(3) Methodology

Following closely the strategy developed by P&W for a single year of data, the approach taken here is to estimate the extent of noncompliance among the self-employed

---

[3]The results without pooling across years tend to be very noisy. In addition, the effect of the business cycle also resulted in a great deal of variance in the estimates from year to year. This was remedied by restricting the sample to full-year full-time workers (a restriction Mirus and Smith did not make).

based on household expenditure patterns pooling across six years of data covering the period 1969 to 1992. Estimation is made possible by a few key assumptions. It is assumed that: i) all respondents report expenditures on food correctly in the FAMEX surveys, ii) all earners report the same income in the survey as they do on their tax return[4], iii) wage and salary earners report income correctly while the self-employed may not, and iv) the relationship between family expenditure and after-tax income is the same for all families in the survey– whether self- or wage and salary employed. Based on these assumptions a two-stage estimation technique can be applied to obtain estimates of the extent of noncompliance. In the first stage a prediction of the relationship between household food consumption and after-tax income, controlling for household characteristics and year effects, is obtained for all years. In the second stage the expenditure function is inverted and estimates of "true" income for the self-employed are derived for each year, demographic and occupational group. When compared to reported income these imputed incomes provide estimates of non-compliance– including unreported income *and* overstated deductions– for the self-employed. Because this method yields a measure of *any* difference between "true" and reported *after-tax* income, both unreported income and overstated deductions will be represented in the estimate of noncompliance[5].

---

[4]This assumption is not critical to the analysis. If, for instance, individuals are more willing to report income correctly in a survey because of assurances of confidentiality, these estimates can be interpreted as a lower bound estimate of income under-reporting on tax returns. Further, in looking across time/characteristics the pattern is not likely to be biased because of violation of this assumption.

[5]This assumes that tax expenses and deductions for the self-employed are allowed by the tax authority to account for expenses etc. incurred for business purposes only. Otherwise, a real differential in the tax treatment of the two groups would result and bias the estimates. What is important to this analysis is the intent of these tax laws and not the interpretation of such laws by the self-employed.

The expenditure equation can be written:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^P + \varepsilon_i \qquad (1)$$

where $\ln Xf_i$ is the natural log of food expenditure for household i, Z_i is a set of household characteristics and year indicators, and Y_i^P is permanent after-tax income.

However, as P&W suggest, measured income will differ from true permanent income for a couple of reasons. First, following the discussion above, true income (Y_i^T) will differ from reported income (Y_i^R) for the self-employed because of noncompliance– but not for wage and salary employed. This relationship can be written:

$$Y_i^T = k_i Y_i^R \qquad (2)$$

by assumption $k_i=1$ for wage and salary earners and $k_i>1$ for the self-employed. A larger k_i indicates a greater degree of noncompliance. Next, true income (Y_i^T) will differ from "permanent"[6] income (Y_i^P) because of transitory shocks. P&W write the correspondence between true income and permanent income as:

$$Y_i^T = p_i Y_i^P \qquad (3)$$

where p_i is a random variable measuring the effect of aggregate events on current income. In any given year, the mean of p_i is assumed to be the same for both the self- and wage and

---

[6]In fact, what one wants here is that measure of income which influences a household's food consumption decision. I follow P&W by labeling this measure "permanent income" without requiring that it conform to the permanent income hypothesis. As P&W suggest, the distinction is made between measured and permanent income to adjust for the likely possibility that the variance of measured income for the self-employed differs from that of the wage employed for a given level of permanent income.

salary employed but its variance is expected to be greater for the self-employed.

Assuming that p_i and k_i are log normal, one can write them as deviations from their means:

$$\ln p_i = \mu_p + u_i \qquad (4)$$

$$\ln k_i = \mu_k + v_i \qquad (5)$$

Where u_i and v_i have zero means and constant variances σ_u^2 and σ_v^2.

Substituting equations (2)-(5) into (1) yields:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^R - \beta(\mu_p - \mu_k) - \beta(u_i - v_i) + e_i \qquad (6)$$

For wage and salary earners k=1, therefore, μ_k=0. For the self-employed k>1. Further, P&W show that, because the variance of p_i is greater for the self-employed than for wage earners, this implies that μ_p (the mean of the log of p_i) is smaller for the self-employed than for wage and salary earners. Thus, equation (6) will differ between the self-employed and wage and salary earners because μ_p-μ_k differs between the two groups. By introducing a self-employment indicator SE_i (=1 if person i is self-employed, 0 otherwise) it can be shown that equation (6) can be altered to isolate μ_p-μ_k in the following estimation equation:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^R + \delta SE_i + e_i \qquad (7)$$

Assuming that α and β are the same for all families, the intercepts for the two groups will differ because μ_p-μ_k will differ. P&W conclude that $\delta = \beta[\mu_k + 1/2(\sigma^2_{uSE} - \sigma^2_{uEE})]$ where

subscripts SE and EE refer to the self- and wage-employed respectively. Therefore, from the estimate of δ one cannot isolate μ_k, because of the differences in the variance of log p_i between the self- and wage-employed. Assuming log normality, the average value of k_i (denoted($\bar{k}$)), which is what one would like to estimate, can be written:

$$\ln \bar{k} = \mu_k + \frac{1}{2}\sigma^2_{vSE} \qquad (8)$$

By substituting the interpretation of δ given above into equation (8), ln $\bar{k}$ can be rewritten:

$$\ln \bar{k} = \frac{\delta}{\beta} + \frac{1}{2}(\sigma^2_{vSE} - \sigma^2_{uSE} + \sigma^2_{uEE}) \qquad (9)$$

However, because p_i and k_i are not observed, it is not possible to obtain estimates of the variances related to aggregate shocks and noncompliance, separately. Hence, no single estimate of the degree of noncompliance can be calculated. Instead, P&W estimate an income equation in order to obtain an estimate of the income variance of errors which is compiled of three types of errors: unexplained variations in permanent income, and the two types of errors arising from deviations from permanent income discussed above (u_i and v_i). They show that, under a reasonable set of assumptions, this "aggregate" measure of the income variance will allow estimates of upper and lower bounds on the extent of noncompliance to be computed. The income equation can be written as follows:

$$\ln Y_i^R = \delta Z_i + \pi X_i + \lambda_i \qquad (10)$$

where Z_i has the same interpretation as in equation (1) and X_i is a set of instruments. Under

the assumption that unexplained variation in permanent income have the same variances for the self- and wage-employed (they have the same σ_u^2), P&W show that the difference between the residual variances of reported income from equation (10) for the two groups can be written:

$$\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE} = \sigma^2_{uSE} + \sigma^2_{vSE} - 2\,\mathrm{cov}(uv)_{SE} - \sigma^2_{uEE} \qquad (11)$$

Assuming that noncompliance and aggregate events are uncorrelated (cov(uv)=0), negative one times the right hand side of equation (11) is approximately equal to the bracketed portion of equation (9). Since σ^2_{vSE} and σ^2_{uSE} are inversely related, a lower bound estimate of noncompliance can be obtained from the residual variances in (10) when σ^2_{vSE} takes on its lowest value and an upper bound when σ^2_{uSE} takes its lowest value. The lowest possible value for σ^2_{vSE} is zero. Therefore, substituting this value in (11) it can be shown that the upper bound $\bar{k}_u$ can be estimated by substituting estimated values of the parameters in:

$$\ln \bar{k}_u = \frac{\delta}{\beta} + \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad (12)$$

Because it is assumed that the variance of self-employed earnings is at least as large as those of wage and salary earners, the lowest possible value for σ^2_{uSE} is when $\sigma^2_{uSE}=\sigma^2_{uEE}$. Substituting this value again in (11) and replacing parameters with estimates yields an estimate of the lower bound ($\bar{k}_l$) which satisfies:

$$\ln \bar{k}_l = \frac{\delta}{\beta} - \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad (13)$$

(4) Data and Estimation

The data used in this analysis are taken from a series of Canadian Family Expenditure Surveys (FAMEX) from 1969 to 1992. Because the FAMEX surveys are conducted somewhat sporadically over time and some key variables are absent in some years of data, only the following six years of data could be used in estimation: 1969, 74, 84, 86, 90, and 92. The FAMEX surveys are conducted in the first few months of the calendar year and provide retrospective information for the previous year on household's income, demographics and detailed expenditures. The unit of analysis used throughout this study is the household spending unit (referred to as a "family), which is assumed to act as a single decision maker for decisions about income reporting and expenditures[7]. In all years, the sample is restricted to private households in 15 metropolitan areas in Canada. The sample studied is further restricted to i) households containing two adults where the spouse of the head is the wife, ii) the household had positive after-tax income, iii) the head was employed full-time full-year (49 weeks or more) in the previous year-- either in the wage and salary or self-employment sector, iv) the head was aged 25 to 64, and v) the head was employed in a non-primary occupation. This was done to eliminate households from the sample deriving much of their self-employment income from farming. Farm households are likely to have much different expenditure patterns on food than those in other occupations. These restrictions left a total of 8,463 households, of which 8.1%, or 682 households were "self-employed" according to

[7] The selection of a sample of family units as opposed to single adult households was made because there are insufficient observations to study the latter.

the definition which follows.

In a regression setting using the data files pooled across the years, equation (7) was estimated for a number of different specifications. The dependent variable is the natural log of expenditure on food (converted to real 1986 dollars using the Consumer Price Index-- All Items), whether purchased for home or in restaurants.[8] The independent variables include the natural log of net after-tax family income[9] at real 1986 values, an indicator variable for self-employment status (defined below), the self-employment indicator interacted with year or household characteristic dummies as indicated below, a series of controls for household characteristics, and year indicator variables. The implicit restriction made in this pooled setting is that the marginal propensity to consume food is the same across time.[10] Controls for household characteristics (Z_i) include: the age of the head, the head's age squared, the number of children present in the household-- separately for young children (aged 6 or less), "middle aged" children (7-12) and older children (13-17), the wife's level of education, indicator variables for region, the number of rooms in the household's dwelling, the value of the house resided in if owned, a dummy variable indicating households that rent and the year that the household was surveyed.

---

[8]I experimented with a number of different expenditure items. These results yielded similar patterns of income under-reporting.

[9]Income after taxes is calculated as the difference between "income before taxes" (which includes wages and salaries, income from self-employment, investment income, government transfers and retirement pensions and other money incomes) and "personal taxes" (individuals are asked to report personal income tax paid in the survey year).

[10]In fact, an informal test of this assumption confirms that it cannot be rejected. In regressions for each year separately the marginal propensities to consume food were not statistically significantly different from one another at standard levels.

The self-employment indicator takes on the value of one for those households classified as self-employed and zero otherwise. Individuals are not asked to self-report self-employment status in the FAMEX surveys. Instead, in this study self-employment status is determined by the fraction of total household income from all sources coming from self-employment. Income from self-employment includes: net income from non-farm self-employment, net income from farm self-employment, and gross income from roomers and boarders not related to the head. Self-employed households are defined as households deriving 30 percent or more of total income from self-employment.[11] With this definition I attempt to capture households with sufficient income from self-employment to allow for the opportunity to conceal income from the tax authorities.

Table 1 presents sample characteristics for households classified as self-employed and wage and salary, separately. The self-employed in this sample tended to be older, had more children and were less likely to rent, on average, than those in wage and salary employment. In addition, the self-employed were more likely to reside in British Columbia and the Prairies and less likely to reside in the Atlantic Provinces and Quebec than wage and salary workers. Incomes were higher among the self-employed as compared to wage and salary earners. Somewhat surprisingly, however, the standard deviations of the natural log of incomes were no higher for the self-employed relative to wage and salary earners, as was expected. Finally, average log expenditures on food by the self-employed were higher than

---

[11]Several other alternatives to the 30% cutoff were tried. The results of these estimates were not significantly different. These results are available from the author.

the same expenditures by the wage and salary employed.

Because permanent after-tax income is measured with error and the error term is assumed to be heteroscedastic, equation (7) is estimated by generalized two-stage least squares.[12] Reported income is treated as endogenous and estimated as equation (10) above, where the instruments, X_i, are a series of indicators for the head's education level, indicators for the wife's work intensity and the self-employment indicator interacted with the number of children, the wife's work intensity and the number of rooms in the dwelling.[13] The results of this particular specification are presented because the test of the over-identifying restrictions was not rejected. However, several alternative specifications were tried and resulted in similar results.

(5) Results

In order to calculate the upper and lower bounds on the estimates of noncompliance in equations (12) and (13) one requires estimates of the coefficients on $\ln Y_i^R$ and SE_i and

---

[12]It is likely that the self-employment indicator, as defined, is also subject to some measurement error. Therefore, because some self-employment income is reported incorrectly by assumption, so is the fraction of income from self-employment. I experimented with a model treating the self-employment indicator as endogenous, using several different instruments (including many very similar to those used by P&W) which led to highly unstable results. Tests of the overidentifying restrictions suggest that these results may be caused by poor instruments. Hence, self-employment is treated as exogenous. The likely result of not instrumenting for the self-employment indicator is that the estimates of non-compliance are biased downward. Households which earn a large fraction of income from self-employment but conceal much of this income will be mis-classified as wage earners. This will likely lead to a greater than expected estimate of food consumption, controlling for income, among wage earners and a lower than expected estimate for the self-employed (i.e. a downward bias of the coefficient on self-employment status).

[13]I experimented with several different specifications and estimated using various combinations of instruments. Unlike P&W, I found that measures such as the number of rooms and the value of the house belonged in both the food expenditure and income equations.

of the residual variance of income for the self- and wage-employed, separately. These coefficients, the residual variances obtained from the first stage regression, and the resulting estimates of noncompliance by year are presented in table 2. Estimates of the upper and lower bounds on $\bar{k}$, the numbers, on average, by which reported self-employment incomes must be multiplied by to derive true income, are calculated for each year of data from the coefficients on the self-employment status indicator interacted with year indicators and estimates of the residual variance of income for the two groups.[14] The estimates of the coefficient on log reported income and the residual variances from income for the two groups are restricted to be the same over the period.[15] (For the full set of coefficients for an end-stage regression see appendix table 1).

The marginal propensity to consume food, $\hat{\beta}$, is positive and significantly different from zero at the five percent level and is estimated to be about 0.39. The coefficient on the self-employment dummy, $\hat{\delta}$, is positive and significant in the base year (1969). While individually each of the interactions between the self-employment indicator and year interactions is statistically insignificant, these coefficients are jointly significant at standard levels. In all years, the implied effect of self-employment on food expenditure is positive. This suggests that, controlling for income and household characteristics, the self-employed consume more food than the wage and salary employed. There is no discernible time trend

---

[14]The variance of permanent after-tax income is assumed to take on a different value for self-employed households than for households engaged in wage and salary employment.

[15]There are few convincing reasons that one would expect these to vary over time. Though not provided here, estimates obtained through generalized 2SLS for each year are consistent with this expectation.

in the estimated degree of income concealing by the self-employed over this period. In 1969 the upper and lower bounds on $\bar{k}$ fall in a somewhat narrow range between 1.14 to 1.26. These imply that self-employed households concealed, on average, between 14 and 26 percent of household income from the tax authorities in 1969. These estimates drop slightly after 1969 and, if anything, trend upward after 1974.

Although there appears to be no pattern of noncompliance over time there is substantial variation across occupations. In order to test for variation in noncompliance across occupations, estimation proceeds as before, with the exception that the coefficient on the self-employment indicator is now crossed with indicator variables for the different occupation groupings and is no longer interacted with the year indicators. The occupation groupings are: managerial and administration, professional and technical, sales, services, product fabricating, construction, and other occupations. Due to changes in the coding of the occupation variable, only the years 1984, 86, 90 and 92, which maintained the same coding, were included in the regression. Table 3 presents estimates of the upper and lower bounds on $\bar{k}$, by occupation for these years. Again, a joint test of the industry dummies crossed with self-employment status suggests that these coefficients are jointly significant at standard levels and significantly different from each other. The estimated bounds on $\bar{k}$ vary from a low of 0.96-1.05 in product fabricating occupations– which suggests that entrepreneurs in these occupations report their incomes correctly– to a high of 1.39-1.53 in construction occupations. The high degree of noncompliance among the self-employed in the construction occupations is followed closely by those in services with estimated bounds

of 1.38-1.52.

Interestingly, while not directly comparable, the results here are similar to those found in the US, presented in section 2, which suggest that construction is among the industries in which the self-employed conceal income the most, and product fabricating is among those with the least amount of income concealing.

One exception in which the results in section 2 differ from those in table 3 arises when the retail sales industries in the US and the sales occupations in Canada are compared. There appears to be a great deal of noncompliance by entrepreneurs in the US in the retail sales (fixed location) industry while the estimates in table 2 suggest that this is not the case for the sales occupations in Canada. These differences arise, most likely, because the industry category for the US is much more narrowly defined than the occupation category in Canada. The occupation category in Canada includes occupations such as wholesale trade– an occupation which has a much lower estimated degree of noncompliance among the self-employed in the US.

These results are consistent with the hypothesis that an entrepreneur's "opportunity" to conceal income as determined by industry/occupation characteristics is correlated with the actual degree of noncompliance. For instance, those involved in the construction trades typically provide services through informal arrangements frequently involving cash transactions. This type of self-employment undoubtably provides greater opportunity for under-reporting. On the other hand, these characteristics do not typify production industries. As the estimates in table 3 show, entrepreneurs in construction tend to conceal a greater

fraction of their income than those in product fabrication. This finding has important implications on the efficiency of the economy. Assuming that there is free entry into and out-of occupations, individuals will be attracted, ceteris paribus, to those occupations which provide the greatest after-tax wages (i.e. those which provide the opportunity to conceal income). Even though pre-tax wages will adjust to equalize after-tax wages, too many resources, in terms of efficiency, will be allocated to occupations which provide the greatest opportunities for noncompliance. This result also suggests an obvious strategy for the development of an effective tax audit system. One which targets the self-employed in construction and service occupations is likely to be more effective.

It is unclear as to how educational attainment, controlling for household income, would affect the degree of noncompliance by the self-employed. It might be the case that those who are more highly educated are more familiar with the tax laws and are, therefore, more able to avoid paying taxes. On the other hand, those with lower levels of education might be more likely to enter industries, such as construction, which provide the opportunity to conceal income. Table 4 presents estimates of noncompliance by the self-employed by the head's level of education applying a similar estimation technique to that used above. One difference is that the self-employment indicator is no longer allowed to vary by year or occupation. Instead, the self-employment indicator is interacted with education category indicators. Also, all six years of data are included in the sample again. There appears to be no clear pattern of noncompliance across the education categories. The degree of noncompliance rises with the level of education beginning with relatively low estimates for

households where the head has less than 9 years of education up to those where the head has some post-secondary education. Households where the head has a post-secondary degree tend to conceal the most, while those where the head has a university degree conceal approximately the same fraction of income as those whose head has a secondary degree.

Finally, to provide guidance to policy makers involved in developing tax policy I provide estimates of noncompliance by some other characteristics which are typically observed on individuals' income tax returns. This essay has previously explored the role of self-employment occupation, on which the tax authority typically has information, in determining tax noncompliance. In addition to occupation, the Canadian federal tax authority also observes the age of the tax filer and information about the tax filer's spouse– including whether or not the spouse is self-employed. Table 5 provides estimates of noncompliance for the self-employed by the head's age category and which of the adult members in the household are self-employed. The estimates provided in table 5 by age category indicate very clearly that the degree of noncompliance by entrepreneurs decreases with age. The estimates increase monotonically with age– from a high of 1.16-1.29 in households headed by an individual aged 25-34, to a low of essentially full compliance (0.95-1.06) among those whose head is aged 55-64. This is consistent with estimates for the US using TCMP data which indicates that the degree and incidence of noncompliance is lower among households headed by older individuals (US General Accounting Office 1990).

I am unaware of any study which looks at non-compliance by the number and sex of the self-employed members of a household. Do households in which there are two members

earning self-employment income conceal a greater fraction of income than households with only a single self-employed member? If there is only one adult member self-employed, does it matter if it is the husband or wife that is self-employed? The estimates provided in the lower half of table 5 suggest that households in which both the head and spouse earn self-employment income conceal a lower fraction of income than households in which only one adult member is self-employed. One possible explanation for this result is the likelihood that increases in the number of self-employed individuals in a household will yield marginally fewer opportunities to conceal income. For instance, if one individual in a household is self-employed and claiming a deduction for a room used as an office, a second self-employed household member will not be able to also claim the same deduction. In response to the second question posed above, it appears that it does not matter whether it is the husband or the wife that is self-employed when only one adult earns self-employment income. Estimates of non-compliance in households where only the head is self-employed are not significantly different from those where only the wife is self-employed. The estimated range is 13-26 percent among self-employed heads and 15-28 percent among self-employed wives.

(6) Conclusions

The rise in self-employment experienced recently in Canada has brought to the forefront some of the problems associated with this form of employment. In particular, the fact that the self-employed are often able to conceal income from tax authorities makes understanding the determinants of non-compliance by this sector of the labour market

increasingly important to those studying labour markets and to tax policy makers. To this point, however, estimates of the degree of income tax noncompliance by the self-employed have revealed little about the factors which influence the decision to conceal income. This essay represents an attempt to overcome this shortcoming. Using family expenditure data for Canada between 1969 and 1992 and employing an approach developed by Pissarides and Weber (1989), I obtain estimates of the degree of noncompliance across various dimensions, focusing primarily on individual and household characteristics commonly reported on income tax returns.

Estimates of the extent of noncompliance averaged over the six years yield an average lower bound estimate of 11 percent of income and an average upper bound estimate of 23 percent. The estimate of noncompliance for the 1990 calendar year found here to be between 12 and 24 percent is somewhat higher than the 10 to 15 percent found in Mirus and Smith (1996) for the same year. The difference in the estimates may be attributable to the fact that the sample contained in this essay includes only full-time full-year workers.

The results when disaggregated across years, occupation, and demographic characteristics provide important insights into the determinants of noncompliance among the self-employed and suggest actions that may be beneficial in designing tax enforcement policy. First, with regard to the determinants of noncompliance among the self-employed in Canada, I find little evidence that the degree of under-reporting has increased over time. However, to the extent that variations in occupational activity represent differences in "opportunity" to hide income for the self-employed, I find evidence which suggests that

opportunity is a determining factor. Those entrepreneurs with the greatest opportunity, those involved in economic activities which often involve cash transactions, tend to conceal income to a greater extent than those with less opportunity. Educational attainment appears to play no role in determining the extent of noncompliance among the self-employed.

Second, for policy makers involved in developing taxation enforcement policies including audit schemes the results offer the following guidance. Given the limited financial resources available to enforce compliance, a tax audit scheme which targets occupations which provide self-employed individuals with the greatest opportunity to conceal income, such as the construction and service occupations, is likely to be more effective. In addition, targeting younger entrepreneurs and households with just one self-employed member for tax audits might also be an effective strategy in enforcing tax compliance. Clearly, such a policy, if continued for a period of time, would lead to the misrepresentation of occupation and other characteristics by self-employed tax filers to avoid being audited. However, if the information from these audits is used effectively, policies can be designed to reduce noncompliance. For instance, the results of such audits could be used to identify the source of noncompliance– overstated expenses or understated income. If, for example, it is revealed that in any given occupational or demographic group most of the noncompliance is achieved through overstated expenses of one form or another, policy makers could then alter which types of expenses to allow in order to reduce noncompliance. Finally, in light of the relative magnitude of the estimates of noncompliance among the self-employed provided above, the ease with which individuals are able to be classified as self-employed is of concern. For

example, the "relabelling" of wage and salary earners as contractors raises concerns in light of these findings. To address this issue, additional restrictions might be placed on the requirements necessary for one to be classified as self-employed.

Table 1 Sample Means		
Employment Status	**Self-Employed**	**Wage and Salary**
Sample Size	682	7781
Heads Age	40.61 (9.24)	39.35 (9.73)
Young Children	0.56 (0.79)	0.56 (0.78)
Middle Children	0.67 (0.83)	0.58 (0.84)
Older Children	0.38 (0.71)	0.32 (0.66)
Head Less Than 9 Years Education	0.13 (0.34)	0.11 (0.31)
Head Some Secondary Education or Degree	0.39 (0.49)	0.42 (0.49)
Head Some Post-Secondary Education	0.09 (0.29)	0.08 (0.27)
Head Post-Secondary Degree	0.15 (0.36)	0.19 (0.39)
Head University Degree	0.24 (0.42)	0.21 (0.40)
Value of House if Owned (In 000's of 1986 real dollars)	83.01 (127. 99)	57.68 (87.30)
Rent Indicator	0.22 (0.41)	0.30 (0.46)
Atlantic Region	0.11 (0.32)	0.16 (0.37)
Quebec	0.20 (0.40)	0.22 (0.41)
Ontario	0.23 (0.42)	0.23 (0.42)

Table 1 Continued		
Prairies Region	0.31 (0.46)	0.29 (0.45)
British Columbia	0.15 (0.35)	0.10 (0.29)
ln Income	9.66 (1.32)	9.49 (1.34)
ln Food Expenditure	7.93 (1.19)	7.71 (1.21)
Values in parentheses are standard deviations.		

Table 2
Key Parameter Estimates and Estimated Degree of Under-Reporting: by Year

Year	$\hat{\beta}$	$\hat{\delta}$ [a]	$\bar{k}_l$	$\bar{k}_u$
1969	0.389 (0.026)	0.070 (0.031)	1.14	1.26
1974	*	0.044 (0.032)	1.06	1.18
1984	*	0.055 (0.037)	1.09	1.21
1986	*	0.060 (0.039)	1.11	1.23
1990	*	0.063 (0.038)	1.12	1.24
1992	*	0.076 (0.039)	1.15	1.28
N	8463			
R-Squared	0.934			

Values in Parentheses are standard errors.

For all estimates $\sigma^2_{\lambda SE} = 0.173$ and $\sigma^2_{\lambda EE} = 0.068$.

a: Estimates are the implied values for each year from the pooled regression.

*: Constrained to be the same in each year.

R-Squared taken from end stage regression.

Table 3
Key Parameter Estimates and Estimated Degree of Under-Reporting by Occupation: Selected Years

Occupation	$\hat{\delta}$	$\bar{k}_l$	$\bar{k}_u$
Managerial and Administration	0.050 (0.042)	1.1	1.2
Professional and Technical	0.055 (0.041)	1.09	1.2
Sales	0.026 (0.047)	1.02	1.12
Services	0.154 (0.065)	1.38	1.52
Product Fabricating	0.0002 (0.077)	0.96	1.05
Construction	0.157 (0.049)	1.39	1.53
Other Occupations	0.029 (0.058)	1.02	1.12
N	3901		
R-Squared	0.42		

Values in Parentheses are standard errors.

$\hat{\beta}$ = 0.415 with a standard error of 0.041.

For all estimates $\sigma^2_{\lambda SE} = 0.161$ and $\sigma^2_{\lambda EE} = 0.068$.

R-Squared taken from end stage regression.

Table 4
Key Parameter Estimates and Estimated Degree of Under-Reporting: by Head's Education Level

Head's Education	$\hat{\delta}$	$\overline{k}_l$	$\overline{k}_u$
Less Than 9 Years	0.049 (0.039)	1.08	1.19
Some Secondary or Secondary Degree	0.069 (0.023)	1.13	1.26
Some Post-Secondary	0.104 (0.047)	1.24	1.38
Post-Secondary Degree	0.017 (0.036)	0.99	1.1
University Degree	0.064 (0.030)	1.12	1.24
N	8463		
R-Squared	0.934		

Values in Parentheses are standard errors.

$\hat{\beta}$ = 0.389 with standard error 0.026

For all estimates $\sigma^2_{\lambda SE} = 0.172$ and $\sigma^2_{\lambda EE} = 0.068$.

R-Squared taken from end stage regression.

Table 5
Key Parameter Estimates and Estimated Degree of Under-Reporting: by Characteristics Observed on Income Tax Returns

	$\hat{\delta}$	$\bar{k}_l$	$\bar{k}_u$
Head's Age Category			
Aged 25-34	0.079 (0.027)	1.16	1.29
Aged 35-44	0.078 (0.022)	1.16	1.29
Aged 45-54	0.025 (0.034)	1.01	1.13
Aged 55-64	0.001 (0.043)	0.95	1.06
N	8463		
R-Squared	0.934		
Adult Members Self-Employed			
Head Only	0.069 (0.016)	1.13	1.26
Wife Only	0.077 (0.066)	1.15	1.28
Both Head and Wife	0.010 (0.036)	0.97	1.08
N	8463		
R-Squared	0.934		

Values in Parentheses are standard errors.

$\hat{\beta}$ = 0.390 with standard error 0.026 in both regressions above.

In both regression $\sigma^2_{\lambda SE}$ = 0.175 and $\sigma^2_{\lambda EE}$ = 0.068.

R-Squared taken from end stage regression.

References

Andreoni, J., Erard, B., and J. Feinstein (1998) "Tax Compliance," *Journal of Economic Literature*, Vol. 36, pp. 818-860.

Auerbach, A., and Martin Feldstein eds. (1998) Handbook of Public Economics.

Baker, P. (1993) "Taxpayer Compliance of the Self-employed: Estimates from Household Spending Data," Working Paper No. W93/14, The Institute for Fiscal Studies.

Erard, B. (1997) "A Critical Review of the Empirical Research on Canadian Tax Compliance," Working Paper 97-6, Business Income Tax Division: Dept. of Finance Canada.

Ethier, M. (1985) "The Underground Economy: A Review of the Economic Literature and New Estimates for Canada," in Income Distribution and Economic Security in Canada, ed. F. Vallaincourt, Toronto, University of Toronto Press, pp. 77-109.

Hill, R. and M. Kabir (1996) "Tax Rates, the Tax Mix and Growth of the Underground Economy in Canada: What Can We Infer?," *Canadian Tax Journal* (44:6) 1552-1583.

Kuhn, P., and Herb J. Schuetze (1999) "Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982-1995," Working Paper, McMaster University.

Mirus, R., and Roger S. Smith (1981) "Canada's Irregular Economy," *Canadian Public Policy* (7:3), pp. 444-453.

Mirus, R., and Roger S. Smith (1996) "Self-Employment, Tax Evasion and the Underground Economy: Micro-Based Estimates for Canada," Working Paper 1002, International Tax Program, Harvard Law School.

Mirus, R., R.S. Smith and V. Karoleff (1994) "Canada's Underground Economy Revisited: Update and Critique," *Canadian Public Policy* (20:3) 235-52.

Pissarides, C.A. and G. Weber (1989) "An Expenditure-Based Estimate of Britain's Black Economy," *Journal of Public Economics*, (39:1) pp. 17-32.

Pouftis, A. (1993) "Estimating the Size of the Underground Economy in Canada," M.A. Dissertation Essay, Queen's University.

Schuetze, H.J., (1998) "Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-U.S. Comparison", Canadian International Labour Network (CILN) Working Paper No. 13, McMaster University.

Slemrod, J. ed. (1992) Why People Pay Taxes: Tax Compliance and Enforcement University of Michigan Press.

Slemrod, J. and S. Yitzhaki (1997) "Tax Avoidance, Evasion and Administration," in Alan Auerbach and Martin Feldstein, eds., Handbook of Public Economics.

Smith, S., (1986) Britain's Shadow Economy, Oxford University Press, Oxford.

Spiro, P.S. (1997) "Taxes, Deficits, and the Underground Economy," in Owen Lippert and Michael Walker, eds., The Underground Economy: Global Evidence of its Size and Impact, The Fraser Institute, Vancouver, 37-52.

U.S. Government Accounting Office (1990) Tax Administration: Profiles of Major Components of the Tax Gap, Document GGD-90-53BR, Washington, DC.

Appendix Table

Appendix Table 1 Typical End Stage Regression: Dependent Variable ln Food Expenditure	
Coefficient	Base Model (Self-Employment Indicator Crossed with Year)
ln(income)	0.389 (0.026)
sed (self-employment indicator)	0.070 (0.031)
sed cross 1974	-0.026 (0.044)
sed cross 1984	-0.015 (0.048)
sed cross 1986	-0.010 (0.049)
sed cross 1990	-0.007 (0.049)
sed cross 1992	0.006 (0.05)
Quebec	-0.039 (0.014)
Ontario	0.083 (0.013)
Prairies	0.019 (0.013)
British Columbia	-0.045 (0.013)
head's age	0.009 (0.001)
head's age squared X 10^3	-0.224 (0.039)
# young children	0.067 (0.005)
# mid aged children	0.090 (0.005)
# older children	0.123 (0.006)
year2 (1974)	0.397 (0.02)
year3 (1984)	1.332 (0.062)

Appendix Table 1
Continued

year4 (1986)	1.424 (0.065)
year5 (1990)	1.585 (0.073)
year6 (1992)	1.608 (0.075)
rent	0.039 (0.009)
value of house X 10^6	0.139 (0.061)
# of rooms	0.008 (0.003)
wife's ed.: Some Secondary	-0.042 (0.013)
wife's ed.: Some Post-Second	-0.055 (0.018)
wife's ed: Post-Second Degree	-0.057 (0.016)
wife's ed.: University Degree	-0.074 (0.02)
Constant	2.695 (0.189)
Sample Size	8463
R - square	0.934

Values in parentheses are standard errors.

Chapter 5

CONCLUSION

This thesis examines the self-employment sector of the labour market, a sector which has expanded appreciably in the past two decades in Canada. Despite its rather dramatic increase in Canada, relatively few studies have been conducted that use Canadian data to study self-employment and its consequences. Using Canadian data as a primary source of information, the first two essays of this thesis examine possible reasons for the recent trends in self-employment while the third essay quantifies and characterizes income tax non-compliance among the self-employed.

In particular, the first essay uses microdata from Canada and the United States for the period 1983-1994 to assess the importance of macroeconomic conditions and the tax environment in explaining the trends in male self-employment in North America. Canada and the United States are, perhaps, more similar in overall institutional structure than any other two countries, but differ substantially in their income tax policy, macroeconomic conditions, and self-employment trends. The findings suggest that higher income tax and unemployment rates are associated with an increase in the rate of male self-employment in the two countries. Changes in the tax environment account for a considerable amount of the secular trends in male self-employment in Canada and the US over this period, while changing economic conditions did not explain much of these trends.

Unlike the first essay, the second essay concentrates only on Canadian self-employment trends and looks at the labour force experiences of both men and women. Also unlike the first essay, which focuses on secular changes in the stock of self-employed individuals, this essay examines the *dynamics* of self-employment among men and women using information on labour force transitions from 13 waves of the Survey of Consumer Finances between 1982 and 1995. The results show that the changes in transition patterns underlying the secular rise in self-employment in Canada were very different for women and men. For women, most of the increase in self-employment is attributable to an increase in their retention rates in self-employment. For men, most is attributable to a decrease in the stability of paid employment.

The final essay establishes estimates of income tax noncompliance by Canadian self-employed households using a series of Canadian Family Expenditure Surveys from 1969 to 1992. These estimates are disaggregated across years, demographic characteristics and occupation to uncover the determinants of such activities and to provide guidance to tax enforcement policy makers. The results show that the degree of noncompliance by the self-employed varies significantly by occupation, age and the number of household members self-employed. In addition, there is no clear time trend in the estimates of noncompliance, and education does not appear to influence these estimates.

Overall, the conclusions drawn from this thesis have important policy implications. The first two essays suggest that it is a poor economic climate, including relatively high income tax rates, and not an increase in entrepreneurial spirit that is compelling Canadian

men to become self-employed. This suggests that, at least for Canadian men, the recent rise in self-employment might not be a desirable trend. In light of these findings, policy analysts considering the effects of changes to the Canadian tax environment and the business cycle on the economy should take into account, not only the outcomes for the wage employed, but also those of the self-employed. One of the more positive findings of the second essay is the result which suggests that the increase in self-employment among women is attributable to an increase in their retention rates in self-employment. This may be particularly important for government sponsored self-employment assistance plans such as that offered by the Employment Insurance system. Such plans may be especially effective if targeted at women.

Finally, the magnitude of the estimates of income tax noncompliance among the self-employed reported in the third essay suggest that changes in taxation enforcement policies, as they relate to the self-employed, might be beneficial. At the very least, the estimates presented in the third essay which are disaggregated across occupation, demographic group and year, suggest certain advantages to an enforcement policy which includes a tax audit scheme targeting occupations which provide self-employed individuals with the greatest opportunity to conceal income, younger entrepreneurs and households with a single self-employed member.