

A Longitudinal Analysis of Early Self-employment in the NLSYs

by

**Yasuyo Abe and Hannah Betesh,
Berkeley Policy Associates and
A. Rupa Datta, NORC at University of Chicago**

for



Under contract no. SBAHQ-08-M-0470

Release Date: August 2010

This report was developed under a contract with the Small Business Administration, Office of Advocacy, and contains information and analysis that was reviewed and edited by officials of the Office of Advocacy. However, the final conclusions of the report do not necessarily reflect the views of the Office of Advocacy.

This study was conducted under a contract funded by the U.S. Small Business Administration (Purchase Order SBAHQ-08-M-0470). The authors would like to thank the staff at the SBA Office of Advocacy for helpful comments on earlier drafts, and especially thank Dr. Ying Lowrey for her guidance and support throughout our research. Catherine Barry also provided skillful assistance with data preparation. Any remaining errors are those of the authors.

Executive Summary

Self-employed workers form a critical segment of the workforce in the United States. These workers—whether small business owners or more informally self-employed—represent an important alternative to employer-based work in today’s uncertain economy. The purpose of this research is to provide policy-relevant analysis of the characteristics and career paths of those Americans who have chosen self-employment.

While existing literature on self-employment offers a wealth of information on the characteristics of self-employed workers at a single point in time, to date few studies have taken as their unit of analysis workers’ patterns of self-employment. Existing literature is also scant in describing how involvement in self-employment is changing for new generations of workers. The specific goals of this study are: (a) to provide new empirical findings regarding the dynamics of self-employment that underpin individual entrepreneurship during early adult work life; and (b) to document generational changes in self-employment patterns in early adult work life between two cohorts born in the second half of the twentieth century. These goals are addressed utilizing the two National Longitudinal Surveys of Youth, the 1979 Cohort (NLSY79) and the 1997 Cohort (NLSY97), which offer unparalleled breadth and depth of information on economic activity, as well as on personal and family backgrounds, and allow detailed longitudinal investigation of self-employment activities.

Our empirical inquiry is guided by a set of research questions that are designed to address the goals noted above, including:

- Does the level of self-employment over time or at a given point in time vary by individual and family characteristics and early labor-market experience?
- Do economic characteristics, measured in terms of income and net worth, alter the probability of self-employment?
- Are economic outcomes, such as income and net worth, affected by the extent of prior self-employment?
- Are there differential patterns of early self-employment among two generations of Americans born in the late twentieth century?

To investigate these questions, we adopt a range of analytical methods, including univariate descriptive analysis and multivariate analysis. These analyses reveal the following findings about the characteristics and self-employment trajectories of individuals over their early adult work life:

- **The demographic characteristics differ considerably between respondents who reported self-employment and those who did not.** Consistent with previous studies, we find that respondents who have ever been self-employed are mostly male and White. Respondents who have ever been self-employed also tended to have parents with higher educational attainment than respondents who did not report any episodes of self-employment. Respondents who reported self-employment were more likely than those who did not to have grown up in a rural area or on a farm.

- **Early exposure to self-employment increases individuals' engagement in self-employment during the ensuing early- to mid-career years.** Our multivariate analyses consistently show a strong positive link between an indicator of self-employment during ages 20-22 and the self-employment outcome measures in the subsequent periods (ages 22-41). These outcome measures include total years of self-employment and the yearly self-employment probability. The estimation results of the yearly self-employment probability also suggest that the effect of the early self-employment indicator is significantly positive and sizable, even after controlling for recent employment status.
- **Measure of financial resources is positively correlated with individuals' engagement in self-employment.** In a multivariate analysis examining each individual's self-employment decision at a given time, we find an indication that the availability of financial resources—family income and net worth—do increase the probability of self-employment, although the estimated effect size is very small.
- **Total self-employment years are positively correlated with economic outcomes, measured in terms of family income, own income, and family net worth.** Our multivariate analyses find that an additional year of self-employment increases the level of income and net worth significantly.
- **There are notable differences and similarities in the self-employment experiences in the early stage of work life between individuals born in 1960-1962 and those born 20 years later in 1980-1982. These include:**
 - The younger NLSY97 cohort has much higher self-employment rates than the older NLSY79 cohort had when the two are compared by age 23. This increase is driven largely by increases in self-employment by Black and Hispanic workers, and to a lesser extent by female workers.
 - The mechanisms through which demographics affect employer-based employment rates seem to work somewhat differently for self-employment rates. Such characteristics as race, gender, and ethnicity do not always exhibit the same patterns for self-employment and employer-based employment.

Research on a Topic of Interest to Small Business Growth and Survival

A Longitudinal Analysis of Early Self-Employment in the NLSYs

Introduction

In this study, we investigate the dynamics of self-employment at the person level, utilizing the National Longitudinal Surveys of Youth (NLSY), which provide longitudinal data on the individuals' yearly employment status in their early and mid-career years. In focusing on person-level self-employment, we regard an individual as an agent of enterprise, an "entrepreneur," and equate self-employment to an incident of entrepreneurship or business ownership. The goals of our study are twofold: (a) providing new empirical findings regarding the experiences of self-employment that underpin individual entrepreneurship, particularly during early adult work life; and (b) documenting generational changes in self-employment patterns in early adult work life between two generations from the second half of the twentieth century. We take advantage of the extensive information available from the NLSY studies and extend the data to address these goals.

In studying self-employment status dynamics over a period of years, this study uses the concept of a career trajectory as a measure to summarize individuals' involvement in self-employment over time. Earlier studies of self-employment have largely focused on each employment episode, with minimal emphasis on the context of life cycle; on the other hand, empirical data suggest that a large number of those who have ever been self-employed entered and have exited self-employment multiple times throughout their work life, and that there may be distinctive pathways into and out of self-employment, triggered by varying circumstances and preferences (see Appendix B for discussion of previous studies).

Our empirical inquiry is guided by a set of research questions that are designed to address the goals described above, including:

- Does the level of self-employment over time or at a given point in time vary by individual and family characteristics and early labor-market experience?
- Do economic characteristics, measured in terms of income and net worth, alter the probability of self-employment?
- Are economic outcomes, such as income and net worth, affected by the extent of prior self-employment?
- Are there differential patterns of early self-employment among two generations of Americans born in the late twentieth century?

These research questions are not intended to test any particular theoretical model of entrepreneurial behaviors and choices; rather, they are used primarily as tools to navigate our empirical inquiry of early- to mid-career self-employment trajectories.

We begin with a discussion of prior literature and an introduction to the data sources, the National Longitudinal Surveys of Youth, 1979 and 1997 Cohorts. We then turn to the NLSY79 for our in-depth investigation of early self-employment in a longitudinal context. This investigation includes both univariate and multivariate analyses. We follow with a brief comparison of early self-employment experiences in the NLSY97 and NLSY79, ending with some concluding remarks.

Review of Previous Studies

There has been considerable research using the individual as a unit of analysis to understand small business dynamics. These studies are underpinned by a theoretical model of entrepreneurial choice, and the empirical investigations typically regard an individual's engagement in self-employment as an entrepreneurial decision to form a business. Economic and other social science theories predict that a wide variety of factors—financial, social, educational, familial, psychological, or biological—will influence an individual's engagement in self-employment. Many of the previous studies of individual entrepreneurs were designed to determine what individual characteristics or other factors explain the “formation” of a business and, to a lesser extent, the closure or duration of a business. Accordingly, there now exists a substantial volume of research findings, albeit not always consistent, on who is more likely to be or to become self-employed and what other factors influence self-employment selection. Below we summarize highlights of the relevant studies that we reviewed (See Appendix B for a fuller review of the literature).

Many of the previous studies investigated what individual characteristics or other factors explain engagement in self-employment. For example, in an earlier empirical study of individual-level data, Evans and Leighton (1989) found, among other things, that: the probability of becoming self-employed was largely independent of age and past experience; those with greater assets were more likely to become self-employed; unemployed and low-wage workers were more likely to become self-employed; and those with a belief that one controls one's own destiny (an internal locus of control) were more likely to be self-employed. More recent studies have investigated the effects of a range of factors on self-employment, including: education, previous work experience, own and parents' assets (Henley 2004, 2005, Moutray 2007), particular types of resources such as homeownership and technology (Moutray 2007, Georgellis et al 2005), military service (Moutray 2007), incarceration (Fairlie 2005B), parental experience of self-employment (Blanchflower and Oswald 2007, Fairlie and Robb 2007B), spousal characteristics (Karoly and Zissimopoulos 2004), and other household members' status, and presence of children (Dickson, et al. 2008; Kim, et al. 2006; Hundley 2006; Cavalluzzo & Wolken 2005; Dunn & Holtz-Eakin 2000; Colombier & Masclet 2008; Taniguchi 2002; Budig 2006; Salazar 2007; Kepler & Shane 2007; Fairlie 2005A; Fairlie & Robb, 2007B).

Previous studies consistently found gender and race as factors influencing self-employment. Empirical data indicate that women and minorities, particularly African Americans and Hispanics, lag behind White men in self-employment rates and other small business performance rates (SBA 2007 and 2006). Overall, past research points to the racial difference in individuals'

access to resources (e.g., financial capital, business experience) as leading explanations for the racial gap (Fairlie and Robb 2007A; Blanchflower, et al. 2003 ; Cavalluzzo 2005; Lowrey 2007). Such findings, combined with other research results demonstrating liquidity constraints as a barrier to starting a business (e.g., Evans and Jovanovic 1989), suggest that African Americans would be less likely to be self-employed than their White counterparts. Past research on gender and self-employment in the United States also documented how females and males differed in their characteristics as entrepreneurs, or described the self-employment selection process for females. An emerging consensus about self-employed females seems to be that they are significantly different from their male counterparts not only in their financial or human capital but, more importantly, in their reasons for entering self-employment, their attitudes, and their preferences. For instance, Kepler and Shane (2007) find gender differences among entrepreneurs regarding pecuniary motivations, expectations as to their future income from the business, risk preferences, technological intensity of their businesses, and their approach to identifying opportunities. Similarly, Schiller and Crewson (1997) showed that marriage increased the probability of self-employment for women while it decreased the probability for men. Dunn and Holtz-Eakin (1995) found that own financial assets modestly increased the probability of self-employment for men but did not have any effect for women. Other studies that focused exclusively on female self-employment also showed that family structures (marriage and children) mattered in women's entry to self-employment (Budig 2006; Carr 1996; Taniguchi 2002).

Aside from the selection into self-employment, a number of studies also investigated other aspects of business dynamics, such as exit from or duration of self-employment. For example, using the NLSY79, Schiller and Crewson (1997) observed that the total years spent in self-employment among young adults was relatively short—less than three years over the 11-year period, among those ever self-employed. Schiller and Crewson also demonstrated that self-employment indicators (both total years spent and income received) vary considerably by gender. This study is corroborated by findings that young women had lower self-employment entry rates and higher self-employment exit rates (Fairlie 2005a). More recently, Rissman (2006) used the same data (NLSY79) to examine factors affecting the duration of self-employment. She found that young men are more likely to exit from self-employment when the economy is better, supporting an argument that self-employment is largely represented not by entrepreneurs but by discouraged wage workers seeking a second-best alternative to unemployment. On the other hand, Bates and Servon (2000) investigated the self-employment experiences of those who started their own businesses because they were unable to find suitable employment elsewhere, finding that this particular population had both lower educational attainment levels and fewer assets than the population generally targeted by small business policies and services. Other studies have also found that assets are an important determinant of self-employment duration and exit rates (Evans and Jovanovic 1989; Evans and Leighton 1989; Bates 1997; Holtz-Eakin, et al. 1994; Holtz-Eakin 1994; Fairlie 1999; Dunn and Holtz-Eakin 2000; and Blanchflower and Oswald 1998).

While most previous studies do not fully describe the dynamics of self-employment over the work life of an individual, there have been studies that focused on self-employment as a process, rather than a static status. For instance, Wennberg, et al. (2007) highlighted that individuals may choose part-time entry to self-employment as a strategy to manage the uncertainty of entrepreneurship, and argue that part-time and full-time employment should be distinguished and that their interrelation needs to be studied in the context of self-employment dynamics. Other

studies focused on the relationships between the employment choices made earlier in the life cycle and later career outcomes. For instance, Williams (2004) identified differences in later career outcomes based on early experiences with self-employment, finding that NLSY79 respondents who had been self-employed as young adults were less likely to complete higher education and had lower earnings over time than their non-self-employed peers.

Existing literature thus offers substantial background on the determinants of self-employment and the demographics of the self-employed population, as well as on pathways to self-employment. Our study builds on these findings, as described in subsequent sections, by analyzing the career paths of self-employed individuals from a life cycle perspective and comparing self-employed workers from two generations.

Research Methods

This study begins with a series of univariate descriptive analyses to compare personal and family characteristics across individuals with varying levels of early self-employment. We apply several different multivariate analysis methods to investigate the determinants of the self-employment trajectories and to assess the economic outcomes attributable to these self-employment measures.

- 1) We use the tobit method to model the level of self-employment over time, measured as total years of self-employment, with the goal of estimating the effects of various background characteristics on total years of self-employment.
- 2) Instead of focusing on respondents' employment experiences over all years in the data, we model the decision to engage in self-employment in any given year. We estimate this model by applying the generalized estimating equation (GEE) method, which allows us to adjust for within-person correlations that arise from repeated observations of individuals in the longitudinal data. Again, the goal of the GEE model is to estimate the effect of various background characteristics.
- 3) We model the effects of the employment trajectory type or of the level of self-employment on economic outcomes using the ordinary least squares (OLS) method. As dependent variables, we use family net income and net worth, both measured in alternative forms (level, growth rate, and coefficient of variation). More details on model specifications and estimation methods used in the study are presented in Appendix C.

We close with another series of univariate descriptive analyses, these comparing early self-employment among birth cohorts from the early 1960s and the early 1980s.

Data Description

This study uses the two National Longitudinal Surveys of Youth, the 1979 Cohort (NLSY79) and the 1997 Cohort (NLSY97). We briefly present the descriptions of these data and highlight the merits of using them. (See Appendix A for additional details on background and descriptions of the NLSYs.)

The NLSY79 and NLSY97 are two in a series of studies sponsored by the National Longitudinal Surveys Program of the Bureau of Labor Statistics, U.S. Department of Labor. Each of the two studies focuses on a cluster of birth cohorts, following them from adolescence into adulthood in order to better understand the dynamics of labor force experience, especially as affected by early life experiences. The studies have a parallel design that permits multi-faceted comparison between two American generations born in the second half of the twentieth century. Exhibit 1 summarizes key characteristics of the two data sets. For both data sets, we apply sample weights to generalize to the target population of each data set, correcting for oversampling of minorities (in the NLSY79) and low-income households.¹

Exhibit 1
Overview of NLSY79 and NLSY97 Data Sets

<u>Characteristic</u>	<u>NLSY79</u>	<u>NLSY97</u>
Birth years	1957-1964	1980-1984
Population represented	Individuals from selected birth years living in U.S. households in 1979	Individuals from selected birth years living in U.S. households in 1997.
First year of data	1979	1997
Initial sample size	12,686	8,984
Ages at first data collection	14-21	12-16
Ages at most recent wave	41-48	23-27
Number of waves of data	22	11
Periodicity of waves	Annual to 1994, biennial since	Annual
Oversamples used in this study	Black, Hispanic	Black, Hispanic

In contrast to such rich sources of data on business ownership and entrepreneurship as the Survey of Business Owners and Self-Employed Persons (part of the Economic Census) and its predecessor, the 1992 Characteristics of Business Owners, or the Panel Study of Entrepreneurial Dynamics, the relative strength of the NLSYs is in contextualizing the self-employment experience within the range of life cycle behaviors, such as education, family formation, assets, and health status.

This study extends the prior studies based on NLSY studies in two important ways. First, we use the underutilized NLSY97, in combination with the NLSY79. More than two dozen papers have been written on self-employment and entrepreneurship using the NLSY79 or its predecessors in the National Longitudinal Surveys Program, the Original Cohorts; however, to our knowledge, few studies on self-employment have been conducted using NLSY97. The use of the NLSY97, a much younger cohort just beginning their labor force participation, can provide new insight into how younger generations today are engaging in self-employment activities.

¹ We use the initial wave sampling weight, which corrects for probability of selection into the sample. For the NLS79, we adjusted the initial wave sampling weight to account for the phaseout of the military sample after the 1984 wave. We use these adjusted baseline weights for our analysis.

Second, we adopt a more intensive and thorough use of available information in the NLSY79 data to identify individuals who are self-employed than approaches taken in the prior studies. Unlike most previous research, which only examines the current or most recent job, our study utilizes information collected on up to five jobs during the survey year, and includes data collected retrospectively from individuals who may have missed an occasional survey round. In later years, where attrition has increased, the use of these retrospective data can reduce nonresponse substantially.² More details on how we define our self-employment measure are provided in Appendix A.

Our more intensive use of the NLSY79 data has a substantial impact on the estimates of self-employment among the respondents. In Exhibit 1, we compare NLSY79 self-employment rates using the methodology used in previous studies and the methodology used in this study for select years. The exhibit shows that our measure generates considerably higher rates of self-employment than the measure previously used, ranging from 16 percent higher in 2000 to more than twice as much in 1980. Increases are to be expected, since counting more jobs gives individuals more opportunities to report self-employment. There are a few implications for this definition change. First, we have a more complete definition of self-employment. Second, we are able to include more individuals in our analysis of self-employment, thus increasing sample sizes and analytical power. Third, because we are including jobs that may have been secondary or for shorter durations, we may be incorporating more workers whose self-employment is either casual or unsuccessful.

Exhibit 2		
Comparison of NLSY79 Self-Employment Rates		
Using Current/Most Recent Job Only versus Using Up to Five Jobs		
(Weighted)		
<u>Percent Currently Self-Employed</u>		
<u>Year</u>	<u>Current/Most Recent job</u>	<u>Up to 5 Jobs</u>
1980	1.16	2.34
1990	5.27	8.73
2000	5.75	6.67
2004	7.04	9.70

Source: NLSY79

For our primary investigations of early self-employment, we include those NLSY79 respondents for whom we can reliably construct a yearly employment status indicator from ages 22 through 41, covering the 20-year span of early adult work life. Of the original sample of NLSY79, 7,482 respondents are retained.

For comparison between the NLSY97 and the NLSY79, we use three birth cohorts of data from each survey: the 1960-1962 birth years from the NLSY79 and the 1980-1982 birth years from

² We are, however, more conservative in one way: we code workers as self-employed or dually employed in a calendar year only if they worked at least 100 hours of self-employment in that year. It is possible that alternative definitions of self-employment would include some workers with very low hours of self-employed work, while our definition would fail to classify them as self-employed.

the NLSY97. Using this narrower age range allows us to make comparisons up to age 23, which is not yet possible for the youngest of the NLSY97 respondents, born through 1984. The NLSY97 analysis sample includes 5,486 respondents, excluding fewer than 50 in the eligible age range who are of mixed race. A comparison group from NLSY79 consists of 4,125 respondents, which may include respondents who were not included in the sample for primary analyses described above (e.g. those who dropped out of the study after age 23).

Yearly Employment Status

Critical to the study is the preparation of a yearly employment status indicator, which is used to identify who is self-employed, employer-employed, or not employed. The employment status indicator was created for the NLSY79 data in several steps, which are designed to capture the most complete information for each survey respondent, many of whom missed interviews in some survey rounds but are generally surveyed again and usually provide information covering the time period(s) of the missed interview(s). These later rounds are used to fill in gaps from noninterviews. Following these steps, we created the employment status indicator, *empstat*, as follows:

- empstat* =
- 1 Exclusively self-employed. Reported 100 or more hours worked per year in one or more self-employed jobs, but did not work or worked less than 100 hours in an employer-employed job.
 - 2 Dual-employed (both self-employed and employer-employed). Reported 100 or more hours worked per year in both types of jobs.
 - 3 Exclusively employer-employed. Reported 100 or more hours worked in one or more employer-based jobs, but did not work or worked less than 100 hours in a self-employed job.
 - 4 Nonemployed. Includes those who were unemployed, out of the labor force, or worked less than 100 hours per year. (This category may be broken down by those in school and not in school.)

We hereafter use the terms “exclusively self-employed,” “dual-employed,” and “exclusively employer-employed” to distinguish the mutually exclusive categories defined for the employment indicator *empstat*. We will also use the term “self-employed” to mean either exclusively self-employed or dual-employed (i.e., *empstat* = 1 or 2). Similarly, we will use the term employer-employed to mean either exclusively employer-employed or dual-employed (i.e. *empstat* = 2 or 3). For the NLSY97 employment status indicator, we use the created variables for yearly hours worked.

Besides the employment status indicator, this study utilizes the NLSYs’ extensive information about the respondents to create demographic profiles of our analysis sample. For the NLSY79, background variables examined in this study include country of birth, country of residence at age 14, South/non-South regional residence at age 14, household structure at age 14, highest grade completed by mother and father by 1979, race/ethnicity, gender, and a variable identifying persons born outside of the U.S. with U.S. parentage, making them citizens at birth. Most variables are self-reported, although the race/ethnicity variable is chosen by the interviewer and

consists of three choices: Black, Hispanic, and non-Black/non-Hispanic. Other demographic variables include characteristics that often change over time, such as completing education, marrying, and having children. The following variables are examined in the study: marital status; school enrollment status; highest grade completed as of survey year; number of children ever born; urban/rural residence; standard metropolitan statistical area (SMSA) residence; regional residence; family size; family poverty status; health limitations on work; health plan coverage; total family income; own income; and family net worth. For the NLSY97, we examined covariates that are equivalent to those listed above. For comparisons between the two surveys, we use only those covariates that can be determined as comparable. Additional discussions on the data used in this study can be found in Appendix A.

Patterns of Self-Employment During Early Adult Work Life

Utilizing the extensive longitudinal data available from the NLSY79 cohort, we address the first three research questions regarding the characteristics of the self-employed individuals and the determinants of self-employment. First, we present a summary of individuals' self-employment experiences from ages 22 through 41, covering the 20-year span of early adult work life.

Exhibit 3					
Distribution of Yearly Employment Status by Age (Weighted)					
A. NLSY79 All Individuals (N=7,482)					
Age	Exclusively Self-employed (<i>empstat</i> =1) (percent)	Dual-employed (<i>empstat</i> =2) (percent)	Exclusively Employer-employed (<i>empstat</i> =3) (percent)	Not Working and Not in School (<i>empstat</i> =4) (percent)	Not Working and In School (<i>empstat</i> =4) (percent)
22	1.4	3.0	81.0	10.5	4.1
27	3.8	5.3	79.2	10.7	1.0
32	5.8	5.3	75.9	12.3	0.7
37	5.8	3.9	76.9	12.9	0.6
41	7.5	3.5	74.4	13.6	1.0
B. NLSY79 Ever Self-Employed (N=2,394)					
	(percent)	(percent)	(percent)	(percent)	(percent)
22	4.1	8.5	73.5	9.8	4.1
27	10.8	15.2	64.1	9.2	0.8
32	16.6	15.1	56.2	11.3	0.8
37	16.6	11.1	59.9	11.8	0.7
41	21.4	10.1	54.2	13.3	1.1

Source: NLSY79

Exhibit 3 provides the yearly employment status indicator, as defined above, at select ages for the full sample, and then for those ever self-employed in the NLSY79. Among all workers, the share of those who were exclusively self-employed steadily increased over time; however, the share of those who were dual-employed (that is, both self-employed and employer-employed) did not increase as steadily. Not surprisingly, the share of those who were not working but in school (full-time students) decreased with age. The share of those who were not working but not in school appears to have increased slightly, probably reflecting the increase in the out-of-labor-force population, especially among women. As one might have expected, the largest category was those who were exclusively employer-employed. The share of this group decreased over

time, mainly due to the shift toward exclusive self-employment. Still, the overwhelming majority—three quarters or more—of the sample was exclusively employer-employed throughout their early adult work lives.

Exhibit 4				
Total Number of Years from Age 22 to 41 in each Employment Status Category by Self-Employment Status (Weighted)				
	<u>Ever Self-Employed</u>		<u>Never Self-Employed</u>	
	<u>Mean</u>	<u>Std.Err.</u>	<u>Mean</u>	<u>Std.Err.</u>
Employment Status				
Self-Employed	2.754	0.090	--	--
Dual-Employed	2.620	0.058	--	--
Employer-Employed	12.162	0.124	17.316	0.072
Not Working	2.465	0.085	2.684	0.072

Exhibit 4 summarizes average years in each employment status for those who were ever self-employed and those who were never self-employed in the study sample. These summary figures reflect the work history of an individual over the 20-year study period. Years spent in each employment status over the 20-year period differ between the ever and never self-employed individuals. Those in the ever self-employed sample report, on average, 2.8 years of exclusive self-employment, 2.6 years of dual employment, 12.2 years of employer-based employment, and 2.5 years of nonemployment; on the other hand, the never self-employed sample reports 17.3 years of employer-based employment and 2.7 years of nonemployment. We can also tabulate the number of spells of employment of each type (not shown). Ever self-employed workers had an average of 2.15 spells of employer-based employment, while never self-employed workers averaged 1.49 spells of employer-based employment.

Exhibit 5 summarizes the variations in early employment experience among this generation between ages 22 and 41. Just over one-third of these individuals reported self-employment at least once over the 20-year period; however, only 14.8 percent spent more than four of the 20 years exclusively self-employed. Among the never self-employed, we see an average of 17.3 years of the 20 spent exclusively in employer-based employment.

Exhibit 5
Total Number of Years during Ages 22-41 in Each Employment Type
By Years of Self-Employment (Weighted)

Years of Self-Employment	Number (Percentage)	Exclusively Self-Employed		Dual-Employed		Exclusively Employer-Employed		Self-Employed (a)+(b)		Employer-Employed (b) + (c)	
		Mean	S.E.*	Mean	S.E.*	Mean	S.E.*	Mean	S.E.*	Mean	S.E.*
10 or more years	379 (6.1%)	8.6	0.26	5.0	0.20	5.4	0.18	13.6	0.17	10.4	0.29
5-9 years	577 (8.7%)	3.3	0.11	3.5	0.11	10.9	0.16	6.7	0.07	14.4	0.21
1-4 years	1438 (20.2%)	0.8	0.03	1.5	0.03	14.7	0.13	2.3	0.03	16.3	0.14
Never self-employed	5088 (64.9%)	n.a.	n.a.	n.a.	n.a.	17.3	0.07	n.a.	n.a.	17.3	0.07

Note: * Linearized standard errors.

Exhibit 6 compares basic demographic characteristics for ever self-employed and never self-employed individuals. The statistics confirm results from previous studies of self-employment, finding that women and Blacks are less likely to be self-employed. The AFQT (Armed Forces Qualifying Test) statistics indicate a small and statistically significant advantage in cognitive ability among the self-employed. A similarly small and statistically significant result is that the ever self-employed are less likely to live in towns or cities (as compared to nonfarm rural areas, farms or ranches). The exhibit also indicates that self-employed individuals have mothers who are significantly more likely to have completed high school than mothers of never self-employed individuals.

Exhibit 6

Demographic Characteristics by Ever/Never Self-Employed (Weighted)

	Ever self-employed	Never self-employed	
Mean age in 1979	18.3	18.2	
Female (percent)	45.5	53.3	***
Race and ethnicity			
Black (percent)	11.3	17.4	***
Hispanic (percent)	6.0	7.3	
Non-Black/Non-Hispanic (percent)	82.7	75.3	
Mean AFQT (score from 0 to 99)	50.3	48.5	**
Mean highest grade completed			
At age 22	12.4	12.5	
At age 41	13.5	13.5	
Completed 12 th grade, at 41 (percent)	91.0	92.1	
Completed 4 yrs in college at 41 (percent)	24.8	26.1	
Residence at age 14			
Lived in town or city	76.3	79.3	**
Lived in rural area (not farm)	17.4	15.9	
Lived on farm or ranch	6.3	4.8	
<i>Father Characteristics</i>			
Percent born outside United States	5.6	6.3	
Percent completed 12th grade or more	68.0	65.8	
<i>Mother Characteristics</i>			
Percent born outside United States	6.6	7.1	
Percent completed 12th grade or more	71.5	66.8	***
N	2,394	5,088	

*Note: *** Difference across columns is statistically significant at the 1 percent level; ** Statistically significant at the 5 percent level.*

Individual and Family Characteristics by Employment Trajectory Types

We next examine how those who are extensively self-employed might differ from those who either did not engage in self-employment or engaged less in self-employment. In Exhibit 7, we compare household characteristics between ever self-employed and never self-employed individuals at different points over the 20 years of this study. The exhibit separately describes the 2,394 individuals who were self-employed at any time from age 22 to 41, and the 5,088 individuals who were never self-employed during that time. Compared to those never self-employed by age 41, the ever self-employed had slightly smaller family size at age 22, and considerably higher net family income at age 32. At both ages 32 and 41, the self-employed were less likely to be covered by a health plan than those who were never self-employed. Note that the ever self-employed category is defined as of age 41 (i.e., the individual has engaged in self-employment at some point before age 41), so the observations about the ever self-employed group at age 22 actually pertain to individuals who might not have their first self-employment experience until many years later. This table, then, provides a glimpse into possible selection effects into self-employment, rather than describing the predictors or effects of self-employment.

Exhibit 7				
Respondent Characteristics by Ever/Never Self-Employed Status				
(Weighted)				
		Ever self-employed	Never self-employed	
Married (percent)	at age 22	27.5	25.3	
	at age 32	61.4	61.0	
	at age 41	65.3	64.6	
Number of children (mean)	at age 22	0.28	0.28	
	at age 32	1.36	1.34	
	at age 41	1.84	1.78	
Family size (mean)	at age 22	3.1	3.4	***
	at age 32	3.1	3.1	
	at age 41	3.2	3.2	
Lived in an SMSA (percent)	at age 22	75.1	74.2	
	at age 32	78.4	79.3	
	at age 41	93.3	94.8	**
Net family income (mean \$)	at age 22	21,366	21,444	
	at age 32	57,309	45,360	***
	at age 41	67,212	64,915	
Poverty rate in a previous year (percent)	at age 22	14.2	15.8	
	at age 32	12.9	10.8	**
	at age 41	11.4	10.5	
Covered by health plan (percent)	at age 32	77.8	86.7	***
	at age 41	79.5	87.8	***
N		2,394	5,088	

*Note: *** Difference across columns is statistically significant at the 1 percent level; ** Statistically significant at the 5 percent level.*

Multivariate Analyses

Examining the patterns of background and outcome characteristics based on empirical self-employment measures suggests that the length of and specialization in self-employment over time are likely to be correlated with individual characteristics both at baseline and later in life. Having established this correlation, we now turn to a multivariate analysis in which we analyze the factors affecting and being affected by self-employment measures.

As outlined earlier, we conducted a range of multivariate analyses designed to investigate a set of research questions about empirical relationships between self-employment and individual and family characteristics and between self-employment and economic outcomes. Our goal is to provide regression-adjusted empirical descriptions of self-employment experiences. Below we summarize the findings from each multivariate analysis model as it relates to the corresponding guiding research question. For details of the estimation models and their results, see Appendix C.

Personal Characteristics, Early Labor Market Experience, and the Extent of Self-Employment

This section investigates the question, “Does the number of total years in self-employment vary by individual and family background characteristics, and/or early labor-market experience?”

To investigate this question, we applied the tobit estimation method to model total years of self-employment, which is viewed as a cumulative self-employment outcome over the respondent’s early adult work life. Our hypothesis is that background characteristics or early labor market experiences are predictive of the total number of years that individuals are self-employed over the ensuing 20-year period.

The results from the tobit estimation of self-employment are shown in Exhibit 8. Column (1) provides estimates for total years of self-employment for the full 20 years of the study, ages 22 through 41. The subsequent columns (2) and (3) look separately at the first ten and latter ten years of the period. Across all three models, being female and being Black are associated with fewer years of self-employment. Other results common to all three models are that early engagement in self-employment is estimated to have a sizable and significant positive effect on total years of self-employment, and that having a parent who graduated from college is associated with an increase in total years of self-employment.

There are some interesting changes between the two ten year periods. For example, we see that the negative effect of being Hispanic on total years of self-employment is primarily driven by the first 10 years. By the time individuals are age 32 or older, the effect of being Hispanic is no longer statistically significant. Consider also the interaction of being female with the number of children the female has at age 22. In column (1), we see no statistically significant effects for this interaction term. Column (2), however, indicates that females with at least one child at age 22 are more likely to be self-employed, while their children are young. For the second ten years, being female and having had children at age 22 is associated with fewer years of self-employment.

Exhibit 8
Tobit Estimation of Total Number of Years Self-Employed
(Weighted)

	(1)		(2)		(3)	
	Between Age 22 and 41		Between Age 22 and 31		Between Age 32 and 41	
	<u>Coef.</u>	<u>Std. Err.</u>	<u>Coef.</u>	<u>Std. Err.</u>	<u>Coef.</u>	<u>Std. Err.</u>
Constant	-1.629	1.125	-2.941	0.814 ***	-3.468	1.088 ***
Female	-1.748	0.296 ***	-1.315	0.205 ***	-1.210	0.285 ***
Black	-2.261	0.334 ***	-1.468	0.238 ***	-1.791	0.327 ***
Hispanic	-1.503	0.482 ***	-1.216	0.340 ***	-0.717	0.468
Married at age 22	0.387	0.327	0.457	0.229 **	0.112	0.320
Number of children at age 22	0.141	0.376	-0.322	0.265	0.490	0.354
Female * Number of children at age 22	-0.082	0.420	0.621	0.298 **	-0.991	0.411 **
Self-employed at ages 20-22	9.206	0.500 ***	7.034	0.304 ***	4.415	0.499 ***
Highest grade completed at 22	-0.099	0.094	-0.013	0.067	-0.092	0.090
AFQT score	-0.003	0.006	-0.004	0.004	0.002	0.006
Family in poverty last yr (at 22)	-0.105	0.362	-0.213	0.263	0.258	0.350
Health problem affecting wk at 22	0.757	0.508	0.791	0.362	0.247	0.493
Parent born outside US	-0.564	0.500	-0.577	0.351 *	-0.324	0.492
Parent graduated college	1.551	0.357 ***	0.774	0.250 ***	1.269	0.341 ***
Spoke non-English growing up	0.473	0.492	0.739	0.329 **	0.013	0.475
Lived in rural area (nonfarm) at 14	0.170	0.349	0.066	0.245	-0.012	0.345
Lived on farm at age 14	1.625	0.588 ***	0.890	0.387 **	1.398	0.528 ***
Test of joint significance of repressors	F(28,7452)=19.8***		(F28,7452)=26.9***		(F27,7453)=9.22***	
Number of observations(a)	7,480		7,480		7,480	

*** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent

Note: The observed total number of years self-employed ranged from 0 to 20 (0 to 10 for columns (2) and (3)). The number of self-employed years are regarded as censored at zero. In addition to covariates listed in the table, each estimation model includes a set of dummy adjustment variables for missing covariates and cohort dummy variables (the calendar year indicator at age 22). Fewer (2 less) control variables were included in the model (3) as they become collinear.

(a) The number of observations used in the estimation is reported. The total analysis sample is 7,482, but two individuals are given zero weight and excluded from the estimation with the use of weights.

Modeling Point-in-Time Self-Employment

The next set of analyses focus on the question, “Does self-employment at a given point in time vary by baseline characteristics and early labor-market experience as well as by contemporaneous events and changes in personal and family circumstances, including asset accumulation and health care coverage?”

In order to assess further how selection into self-employment varies by individual and family background and other key factors suspected of affecting self-employment (such as assets and health care coverage), we modeled the probability of self-employment utilizing the panel of individual data. We applied the generalized estimating equations (GEE) method to take into account within-individual correlation in the panel data.

To aid the interpretation of the results of our GEE analyses, we report in Exhibit 9 the estimated marginal effect size, evaluated at the sample means of independent variables. The self-employment probability evaluated at the means of independent variables is 4.7 percent. The exhibit shows the marginal effects associated with a unit increase in a covariate (i.e., a background factor) given this expected probability of self-employment. For the dummy variables, a unit change refers to a change from 0 to 1. In continuous variables, it refers to the unit used in estimating the original model. For example, the exhibit shows that if a person was self-employed between age 20 and 22, her current probability of self-employment increases by 7.9 percentage points, while a \$1,000 increase in the annual total income only leads to a 0.001 percentage point increase in the likelihood of self-employment. (See Appendix C for additional discussion on our GEE analysis and estimation results.)

Our GEE analyses indicate that, not surprisingly, the most important predictor of self-employment is previous employment status. Those who were self-employed (either exclusively self-employed or dual-employed) in the previous year are more likely to be self-employed in the current year than those who were employer-employed. For example, we estimated that the probability of self-employment increases by 56 percentage points (from the mean likelihood of self-employment of 4.7 percent) if individuals were exclusively self-employed in the prior year. The exceptionally large effect underscores that the previous year’s self-employment status is a strong predictor of the outcome. The exhibit also shows that early exposure to self-employment is another important factor. As noted above, the likelihood of self-employment among those who were self-employed between ages 20 and 21 is estimated to be 7.9 percentage points higher than those who did not have such experiences.

The GEE estimates also indicate that the probability of being self-employed is influenced by other early life experiences and family history. For example, the probability of being self-employed is significantly higher among those who had at least one college-educated parent, who lived on a farm when they were young, and whose families were in poverty at age 22. However, as shown in Exhibit 9, the estimated effect sizes are modest.

The GEE estimates also indicate that the probability of being self-employed varied significantly across individuals’ demographic and current characteristics. For example, we found that the likelihood of being self-employed was significantly lower among females, minorities, and those who report a health problem that affects the ability to work. On the other hand, the likelihood of self-employment is higher among those who are currently married and live in an SMSA.

Consistent with our earlier findings, the self-employment probability also increases with age. Completing a four-year college degree is negatively related to self-employment probability, suggesting that college graduates are less likely to be self-employed; however, controlling for completions of a college degree, an additional year of education (measured by the highest grade completed) increases the probability of being self-employed. One interpretation of this result is that additional years of schooling help equip individuals with skills needed to be self-employed, but higher degrees do not necessarily increase the likelihood of self-employment. Another way to look at the results is that college graduates are less likely to be self-employed than noncollege graduates, but among college graduates, those with more schooling are still more likely to be self-employed. As in the case of family background effects, the estimated effects are, however, relatively small in size.

The effects of financial resources were examined using family income and net worth measures. These variables are measured in the previous year to minimize the endogeneity problem arising from self-employment statuses and income. The estimation results show that the effects of financial resources are positive and significant, but the effect is so small that an increase in \$1,000 or even \$100,000 does not make any consequential difference, controlling for other factors, in the probability of being self-employed.

Exhibit 9			
Marginal Effects on the Probability of Self Employment			
Change in Percentage Points by a Unit Change in Covariate			
	<u>dY/dx</u>	<u>Std. Err.</u>	
<i>Background/History</i>			
Parent born outside U.S.	-0.00144	0.00315	
Parent completed college	0.00612	0.00272	**
Lived on farm at age 14	0.01979	0.00504	***
Family in poverty at age 22	-0.00513	0.00260	**
Self-employed ages 20-22	0.07933	0.00732	***
<i>Current year characteristics</i>			
Age	0.00824	0.00273	***
Age squared	-0.00011	0.00004	***
Female	-0.02151	0.00281	***
Black	-0.01829	0.00191	***
Hispanic	-0.01389	0.00222	***
Highest grade completed	0.00211	0.00079	***
Did not complete 12th grade	0.00505	0.00401	
Completed four-year college	-0.00846	0.00326	**
Number of children	0.00017	0.00094	
Married	0.00482	0.00161	***
Health problem affecting work	-0.00775	0.00270	***
Live in an SMSA	0.00528	0.00203	***
<i>Vbls from Previous Year (t-1)</i>			
Family income (\$'000)	0.00001	0.00001	**
Family net worth (\$'000)	0.00001	0.00000	***
Exclusively self-employed	0.56658	0.00906	***
Dual-employed	0.28348	0.00912	***
Not employed	0.02701	0.00343	***

Notes: The GEE estimation results from Exhibit F.8 (3) are used to compute the marginal effects. Estimates are based on the sample means of the covariates. The expected probability at the covariate means is 0.047.

**** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent*

Self-Employment as a Predictor of Financial Outcomes

Whereas the prior set of estimates model self-employment as an outcome of an individual's financial resources, the final set of analyses considers economic outcomes as products of self-employment experiences. The question we investigate here is, "Do economic outcomes, measured in terms of income and net worth, vary by self-employment trajectory type and by years of self-employment?" In exploring this question, we shift the focus to economic performance resulting from self-employment choices over time. Our hypothesis is that the difference in how individuals experience self-employment over time leads to different financial experiences and outcomes. In particular, we focused on family income, individuals' own income, and family net worth as outcomes and estimated them using the OLS method.

Exhibit 10

OLS Estimation of Family and Individual Income Measures (Weighted)

	(1) Family Income at Age 41 (in '000s)			(2) Average Family Income (in '000s)		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Number of years self-employed	1.10	0.31	***	1.39	0.21	***
Number of years employer-employed	0.92	0.19	***	0.84	0.14	***
Female	-9.23	1.65	***	-18.28	1.25	***
Black	-4.84	1.41	***	4.93	0.95	***
Hispanic	1.84	1.72		5.92	1.16	***
AFQT	0.54	0.04	***	0.20	0.02	***
Parent's highest grade completed	1.63	0.33	***	1.50	0.22	***
Spouse income (in '000)	0.87	0.05	***	2.05	0.10	***
Constant	-14.97	4.93	***	-12.51	3.30	***
Number of observations	6,233			5,293		
	(3) Own Income at Age 41 (in '000s)			(4) Average Own Income (in '000s)		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Number of years self-employed	1.17	0.26	***	0.63	0.08	***
Number of years employer-employed	1.54	0.14	***	0.58	0.09	***
Female	-20.51	1.49	***	-9.04	0.39	***
Black	1.49	1.32		-0.55	0.35	
Hispanic	2.93	1.56	*	1.01	0.49	**
AFQT	0.37	0.03	***	0.14	0.01	***
Parent's highest grade completed	0.82	0.29	***	0.44	0.08	***
Constant	-8.90	3.86	**	1.32	1.98	
Number of observations	3,523			5,510		

*Source: NLSY79. Notes: *** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent. For (1) and (3), we include only those for whom income was reported or could be imputed. For (2) and (4), we include only those for whom the data were reported or imputed for 12 or more out of 20 years. Family and individual incomes are constructed from different sources and not always consistent. For family income, nonsurvey year incomes were imputed; for individual income, they were not imputed; consequently, the number of observations available for a year-to-year estimate of family income was fewer than the number available for yearly individual income. Each equation contains dummy indicators for missing variables for AFQT, parents' highest grade completed, and spouse's income (if applicable). For (1) and (2), spouse's income reported at age 41 is included. For other equations, average spousal income over the 20 year period is included. The table reports robust (Huber-White) standard errors.*

As income measures, we use annual net family income, as well as the respondent's own reported income. We examine family income, assuming that an individual's employment decision is motivated to maximize economic gains for his/her family unit. As such, family income can be viewed as one measure for an individual's success as an economic agent. The family income variable used here includes both earned income (e.g. wages from employer-based work and earnings from an individual's own business) as well as unearned income (e.g. interest, capital gains). We also examine how individual self-employment patterns are related to the respondent's own income. The own income measure used in our study includes wages and salaries as well as

income (earnings) from the respondent's own business or farm. The use of own income allows us to assess a more direct link between financial outcomes and respondents' own career pathway choices.³

Based on the OLS estimations of economic outcome measures, we find that family or own incomes are affected by years of self-employment, suggesting that an additional year of self-employment increases the level of income significantly. We also conduct separate analyses of family net worth (see Appendix C). Those analyses are consistent with the estimates shown in Exhibit 10. We find that an additional year of self-employment increases family net worth significantly, although an additional year of employer-based work does not. Overall, we find that total self-employment years are positively related to financial outcomes.

Comparing Early Self-Employment in the NLSY79 and NLSY97

In addition to investigating self-employment dynamics during early adult work life, we wish to document generational changes in self-employment patterns in early adult work life in the second half of the twentieth century. To this end, we use the NLSY79 along with its successor survey, the National Longitudinal Survey of Youth, 1997 Cohort (NLSY97). As noted earlier, the NLSY97 adopts the same sampling design as the NLSY79, and is designed specifically for generational comparisons with the NLSY79, as well as stand-alone analyses. The NLSY97 respondents are about 20 years younger than the NLSY79 respondents, with the younger cohort having birth years between 1980 and 1984. The NLSY97 respondents are still early in their labor force participation, but adequate data exist to compare their entry into the labor market with the older, NLSY79 respondents.

For comparison of the two surveys, we use subsets of data from each survey: the 1960-1962 birth years from the NLSY79 and the 1980-1982 birth years from the NLSY97. The use of these subsets allows us to make comparisons up to age 23. The sub-sample for the NLSY97 consists of 5,486 respondents. The sub-sample for the NLSY79, 4,125 respondents, may include respondents who were not included in the benchmark analysis sample (e.g. those who dropped out of the study after age 23).

Exhibit 11 compares employment statuses of NLSY79 and NLSY97 respondents at similar young ages. The taxonomy used in the earlier sections of this report is simplified into three categories: self-employed includes those self-employed for at least 100 hours in a given calendar year, including those dual-employed; employer-based only includes those who worked at least 100 hours for an employer in the calendar year with no self-employment, and not employed includes all others, including full-time students as well as individuals not in the labor force.

³ However, we cannot make a precise correlation between a particular job and its portion of total own income received due to the limitations of the data.

Exhibit 11
Percentages of NLSY97 (born 1980-1982) and NLSY79 (born 1960-1962) Respondents
by Self-Employment Status at Ages 18-23 (weighted)

Male									
<u>Age</u>	<u>Self-employed</u>			<u>Employer-based only</u>			<u>Not employed</u>		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	1.9	2.9	***	78.6	74.1	***	19.4	23.0	***
19	3.0	4.0	***	82.7	81.2	***	14.3	14.8	
20	4.7	4.6		79.4	81.6	***	15.9	13.7	***
21	6.7	5.2	***	78.0	84.4	***	15.3	10.4	***
22	8.1	6.1	***	76.3	83.5	***	15.6	10.4	***
23	8.2	7.7		77.7	83.0	***	14.0	9.3	***

Female									
<u>Age</u>	<u>Self-employed</u>			<u>Employer-based only</u>			<u>Not employed</u>		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	1.9	0.8	***	78.8	70.6	***	19.3	28.6	***
19	3.4	1.7	***	83.7	76.2	***	12.9	22.1	***
20	5.4	1.7	***	82.6	77.7	***	12.1	20.7	***
21	5.9	2.1	***	81.8	76.5	***	12.3	21.4	***
22	5.3	2.5	***	81.8	77.5	***	12.9	20.0	***
23	5.2	3.2	***	82.0	79.5	***	12.7	17.3	***

Sources: NLSY79 and NLSY97.

Notes: *** Significant at 1 percent; ** Significant at 5 percent.

From ages 18 to 23, we see that men in the two generations had similar rates of self-employment, but that the older men had higher rates of employer-based employment, implying lower rates of nonemployment for the older men at these young ages. Among women, the data indicate higher rates of self-employment and employer-based employment among the 1980s cohorts compared to the 1960s cohorts, and therefore lower rates of nonemployment. There are modest gender gaps in self-employment in both generations, with men more likely to be self-employed than women at these ages.

Exhibit 11 seems to suggest that if the younger generation follows the pattern of secular increases in self-employment with age, rates of self-employment will be quite high among the 1980s birth cohorts as they reach prime working age. To further explore the increase in self-employment from the NLSY97 to the NLSY79, we re-visit Exhibit 11, this time separating out Blacks, Hispanics, and Whites (and other non-Black, non-Hispanic individuals). Exhibits 12 and 13 show age-specific employment status rates for men and women respectively.

The first observation from Exhibits 12 and 13 is that Black and Hispanic rates of self-employment have indeed increased dramatically. Among women in the NLSY97, there is virtually no race/ethnicity difference in self-employment rates. There are some differences among men in the NLSY97, but both Blacks and Hispanics have made significant gains in catching up with White males. It is also worth noting that the self-employment rates of White males do not seem to have increased much from NLSY79 to NLSY97.

Exhibit 12
Percentages of NLSY97 (born 1980-1982) and NLSY79 (born 1960-1962) Males
by Race/Ethnicity and Self-Employment Status at Ages 18-23
 (weighted)

Black Male									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	1.9	1.5		68.3	58.3	***	29.7	40.3	***
19	2.8	2.1	**	74.5	69.5	***	22.8	28.4	***
20	5.5	2.6	***	72.3	76.5	***	22.2	20.8	
21	7.9	3.3	***	66.6	79.0	***	25.5	17.7	***
22	7.5	3.7	***	67.6	79.4	***	24.9	16.9	***
23	7.2	3.9	***	67.8	80.1	***	25.0	16.0	***
Hispanic Male									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	1.1	1.0		77.1	74.2	**	21.8	24.8	**
19	1.9	1.8		81.4	81.8		16.6	16.4	
20	3.5	2.0	***	82.2	80.1	**	14.3	17.8	***
21	4.6	2.1	***	81.8	80.2	**	13.6	17.7	***
22	6.8	3.8	***	77.2	85.9	***	15.9	10.2	***
23	6.4	4.1	***	80.3	85.8	***	13.2	10.1	***
White Male									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	2.1	3.3	**	81.3	77.0	***	16.7	19.6	***
19	3.2	4.5	**	84.8	83.5		11.9	12.0	
20	4.8	5.3		80.4	83.3	**	14.8	11.4	***
21	6.8	5.9		79.7	86.6	***	13.4	7.4	***
22	8.4	6.9	**	78.1	85.3	***	13.5	7.8	***
23	8.8	8.8		79.4	84.9	***	11.8	6.3	***

*Notes: NLSY97 data for 1446 White (non-Black, non-Hispanic) males, 584 Hispanic males, and 724 African-American males. NLSY79 data for 1040 White (non-Black, non-Hispanic) males, 395 Hispanic males, and 651 African-American males. *** Significant at 1 percent; ** Significant at 5 percent.*

Among males, there appear to be differences by race and ethnicity in employment status, but the patterns are not uniform. For employer-based employment, White and Hispanic men have comparable rates, while Black men have a considerably lower rate of employer-based employment. In self-employment, the relationship flips, with Hispanic men having the lowest rates of self-employment, exceeded somewhat by Black men and even more so by White men. Among females, we see race/ethnicity differences in employer-based employment and nonemployment, but we see essentially no differences across race and ethnicity in self-employment rates.

Exhibit 13
Percentages of NLSY97 (born 1980-1982) and NLSY79 (born 1960-1962) Females
by Race/Ethnicity and Self-Employment Status at Ages 18-23
 (weighted)

Black Female									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	2.2	0.3	***	68.5	52.7	***	29.3	47.0	***
19	5.0	0.4	***	71.7	61.3	***	23.3	38.3	***
20	6.1	1.1	***	73.2	63.7	***	20.8	35.1	***
21	5.2	1.2	***	74.5	64.9	***	20.3	33.9	***
22	5.0	1.1	***	74.7	68.2	***	20.2	30.7	***
23	6.0	1.8	***	76.2	68.6	***	17.9	29.7	***

Hispanic Female									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	2.1	0.4	***	70.6	59.5	***	27.3	40.2	***
19	3.1	1.1	***	81.3	68.4	***	15.6	30.5	***
20	5.5	0.6	***	79.7	72.9	***	14.8	26.5	***
21	6.5	0.7	***	76.5	71.4	***	16.9	27.9	***
22	5.4	0.5	***	77.2	71.2	***	17.4	28.3	***
23	5.6	2.5	***	78.9	69.9	***	15.4	27.6	***

White Female									
<u>Age</u>	Self-employed			Employer-based only			Not employed		
	<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>		<u>NLSY97</u>	<u>NLSY79</u>	
18	1.7	0.9	**	82.6	75.1	***	15.6	23.9	***
19	3.1	2.0	**	86.8	80.0	***	10.0	18.0	***
20	5.2	1.9	***	85.2	81.2	***	9.6	16.9	***
21	6.0	2.4	***	84.4	79.8	***	9.6	17.8	***
22	5.4	3.0	***	84.2	80.7	***	10.4	16.3	***
23	5.0	3.6	***	83.9	83.6		11.1	12.8	*

Notes: NLSY97 data for 1388 White (non-Black, non-Hispanic) females, 557 Hispanic females, and 707 African-American females. NLSY79 data for 1031 White (non-Black, non-Hispanic) females, 387 Hispanic females, and 621 African-American females.

**** Significant at 1 percent; ** Significant at 5 percent.*

Comparing men and women within race/ethnicity groups, we see that Hispanic men and women are approximately equally likely to be self-employed, but Hispanic men are very slightly more likely to have employer-based employment. Among Whites, White males quickly increase self-employment well above White women's rates, but White women are more likely to have employer-based employment. Both Black men and Black women have high rates of nonemployment, but Black males have very slightly higher self-employment, while Black women have higher employer-based employment. Evidently, the mechanisms that associate race and ethnicity with employment rates work differently for self-employment and for employer-based employment.

The similarities and differences in early self-employment experiences between these two generations of American workers will continue to be a topic of great policy and economic import over the coming decades. This analysis has just begun to investigate what those similarities and differences might be at the point of labor market entry. Our notable findings include:

- The younger NLSY97 cohort has much higher self-employment rates than the older NLSY79 cohort had at similar ages. This increase is driven largely by increases in self-employment by Black and Hispanic workers, and to a lesser extent by female workers.
- The mechanisms through which demographics affect employer-based employment rates seem to work somewhat differently for self-employment rates. Such characteristics as race, gender, and ethnicity do not always exhibit the same patterns for self-employment and employer-based employment.

Conclusion

This study conceptualizes self-employment as part of a career trajectory, emphasizing a life cycle perspective in the analysis of self-employment patterns. Using the NLSY79, a data set that represents Americans born between 1957 and 1963 and living in the U.S. in 1979, we begin by confirming demographic patterns in self-employment that have been documented in other studies. Through three different multivariate analysis approaches, we then investigate the determinants and outcomes of different patterns of self-employment in early career, especially the effect of early self-employment on later self-employment. Across our analyses, we find evidence that the determinants and outcomes of self-employment do vary over the life cycle. We find that early self-employment is associated with increased likelihood of self-employment later in one's career. We also find that increased financial resources have a modest contribution to the probability of self-employment. Regarding the contribution of self-employment to income, our results indicate that additional years of self-employment increase income levels significantly.

We close with some comparisons of this group to a group approximately 20 years younger. NLSY97 data indicate that youths born from 1980 to 1982 are experiencing higher levels of self-employment compared to the experiences of individuals born 20 years earlier. The increase is primarily among women and minorities, with White men's participation rates rather similar across the two groups.

Our analyses indicate strong support for continued modeling of self-employment in a trajectory context rather than as isolated employment events. They also suggest that self-employment patterns observed in earlier generations may not persist with younger Americans, and that the NLSYs will be a valuable resource in monitoring these shifts in self-employment dynamics.

References

The list of articles below includes only those referenced in the text above. Additional references used in the study are provided in Appendix A.

- Bates, Timothy. 1997. *Race, Self-Employment, and Upward Mobility: An Illusive American Dream*. Washington, D.C: Woodrow Wilson Center Press; Baltimore: John Hopkins University Press.
- Bates, Timothy and Lisa Servon. 2000. "Viewing Self Employment as a Response to Lack of Suitable Opportunities for Wage Work." *National Journal of Sociology* 12(2), 23-55
- Blanchflower, David G. and Andrew J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1), 26-60
- Blanchflower, David G. and Andrew J. Oswald. 2007. "What Makes a Young Entrepreneur?" IZA Discussion Paper #3139.
- Blanchflower, David G., Phillip B. Levine, and David J. Zimmerman. 2003. "Discrimination in the Small-Business Credit Market." *The Review of Economics and Statistics* 85(4), 930-943
- Budig, Michelle. 2006. "Intersections on the Road to Self-Employment: Gender, Family, and Occupational Class." *Social Forces* 84(4), 2223-2239
- Carr, Deborah. 1996. "Two Paths to Self-Employment?" *Work and Occupations* 23(1), 26-53
- Cavalluzzo, Ken and John Wolken. 2005. "Small Business Loan Turndowns, Personal Wealth, and Discrimination." *Journal of Business* 78(6), 2153-2178
- Colombier, Nathalie and David Masclet. 2008. "Intergenerational Correlation in Self Employment: Some Further Evidence from French ECHP Data." *Small Business Economics* 30(4), 423-437
- Dickson, Pat H., George T. Solomon, and K. Mark Weaver. 2008. "Entrepreneurial Selection and Success: Does Education Matter?" *Journal of Small Business and Enterprise Development* 15(2), 239-258
- Dunn, Thomas and Douglas Holtz-Eakin. 2000. "Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links." *Journal of Labor Economics* 18(2), 282-305
- Dunn, Thomas and Douglas Holtz-Eakin. 1995. "Capital Market Constraints, Parental Wealth, and the Transition to Self-Employment Among Men and Women." Center for Policy Research, Syracuse University.
- Evans, David S. and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy* 97(4), 808-827
- Evans, David S. and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79(3), 519-535
- Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment." *Journal of Labor Economics* 17(1), 80-108

- Fairlie, Robert W. 2005A. "Technology and Entrepreneurship: A Cross-Industry Analysis of Access to Computers and Self-Employment." The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Fairlie, Robert W. 2005B. "Self-Employment, Entrepreneurship, and the NLSY79." *Monthly Labor Review* February 2005, 40-47
- Fairlie, Robert W. and Alicia M. Robb. 2007A. "Why Are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital." *Journal of Labor Economics* 25(2), 289-323
- Fairlie, Robert W., and Alicia Robb. 2007B. "Families, Human Capital, and Small Business: Evidence from the Characteristics of Business Owners Survey." *Industrial and Labor Relations Review* 60(2), 225-245
- Georgellis, Yannis, John G. Sessions, and Nikolaos Tsitsianis. 2005. "Windfalls, Wealth, and the Transition to Self-Employment." *Small Business Economics* 25(5), 407-428
- Henley, Andrew. 2004. "Self-Employment Status: The Role of State Dependence and Initial Circumstances." *Small Business Economics* 22(1), 67-82
- Henley, Andrew. 2005. "Job Creation by the Self-Employed: The Roles of Entrepreneurial and Financial Capital." *Small Business Economics* 25(2), 175-196
- Holtz-Eakin, Douglas. 1994. "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." *Journal of Political Economy* 102(1), 53-75
- Holtz-Eakin, Douglas, David Joulfaian, and Harvey Rosen. 1994. "Entrepreneurial Decisions and Liquidity Constraints." *Rand Journal of Economics* 23(2), 334-347
- Hundley, Greg. 2006. "Family Background and the Propensity for Self-Employment." *Industrial Relations: A Journal of Economy and Society* 45(3), 377-392
- Karoly, Lynn A. and Julie Zissimopoulos. 2004. "Self-Employment and the 50+ Population." Washington, DC: AARP Public Policy Institute.
- Kepler, Erin and Scott Shane. 2007. "Are Male and Female Entrepreneurs Really That Different?" The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Kim, Phillip, Howard Aldrich, and Lisa Keister. 2006. "Access (Not) Denied: The Impact of Financial, Human, and Cultural Capital on Entrepreneurial Entry in the United States." *Small Business Economics* 27(1), 5-22
- Lowrey, Ying. 2007. "Minorities in Business: A Demographic Review of Minority Business Ownership." U.S. Small Business Administration, Office of Advocacy.
- Moutray, Chad. 2007. "Educational Attainment and Other Characteristics of the Self-Employed: An Examination Using Data from the Panel Study of Income Dynamics." U.S. Small Business Administration, Office of Advocacy.
- Rissman, Ellen R. 2006. "The Self-Employment Duration of Younger Men over the Business Cycle." *Economic Perspectives* Q III, 14-27
- Salazar, Maritza. 2007. "The Effect of Wealth and Race on Start-up Rates." U.S. Small Business Administration, Office of Advocacy.

- Schiller, Bradley R. and Philip E Crewson. 1997. "Entrepreneurial Origins: A Longitudinal Inquiry." *Economic Inquiry* 35(3), 523-531
- Taniguchi, Hiromi. 2002. "Determinants of Women's Entry into Self-Employment." *Social Science Quarterly* 83(3), 875-893
- U.S. Small Business Administration, Office of Advocacy. 2007. "The Small Business Economy for Data Year 2006: A Report to the President."
- U.S. Small Business Administration, Office of Advocacy. 2006. "The Small Business Economy for Data Year 2005: A Report to the President."
- Wennberg, Karl J., Timothy Folta, and Frederic Delmar. 2007. "A Real Options Model of Stepwise Entry into Self-Employment." The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Williams, Donald R. 2004. "Youth Self Employment: Its Nature and Consequences." *Small Business Economics* 23(4), 323-336

Appendices

Appendix A: Study Design and Data

Appendix B: Review of Existing Literature

Appendix C: Multivariate Analyses

Appendix A Study Design and Data

Research Purpose and Questions

In our research on the survival and growth of small businesses, we regard an individual as an agent of enterprise, an “entrepreneur,” and equate self-employment to an incident of entrepreneurship or business ownership. We thus focus on individuals, rather than firms, as our unit of analysis, and investigate the dynamics of self-employment at the person level, utilizing the National Longitudinal Surveys of Youth (NLSY), which provides longitudinal data on the individual’s yearly employment status. The goals of our study are twofold: (a) providing new empirical findings regarding the engagement in self-employment that underpin individual entrepreneurship, particularly during early adult work life; and (b) documenting generational changes in self-employment patterns in early adult work life between young adults in the early 1980s and those entering their twenties in the early 2000s. As a particular interest of the study, we investigate the impact of gender, race, financial resources, self-employment experiences at early age, as they relate to the extent to which individuals engage themselves in self-employment.

The study is exploratory in nature. Its purpose is to broaden the existing empirical knowledge regarding self-employment as an entrepreneurial activity,⁴ with the specific goals described above. Our empirical inquiry is guided by a set of research questions arising from a series of hypotheses about self-employment. The hypotheses, outlined below, are not intended to reflect any particular theoretical model of entrepreneurial behaviors and choices to be tested; rather, they are used simply as tools to organize the many empirical questions. Below we discuss the key testable hypotheses that form our research questions and guide our exploratory investigation.

Hypothesis #1: Individual characteristics vary by the extent to which individuals are engaged in self-employment over their early work life.

We expect that the observable profile of an average person varies by the level of his/her engagement in self-employment over the year. As implied by previous research, individuals are likely to approach self-employment differently according to their preferences, resources, and life circumstances. We hypothesize that this will lead to disparate pathways into and out of self-employment according to individuals’ characteristics, resulting in the varying extent to which individuals are self-employed (e.g. total years of self-employment).

Hypothesis #2: The extent to which individuals are engaged in self-employment over time or at a given point in time varies by individual and family characteristics and early labor-market experience.

We extend Hypothesis #1 by asking the question of who is more likely to engage in self-employment. To assess the extent of self-employment experiences, we consider two alternative

⁴ See Appendix B for a review of the relevant literature.

measures for this study: (a) years of self-employment and (b) the probability of self-employment at a given period. Total years of engagement in self-employment represents an eventual outcome over time. The point-in-time probability of self-employment, on the other hand, does not provide a long-term view, but provides insight into the choice of employment type in each period along the way

Our hypothesis is that certain background characteristics or early labor market experiences will predict the total number of years that individuals are self-employed over the ensuing 20-year period, as well as the probability that individuals will be self-employed at a given time. For example, we expect that certain demographic characteristics such as at gender and race will affect an individual's likelihood of self-employment, based on the previous literature suggesting less involvement in entrepreneurial activities among women and minorities. In addition to background characteristics, we also expect that contemporaneous factors that measure changes in life circumstances will predict the point-in-time probability of self-employment. In exploring the effect of contemporaneous factors, we pay particular attention to early age exposure to self-employment and financial resources.

Hypothesis #3: Economic outcomes, measured in terms of income and net worth, vary by the extent to which individuals are self-employed.

Having analyzed how background characteristics relate to self-employment, we shift the focus to the economic outcomes resulting from self-employment choices over time. Since we are interested in assessing the performance of self-employed individuals over time, we construct outcome measures that summarize economic experience across all observed years of work life. We consider a range of outcome measures, including: family income, individuals' own income, and family net worth. Our hypotheses is that economic outcomes will vary by how individuals experience self-employment over time, as measured by total years of self-employment.

Hypothesis 4: Self-employment experience in the early career has changed from the early 1980s to the early 2000s.

We expect that the previous generation of entrepreneurs and the new generation of entrepreneurs will look different for a number of reasons, including technological changes, demographic shifts, and change in attitudes toward and preferences for work. Taking advantage of the data available to the study, we describe the characteristics of a younger generation and their early self-employment experiences and then compare these findings with an older generation when they were at a comparable age. We expect that the two cohorts will be different in how they pursue self-employment at comparable ages.

Testing of the hypotheses listed above forms the basis of the research questions to be investigated in our study. By addressing these questions, the study aims to provide new insight into small business survival and growth in terms of self-employment trajectories.

Research Methods

This study involves several methodological steps to explore the research questions arising from the hypotheses discussed above. We briefly review these steps below.

Univariate Descriptive Analysis.

We primarily use simple univariate descriptive analyses to explore Hypothesis #1, and Hypothesis #4. This amounts to tabulating descriptive statistics (e.g. mean, percentage distribution) for the variables that we expect to be related to being self-employed, one variable at a time. In doing so, we do not control for other variables. For example, we tabulate and compare the mean characteristics by the extent to which individuals are self-employed over time. Univariate descriptive analyses are also used to compare self-employment patterns between the two generations of workers.

In these descriptive analyses, we therefore do not limit ourselves to examining types of variables that have specific formal theoretical bases that link them to self-employment (e.g. human capital and financial assets). As discussed below, the data that we use in the study offer a wide range of possible background measures to be analyzed. We thus can look at variables that may not have been closely reviewed by previous self-employment research (such as an ability measure). We also examine those variables that have commonly been used in previous research (e.g. own education level, parents' education level, family structure) and those factors known or suspected to be critical for self-employment (e.g. health plan coverage and financial resources).

Multivariate Analyses

Using several different estimation methods, we conduct multivariate analyses to investigate Hypotheses #2 and #3. First, we use the tobit method to model the extent of self-employment over time, measured as total years of self-employment. The goal of the tobit model is to gauge the effects of various background characteristics. Second, instead of focusing on respondents' employment experiences over all years in the data, we model the decision to engage in self-employment in each given year. We estimate this model by applying the generalized estimating equation (GEE) method, which allows us to adjust for within-person correlations that arise from repeated observations of individuals. Again, the main goal of the GEE model is to estimate the effect of various background characteristics. Third, we model economic outcomes to investigate whether and how they vary by the extent of self-employment. As dependent variables, we use family net income and net worth, both measured in alternative forms (level, growth rate, and coefficient of variation). The model is estimated using the ordinary least squares (OLS) method. The goal is to estimate the effects of various background variables on financial outcome measures. In all of these multivariate analyses, the explanatory variables included are similar to those used in basic univariate descriptive analyses. More details on model specifications and estimation methods are presented in Appendix F.

Data Description

This study uses the two National Longitudinal Surveys of Youth, the 1979 Cohort (NLSY79) and the 1997 Cohort (NLSY97). With unparalleled breadth and depth of information on economic activity as well as personal and family backgrounds, these surveys provide an outstanding opportunity for studying self-employment and entrepreneurship.⁵ Indeed, the strength and unique contribution of this study is our extensive utilization of these excellent yet complex longitudinal data, especially the underutilized NLSY97, to generate new empirical evidence in small business research.

Background of NLSY Studies

The NLSY79 and NLSY97 are two in a series of studies sponsored by the National Longitudinal Surveys Program of the Bureau of Labor Statistics, U.S. Department of Labor. Each of the two studies focuses on a cluster of birth cohorts, following them from adolescence into adulthood in order to better understand the dynamics of labor force experience, especially as affected by early life experiences. The studies have a parallel design that permits multi-faceted comparison between two American generations born in the second half of the twentieth century.

The NLSY79 is a longitudinal survey of individuals born between 1957 and 1964 and living in American households in 1979. The NLSY79 originally consisted of 12,686 young women and men interviewed annually from 1979 through 1994, and interviewed biennially since then. This study makes use of data from the first 22 rounds (26 years) of the NLSY79, from 1979 to 2005. The NLSY79 sample is nationally representative, including oversamples of African Americans (Blacks) and Hispanics. Respondents were recruited from households using a multi-stage, stratified random sample. Interviewers screened households for at least one youth or young adult aged 14-21 by December 31, 1978, who lived as a usual resident of the household.

The NLSY79 survey includes three main sub-samples. The first sub-sample includes a cross-section of 6,111 youths born between January 1, 1957 and December 31, 1964 that represents noninstitutionalized civilian youths. The second sub-sample includes a supplemental over-sample of 5,295 Hispanics, Blacks, and economically disadvantaged non-Hispanic/non-Blacks. The final sub-sample consists of 1,280 youth enlisted in the military. However, 1,079 respondents from the military subsample were eliminated after 1984 due to funding limitations, as was the entire supplemental sample of 1,643 economically disadvantaged non-Hispanics/non-Blacks after the 1990 survey. Because we are unable to follow the labor force progress of these eliminated groups across waves, we have removed them from our sample and this study focuses on the remaining 9,964 respondents. The 201 individuals of the military sub-sample retained in the NLSY79 included 51 persons who would have been integrated into the original study if it were not limited to civilian youth, as well as an additional 150 individuals included to reduce variability in the retained sub-sample.

The survey response rate is at least 90 percent until 1993, after which a steady decline brings the response rate to 76.8 percent in 2006. However, many respondents who are unavailable during

⁵ Fairlie, Robert W. 2005B. "Self-Employment, Entrepreneurship, and the NLSY79," *Monthly Labor Review*, February, 40-47.

one or several survey waves are thoroughly searched for and re-recruited when possible, permitting researchers to reconstruct more complete information on respondents because respondents are asked to provide answers to questions for years in which they were unavailable.

The NLSY97 is a longitudinal survey of individuals born between 1980 and 1984 and living in an American household in 1997. This relatively young sample consists of 8,984 individuals who have been interviewed annually since 1997. This study makes use of the data from the first 11 rounds (11 years) of the NLSY97. The NLSY97 sample is nationally representative, including oversamples of African Americans and Hispanics. In-person interviews conducted using a computer-assisted personal interviewing (CAPI) system account for the majority of the interviews in each survey round (88 percent or higher). Retention rates hovered above 90 percent in the first 3 rounds, falling to 86.3 percent in Round 7. NLSY97 respondents were recruited from two independently selected, stratified multistage area probability samples for the cross-sectional and supplemental samples. Interviewers screened households and included those with at least one youth, aged 12 – 16 by December 31, 1996, who usually resided in the household.

Both surveys are notable for their tremendous breadth of topical coverage. Because they focus on labor force experiences and their determinants, the surveys include data on virtually every domain of life, including education, relationship and family formation, migration/residence, criminal activity, mental and physical health, income and assets, participation in government programs, and attitudes and expectations. In addition, both surveys employ an event history design in which events that occur in the time between interviews are captured as well as the current status at the time of the interview. This approach, more extensive in the NLSY97 than in the NLSY79, yields data that provide tremendous detail on the relative timing of different life activities, for example, the interweaving of marriage dates with when an individual transitioned from employer-based employment to self-employment.

NLSY97 also includes a parent interview, providing considerable detail on the youth's early life experiences and family background. For example, the quality and extent of data on assets and income of the youth's household are much greater than could be achieved through just a youth interview. Unfortunately, data on the self-employment behavior of parents are not available. In addition, the NLSY97 made an unprecedented effort to collect employment behavior of individuals aged 12-16, ages at which one is not technically in the labor force. These data, generally pertaining to "freelance" labor force activity (e.g. babysitting, lawn mowing, etc.) may shed light on another domain of pre-self-employment activity that is predictive of entry into and success in self-employment.

Both the NLSYs include considerable data on self-employment, including wages earned, duration of self-employment spells, concurrent employment in employer-based jobs, industry and occupation of own business, incorporation status of business, and usual hours worked per week at business. Indeed, more than two dozen papers have already been written on self-employment and entrepreneurship using the NLSY79 or its predecessors, the Original Cohorts of Longitudinal Studies. (In Appendix B, we review the most relevant examples of these previous studies.)

Merits of Using NLSY Data

This study extends the prior literature based on NLSY studies (as well as the extant literature using other data sets) in a few important ways. In addition, this study extends the prior studies based on NLSY studies in two important ways.

First, we adopt a more intensive and thorough use of available information in the data to identify individuals who are self-employed than the approaches taken in the prior studies. Unlike most previous research, which only examines the current or most recent so-called CPS job,⁶ our study utilizes information collected on up to five jobs during the survey year, and includes data collected retrospectively from individuals who may have missed an occasional survey round. We are, however, more conservative in one way: we code workers as self-employed or dually employed in a calendar year only if they worked at least 100 hours of self-employment in that year. It is possible that more traditional definitions of self-employment would include some workers with very low hours of work, while our definition would fail to classify them as self-employed. As we demonstrate in the subsection below, our more intensive use of the NLSY79 data has substantial impact on the estimates of self-employment among NLSY79 respondents.

Second, we use the previously underutilized NLSY97, in combination with the NLSY79. More than two dozen papers have been written on self-employment and entrepreneurship using the NLSY79 or its predecessors in the National Longitudinal Surveys Program, the Original Cohorts; however, to our knowledge, few studies on self-employment have been conducted using NLSY97. The use of the NLSY97, a much younger cohort just beginning their labor force participation, can provide new insight into how younger generations today are engaged in self-employed activities. As we discuss below, this younger generation seems already to be more involved in self-employment than the NLSY79 generation was at similar ages. The bi-generational comparisons permitted by using these two data sets in tandem shed light both on the trajectories followed by the older generation in their first 20 or so years of labor force experience, as well as on the differences and potential implications of self-employment patterns among the younger generation that is just entering the labor force.

In contrast to such rich sources of data on business ownership and entrepreneurship as the Survey of Business Owners and Self-Employed Persons (part of the Economic Census) and its predecessor, the 1992 Characteristics of Business Owners, or the Panel Study of Entrepreneurial Dynamics, the relative strength of the NLSYs is in contextualizing the self-employment experience within the range of life cycle behaviors, such as education, family formation, assets, and health status. The NLSYs' extensive detail in these domains—all known or hypothesized to be relevant to entry into and duration of self-employment—is the core strength of this study.

⁶ NLSYs ask a set of questions that will allow researchers to construct labor market indicators that are equivalent to those reported by the Current Population Survey.

Identification of the Self-Employed Using NLSY79

As we mention above, our study makes more intensive use of the NLSY79 employment data than have prior studies. Most importantly, we use up to five jobs reported at each interview rather than only the current or most recent.

Exhibit A.1
Comparison of NLSY79 Self-Employment Rates
Using Current/Most Recent Job Only versus Using Up to Five Jobs
(Weighted)

Percent Currently Self-Employed

<u>Year</u>	<u>Current/Most Recent job</u>	<u>Up to 5 Jobs</u>
1979	1.59	2.28
1980	1.16	2.34
1981	1.32	3.1
1982	2.01	3.44
1983	2.34	4.08
1984	2.75	4.87
1985	3.25	5.62
1986	3.6	6.56
1987	4.29	7.44
1988	4.84	7.96
1989	5.08	8.45
1990	5.27	8.73
1991	5.95	9.24
1992	6.23	9.37
1993	6.39	9.08
1994	5.65	8.93
1996	6.02	8.78
1998	5.84	7.7
2000	5.75	6.67
2002	6.38	8.83
2004	7.04	9.7

Source: NLSY79

In Exhibit A.1, we compare NLSY79 self-employment rates using the methodology used in previous studies and the methodology used in this study. In prior studies, researchers used only the current or most recent job. Where there is more than one current job, the job with the most hours is selected. In contrast, in this study, we use all five of the jobs recorded in the data, and all

self-employment is reflected here, whether dual-employment or self-employment alone. In constructing our employment status variables, we also make use of all retrospective data reported for individuals who have missed one or more interviews. In later years, where attrition has increased, the use of these retrospective data can reduce nonresponse substantially.

Reviewing the data in Exhibit A.1, we see that our measure generates considerably higher rates of self-employment than the measure previously used, ranging from 16 percent higher in 2000 to more than twice as much in 1980 and 1981. Increases are expected, since counting more jobs gives individuals more opportunities to report self-employment. There are a few implications for this definition change. First, we have a more complete definition of self-employment. Second, we are able to include more individuals in our analysis of self-employment, thus increasing sample sizes and analytical power. Third, because we are including jobs that may have been secondary or for shorter durations, we may be incorporating more workers whose self-employment is either casual or unsuccessful.

In Exhibit A.2, we examine selected demographic characteristics of workers who are classified as self-employed according to the two definitions discussed above. In general, we see that the two groups of self-employed workers resemble each other reasonably closely. In 1984, when these individuals were aged 20-27, our broader definition of self-employment shows slightly more Blacks and Hispanics self-employed than the more restrictive definition. Because of the possibility of dual employment, slightly more of the self-employed workers in our classification have ever had employer-based employment. Our definition also shows more self-employed workers with at least some college, whereas in the more restrictive definition using current or most recent job, a larger fraction of self-employed workers have a high school diploma or less. When using up to five jobs' worth of data, self-employed workers are less likely to be married currently, and they have about 10 percent lower family income than those self-employed in their current or most recent job.

Similar patterns emerge when we look at individuals classified as self-employed in 2004. By that time, almost all self-employed workers had had employer-based work at some point in their lives. We see that counting up to five jobs increases the fraction of male self-employed workers somewhat, and lowers the fraction who are Hispanic. Again, in the more inclusive definition of self-employment, fewer workers have 12 or fewer years of completed education. Workers self-employed based on up to five jobs are slightly less likely to be currently married at the later ages, and they have about 5 percent higher average total net family income than workers self-employed in their current or most recent job.

The two exhibits together suggest that although the fraction of self-employed increases substantially with our classification relative to more traditional, restrictive ones, the characteristics of those classified as self-employed do not change dramatically.

Exhibit A.2
Characteristics of NLSY79 Individuals Self-Employed
According to Current/Most Recent Job vs. Up to Five Jobs
(Weighted)

	1984		2004	
	Current/Most Recent job	Up to 5 jobs	Current/Most Recent job	Up to 5 jobs
Percent dually employed	n/a	3.2	n/a	3.3
Percent ever employer-based employed	91.7	94.3	98.5	98.5
Percent male	68.3	68.0	61.9	63.6
Percent Black	16.8	18.4	25.4	26.0
Percent Hispanic	14.6	15.5	18.7	17.1
Completed 11 or less years of schooling	23.4	21.9	9.3	8.3
Completed 12 years of schooling	44.5	42.9	45.2	40.0
Completed 13-15 years of schooling	20.8	22.5	22.0	22.3
Completed 16 years of schooling	9.5	7.4	13.6	12.7
Completed 17 or more years of schooling	1.8	1.4	10.0	11.2
Percent married currently	43.1	38.1	61.1	59.6
Average Total Net Family Income	\$ 24,349	\$ 22,360	\$ 78,745	\$ 82,199

Source: NLSY79.

Key Variables from NLSY79

For the current study, we use the NLSY79 as our main source of information, as it provides the most extensive longitudinal data. The NLSY97 is used as a secondary data set, primarily to create a comparison group for the inter-generational analyses. As such, for the remainder of this appendix, we focus our data discussions on the NLSY79, our main data source. (Discussion of the NLSY97 data is included in Appendix F.) Generally, whenever the NLSY97 is being used, our basic approach is to create variables comparable to those in the NLSY79.

We utilize the rich information available from the NLSY79 to conduct our base analysis represented by Hypotheses #1 through #3. Critical to this analytical process is the preparation of an employment status indicator, which is used to identify who is self-employed, employer-employed, or not employed. As demonstrated above, how we construct such an indicator can make differences in our ability to determine who should be considered self-employed (or employer-employed). Below, we present the classifications we have used to construct an

employment status indicator from the NLSY79, and discuss the types of other covariates (e.g. background variables) that we have selected to use in this study.

Employment Status Indicator

Self-Employment

Self-employed persons in the NLSY79 are identified by responses to the class-of-worker question asked in each survey round and available for up to five jobs each year. Unlike most other research analyzing the NLSY79 and self-employed persons, which focuses on the CPS job for each year, this study takes advantage of the information provided on up to five jobs held during each year. The NLSY79 classifies respondents as self-employed in a given job if:

. . . he or she owned at least 50 percent of the business, was the chief executive officer or principal managing partner of the business, or was supposed to file a form SE for federal income taxes. Respondents are also classified as self-employed if they identify themselves as independent contractors, independent consultants, or freelancers. This information is available for an employed respondent's current/most recent job, as well as for each job held since last interview in which he or she worked for more than 10/20 hours a week and for more than nine weeks since last interview. Prior to 1988, information was collected for jobs worked more than 20 hours a week. After 1988, the number of hours was reduced to 10" (NLSY79 User's Guide).

This study categorizes a person as ever self-employed if he or she has worked at least 100 hours in self-employment during a single calendar year from 1979 through 2005. A respondent is categorized as self-employed in a given year if he or she worked at least 100 self-employment hours in that year. These 100 hours could conceivably be worked across multiple jobs within the year, although that is rarely the case.

Employer-based Employment

Employer-based employment refers to employed persons categorized in one of the other available class-of-worker classifications. The classifications refer to private companies, the government, nonprofit companies and household farms or businesses that the respondent does not own and from which the respondent receives no wages. Active-duty military personnel are also classified as having employer-based employment. These respondents are identified as those responding to the question regarding the number of weeks of active service during a calendar year.

Yearly Employment Status

The main sequence analysis examines yearly employment status variables created for this study. The employment status variables were created in several steps. These steps capture the most complete information for each survey respondent, many of whom missed interviews in some survey rounds but are generally surveyed again and often provide information covering the time

period(s) of the missed interview(s). These later rounds are used to fill in gaps from noninterviews.

The first step in creating the employment status variable involved creating the number of weeks that each respondent was employed for each individual year between 1979 and 2005 by examining start and stop dates for each job in each year. Second, the number of hours worked in each job each year was created by multiplying the number of weeks worked in that job that year with the usual number of hours worked per week in that job. Then, the number of hours worked during the year for self-employed jobs are added together, as are the number of hours worked for the year for employer-based jobs. Respondents who work at least 100 hours during the year at a self-employed job and have fewer than 100 hours of employer-based work are classified as self-employed only. Persons who work at least 100 hours during the year at an employer-based job and worked fewer than 100 hours at self-employed jobs are classified as employer-based only. Persons who work at least 100 hours in self-employed jobs and at least 100 hours in employer-based jobs during a calendar year are considered to be dually employed during that year.

The two categories of nonemployment focus on those who were employed for less than 100 hours during the survey round. Those persons who were unemployed and did not attend school, such as a high school, college or university, graduate program, or GED program are coded as nonemployed and not in school. Respondents who were not employed but attended some kind of high school, college or university, graduate program or GED program during the survey year are coded as not employed but in school in the employment status variable.

All persons with missing data for an interview that was not recaptured in a later round, and those who refused to answer, answered “Don’t know” or were skipped for invalid reasons, are classified as “unknown” in the employment status category.

Covariates

The NLSYs provide rich, detailed information about study participants, enabling the creation of demographic profiles of both ever self-employed respondents and respondents who have never been self-employed. Background characteristics and yearly demographic variables are examined to evaluate differences and similarities between the ever self-employed and never self-employed populations. Background variables include country of birth, country of residence at age 14, South/non-South regional residence at age 14, household structure at age 14, highest grade completed by mother and father by 1979, race/ethnicity, gender, and a variable identifying persons born outside of the U.S. with U.S. parentage, making them citizens at birth. Most variables are self-reported, although the race/ethnicity variable is chosen by the interviewer and consists of three choices: Black, Hispanic, and non-Black/non-Hispanic.

Other demographic variables include characteristics that often change over time, especially during the young adult years when many life transitions take place, such as completing education, entering the workforce, marrying, and having children. The following variables were collected in each wave: age of R (respondent) at interview date; marital status; enrollment status as of May 1 survey year (revised version); highest grade completed as of survey year (revised version); number of children ever born; urban/rural residence; standard metropolitan statistical

area (SMSA); regional residence; total net family income (truncated); family size; family poverty status; and health limitations on work.

An additional data item is the Armed Forces Qualifying Test (AFQT) normed percentile score, a measure of ability derived from the Armed Services Vocational Aptitude Battery (ASVAB). This is a robust test that is often used as measure of ability, and whose psychometric properties are well known.

NLSY79 Analysis Sample

The original NLSY79 sample consisted of 12,686 youths. As mentioned earlier, the original sample included both a civilian sample and a military sample, and included an oversample of Black, Hispanic, and economically disadvantaged non-Black non-Hispanic civilians. The military sample was reduced in size by roughly half after 1984. Accordingly, we limited our study sample to 9,964 by excluding those respondents in the military subsample who were later dropped from the study. Of the total NLSY97 cohort of 9,964 respondents who ranged in age from 15 to 22 in 1979, we could potentially observe the employment status from age 22 to 41 for 8,860 respondents based on the currently available public use files, excluding the youngest subgroup for whom we could not observe 20 years. Of those 8,860, 7,482 individuals reported employment information that allowed us to construct the yearly employment status for each year corresponding to ages 22 through 41.⁷ We refer to this cohort as the benchmark analysis sample.⁸

In investigating the question of what self-employment paths individuals take during the first two decades of adult life, we focus on those who have experienced self-employment at least once during the 20-year period. Of the benchmark sample of 7,482, 2,394 reported having been self-employed in one or more years. We refer to this subgroup as the ever self-employed sample, and conduct the sequence analyses for these 2,394 respondents to develop self-employment trajectory types.⁹

The NLSY79 cohort is designed to be nationally representative of men and women who were 14 to 22 years old and living in the U.S. in 1979, who are part of the later group of baby boomers, born in the 1950s and 1960s. As such, our analysis of employment trajectories, with appropriate weights, is used to illustrate representative experiences among the target population of late baby

⁷ Using all available surveys through the 2006 wave, we were able to construct yearly employment status information reliably through 2004. For 2005 and 2006, we were not able to construct reliable employment status variables.

⁸ For comparison between the NLSY97 and the NLSY79, we use three birth cohorts of data from each survey: the 1960-1962 birth years from the NLSY79 and the 1980-1982 birth years from the NLSY97. Using this narrower age range allows us to make comparisons up to age 23, which is not yet possible for the youngest of the NLSY97 respondents, born through 1984. The NLSY97 analysis sample includes 5,486 respondents, excluding fewer than 50 in the eligible age range who are of mixed race. A comparison group from NLSY79 consists of 4,125 respondents, which may include respondents who were not included in the benchmark analysis sample (e.g. those who dropped out of the study after age 23).

⁹ We have also examined a smaller subsample of respondents for whom we can observe employment status from ages 22 to 45. Findings from this smaller sample are consistent with those reported here, based on the benchmark sample consisting of those whom we can observe from age 22 to 41.

Exhibit A.3
Distribution of Yearly Employment Status by Age
Benchmark Sample (N=7,482)
(Weighted)

Age	Exclusively Self- employed (percent)	Dual-employed (percent)	Exclusively Employer- employed (percent)	Not Working and Not in School (percent)	Not Working and In School (percent)
22	1.4	3.0	81.0	10.5	4.1
23	1.9	3.6	81.0	11.0	2.6
24	2.2	4.4	80.6	10.7	2.1
25	2.9	4.2	81.0	10.5	1.4
26	3.5	4.4	80.2	10.7	1.3
27	3.8	5.3	79.2	10.7	1.0
28	3.9	5.9	78.3	11.0	1.0
29	4.7	5.4	77.2	11.7	1.0
30	5.1	5.9	76.2	11.7	1.1
31	5.5	5.6	75.9	12.1	1.0
32	5.8	5.3	75.9	12.3	0.7
33	5.8	5.5	75.7	12.2	0.8
34	5.8	5.3	75.8	12.5	0.6
35	5.7	5.4	75.6	12.7	0.7
36	5.8	4.0	76.8	12.7	0.6
37	5.8	3.9	76.9	12.9	0.6
38	6.0	3.8	76.7	12.9	0.5
39	6.8	3.8	76.7	12.2	0.5
40	6.8	3.8	75.7	13.4	0.3
41	7.5	3.5	74.4	13.6	1.0

Source: NLSY79

Exclusively self-employed=self-employed, but not employer-employed; dual-employed=self-employed and employer-employed; exclusively employer-employed=employer-employed but not self-employed.

boomers.¹⁰ In presenting the analysis results, we will report weighted outcomes where our intention is specifically to make inferences about the general population.

Exhibits A.3 to A.5 provide the yearly employment status, as defined above, for the benchmark sample, as well as for the ever- and never self-employed samples. For the benchmark sample (See Exhibit A.3), the share of those who were exclusively self-employed steadily increased over time; however, the share of those who were dual-employed (that is, both self-employed and employer-employed) did not increase as steadily. Not surprisingly, the share of those who were not working but in school (full-time students) decreased with age. The share of those who were not working but not in school appears to have increased slightly, probably reflecting the increase in the out-of-labor-force population, especially among women. As one might have expected, the largest category was those who were exclusively employer-employed. The share of this group

¹⁰ We adjusted the original sample weight to account for the sample size reduction in the military sample after the 1984 wave. We use these adjusted baseline weights for our analysis. The weights correct for oversampling of Blacks, Hispanics, and economically disadvantaged non-Blacks/non-Hispanics.

decreased over time, mainly due to the shift toward exclusive self-employment. Still, the overwhelming majority—about three quarters—of the sample was exclusively employer-employed after age 40. A parallel trend was observed for those who were ever self-employed (See Exhibit A.2); however, compared to those in the benchmark, the ever self-employed sample consistently had a higher share of those exclusively self-employed as well as dual-employed, and a considerably lower share of those who were exclusively employer-employed. Exhibit A.5 suggests that those who were never self-employed largely stayed in the exclusively employer-employed category over time.

Exhibit A.4
Distribution of Yearly Employment Status by Age
Ever Self-employed Sample (N=2,394)
 (Weighted)

Age	Exclusively Self- employed (percent)	Dual-employed (percent)	Exclusively Employer- employed (percent)	Not Working and Not in School (percent)	Not Working and In School (percent)
22	4.1	8.5	73.5	9.8	4.1
23	5.3	10.2	71.4	10.5	2.5
24	6.1	12.7	69.2	10.1	1.9
25	8.3	12.0	68.5	10.0	1.2
26	9.9	12.5	66.3	9.8	1.5
27	10.8	15.2	64.1	9.2	0.8
28	11.0	16.8	61.8	9.6	0.8
29	13.4	15.3	59.6	10.9	0.8
30	14.4	16.8	55.4	11.9	1.5
31	15.6	15.8	56.0	11.8	0.8
32	16.6	15.1	56.2	11.3	0.8
33	16.6	15.7	56.5	10.4	0.9
34	16.6	15.2	56.5	11.3	0.4
35	16.2	15.3	56.6	11.4	0.5
36	16.5	11.5	59.7	11.9	0.5
37	16.6	11.1	59.9	11.8	0.7
38	17.2	10.9	58.1	13.4	0.4
39	19.4	10.8	57.2	12.0	0.5
40	19.3	10.8	55.6	14.2	0.2
41	21.4	10.1	54.2	13.3	1.1

Source: NLSY79

Exclusively self-employed=self-employed, but not employer-employed; dual-employed=self-employed and employer-employed; exclusively employer-employed=employer-employed but not self-employed.

Exhibit A.5
Distribution of Yearly Employment Status by Age
Never Self-employed Sample (N=5,088)
(Weighted)

Age	Exclusively Self- employed (percent)	Dual-employed (percent)	Exclusively Employer- employed (percent)	Not Working and Not in School (percent)	Not Working and In School (percent)
22	0.0	0.0	85.0	10.9	4.1
23	0.0	0.0	86.2	11.2	2.6
24	0.0	0.0	86.8	11.0	2.2
25	0.0	0.0	87.7	10.8	1.5
26	0.0	0.0	87.7	11.1	1.2
27	0.0	0.0	87.4	11.5	1.1
28	0.0	0.0	87.2	11.7	1.1
29	0.0	0.0	86.6	12.2	1.2
30	0.0	0.0	87.5	11.6	1.0
31	0.0	0.0	86.7	12.3	1.0
32	0.0	0.0	86.5	12.8	0.7
33	0.0	0.0	86.1	13.2	0.7
34	0.0	0.0	86.2	13.2	0.6
35	0.0	0.0	85.8	13.4	0.9
36	0.0	0.0	86.1	13.2	0.7
37	0.0	0.0	86.0	13.4	0.5
38	0.0	0.0	86.8	12.7	0.6
39	0.0	0.0	87.2	12.3	0.5
40	0.0	0.0	86.7	12.9	0.4
41	0.0	0.0	85.3	13.8	0.9

Source: NLSY79

Exclusively self-employed=self-employed, but not employer-employed; dual-employed=self-employed and employer-employed; exclusively employer-employed=employer-employed but not self-employed.

Appendix B

Review of Existing Literature

To provide topical context for our research, we present a review of the existing literature in this appendix. There has been considerable research using the individual as a unit of analysis to understand small business dynamics.¹¹ These studies are underpinned by a theoretical model of entrepreneurial choice, and the empirical investigations typically regard an individual's engagement in self-employment as an entrepreneurial decision to form a business. Economic and other social science theories also predict that a wide variety of factors—financial, social, educational, familial, psychological, or biological—influence entrepreneurship. Many of the previous studies of individual entrepreneurs were designed to determine what individual characteristics or other factors explain the “formation” of a business and, to a lesser extent, the closure or duration of a business. Accordingly, there now exists a substantial volume of research findings, albeit not always consistent, on who is more likely to be or to become self-employed and what other factors influence self-employment selection.

Self-Employment as a Unit of Study

Existing literature identifies numerous instances of and caveats to using self-employment to measure entrepreneurship. Entrepreneurship is a broad concept not entirely captured by the narrow definition of self-employment (Parker 2004). Small business activity is often used as a measure of entrepreneurship; however, as Holtz-Eakin (2000) notes, there may be entrepreneurs who have large firms and small businesses not run by entrepreneurs; hence, self-employment has emerged as a frequent proxy for entrepreneurship.

Measuring self-employment is not without challenges, as estimates of self-employment are not consistent across major data sets. For example, despite the fact that both the Current Population Survey (CPS) and the Characteristics of Business Owners (CBO) datasets have been used in numerous studies of self-employment, Boden and Nucci (1997) found substantial differences in the estimates of self-employed workers between two waves of these datasets for the same reference years, even after correcting for differences in the coverage and unit of analysis between the two surveys. As Headd and Saade (2008) note, the definition of self-employment varies across datasets, and even for the same data set, researchers conducting secondary analysis may use varying definitions of self-employment. While many data sets may not always capture those who have a primary employer-based job and run their own business as a second job, both Holtz-Eakin (2000) and Giannetti and Simonov (2004) stress that it is important to include this group because most businesses start as small units and it is impossible to predict in advance which businesses will be successful.

The definition of self-employment used in previous studies of the NLSY79 varies; Fairlie (2005B) observed that in the majority of previous studies, “self-employed workers are defined as those individuals who identify themselves as self-employed...in response to the class of worker question relating to the current or most recent job (p.41)” and definitions exclude unpaid family

¹¹ For firm-level investigation of business formation, expansion, contraction, and closure, previous studies focused on topics of growth rates, survival rates, and job creation (Headd and Kirchoff 2007).

workers, individuals enrolled in school and those who worked fewer than 300 hours in the previous year. Looking at specific studies using the NLSY, Rissman (2003), Evans and Jovanovic (1989), Dunn and Holtz-Eakin (2000), Taniguchi (2002) and Williams (2004) all define an individual as self-employed if he or she self-identifies as such; Ferber and Waldfogel (1998) use the same definition but further subdivide the self-employed by whether their business is incorporated. In our study, we will use self-employment as a basis for analyzing how individuals engage in own-initiated business activities over their life cycle. While this approach may not strictly capture behaviors of a firm as an institution, it provides important insights into an individual as an agent to generate business activities. As we will discuss in the report, the NLSY79 and NLSY97 provide extensive details on employment and activities at the individual level, which we will use to construct a self-employment measure. As our study objective is to analyze the self-employment trajectories, we will emphasize the consistency of the measure over time.

Previous Studies Focusing on Transition Into and Out of Self-Employment

Many of the previous studies of individual entrepreneurs were designed to determine what individual characteristics or other factors explain the “formation” of business and, to a lesser extent, the closure or duration of business. For example, in an earlier empirical study of individual-level data, Evans and Leighton (1989) found, among other things, that the probability of becoming self-employed was largely independent of age and past experience; those with greater assets were more likely to become self-employed; unemployed and low-wage workers were more likely to become self-employed; and those with a belief that one controls one’s own destiny (an internal locus of control) were more likely to be self-employed. More recent studies have investigated the effects on self-employment of a range of factors, including: education (Henley 2005, Moutray 2007), previous work experience, own and parents’ assets (Henley 2005, Moutray 2007), particular types of resources such as homeownership and technology (Moutray 2007, Georgellis et al 2005), military service (Moutray 2007), incarceration (Fairlie 2005B), parental experience of self-employment (Blanchflower and Oswald 2007, Fairlie and Robb 2007B), spousal characteristics (Karoly and Zissimopoulos 2004), other household members’ status, and presence of children (Dickson, et al. 2008; Kim, et al. 2006; Hundley et al. 2006; Cavalluzzo & Wolken 2005; Dunn & Holtz-Eakin 2000; Colombier & Masclat 2008; Taniguchi, 2002; Budig 2006; Salazar 2007; Kepler & Shane 2007; Fairlie 2005A; Fairlie & Robb, 2007B).

Bates and Servon (2000) also investigated the self-employment experiences of those who started their own businesses because they were unable to find suitable employment elsewhere, finding that this particular population had both lower educational attainment levels and fewer assets than the population generally targeted by small business policies and services. Williams (2004) found that self-employment experiences for youth differed significantly from those of their older counterparts; for example, while self-employment is more prevalent among part-time than full-time workers for older entrepreneurs, this pattern is not the case for teenagers. These studies were also frequently conducted in the context of documenting and explaining racial or gender gaps in self-employment (Fairlie & Robb 2007A and 2007B; Taniguchi 2002; Salazar 2007; Cavalluzzo & Wolken 2005), as well as educational attainment (Bates and Servon 2000).

While a large number of past individual-level studies have been dedicated to examining selection into self-employment, there has been less research on other aspects of business dynamics, such as exit from or duration of self-employment. This may be at least partly due to the lack of suitable data with which to study these topics. Meager (1992) notes that suitable flows data on self-employment is scarce, causing many previous studies of self-employment dynamics to rely on stocks data to measure the life cycle of self-employment; however, longitudinal data increasingly available in developed countries remedies this issue somewhat. Frequently used longitudinal data sets are the CBO, the British Household Panel Survey (BHPS), the National Longitudinal Survey (NLS) series, and the Panel Study of Income Dynamics.¹²

Several of these studies have examined the success of self-employment, as measured by either income generation or job creation. Using 1982 CBO data, Bates (1990) found that highly educated entrepreneurs were more likely to have businesses that were still in operation four years later. Schiller and Crewson (1997) used the National Longitudinal Study of Youth of 1979 (NLSY79) to find that businesses started by young adults were only moderately successful in generating a significant income. Henley (2005) focused on job creation using the BHPS and found that job creation by the self-employed was significantly associated with housing wealth, having parents who were self-employed themselves (and in particular, self-employed parents who also employed others), and university-level educational attainment.

Another portion of the literature focuses on duration of and exit from self-employment. Evans and Leighton (1989) used the NLS to estimate exit rates from self-employment, and found that the exit probability decreased with duration. Using the NLSY79, Schiller and Crewson (1997) observed that the total years spent in self-employment among young adults was relatively short—less than three years over the 11-year period—among those ever self-employed. Schiller and Crewson also demonstrated that self-employment indicators (both total years spent and income received) vary considerably by gender. This study is corroborated by findings that young women had lower self-employment entry rates and higher self-employment exit rates (Fairlie 2005A) and women-owned businesses were smaller and less successful on average than male-owned businesses (Fairlie and Robb 2007B). Fairlie (2005B) also noted that in studies of self-employment using the NLSY79, self-employment increases as the cohort ages. More recently, Rissman (2006) used the same data (NLSY79) to examine factors affecting the duration of self-employment. She found that young men's exits from self-employment are positively affected by aggregate and local economic conditions, supporting an argument that self-employment is largely represented not by entrepreneurs but by discouraged wage workers seeking a second-best alternative to unemployment. Fairlie (1999) studied self-employment dynamics using the PSID and found that having a self-employed father resulted in a large, negative, and statistically significant effect on the probability of exiting from self-employment for White men. Many studies have also found that assets are an important determinant of self-employment duration and exit rates (Evans and Jovanovic 1989, Evans and Leighton 1989, Bates 1997, Holtz-Eakin et al 1994a, 1994b, Fairlie 1999, Dunn and Holtz-Eakin 2000, and Blanchflower and Oswald 1998). Using the BHPS, Georgellis et al (2005) found that while wealthier individuals were more likely

¹² These sources all use the individual as a unit of analysis. Using the firm as a unit of analysis, Headd and Kirchhoff (2007) have also used the U.S. Census Bureau's Statistics of U.S. Businesses (SUSB) database to analyze small business dynamics; however, these findings are not specific to self-employment.

to transition to self-employment, they are also more likely to exit self-employment upon receipt of a windfall payment (such as lottery winnings or an inheritance) than less wealthy individuals.

Previous Studies Focusing on Trajectories of Self-Employment

From the perspective of an individual, self-employment is not always a single event of life, but is better understood as part of a lifelong process that forms the individual's career trajectory. The end of a self-employment episode is not always equal to the end of entrepreneurship by the individual; in fact, empirical evidence suggests that a large proportion of the self-employed have repeated self-employment episodes over their life cycle. For example, Rissman (2006) reports, based on NLSY97, that of 1,479 men who experienced self-employment in her sample, 938 had only one spell of self-employment and the rest had multiple spells. Such findings suggest that, in order to understand how a unit of entrepreneurship (defined as self-employment here) persists and develops over time, is it important not only to examine each self-employment episode separately but also to study self-employment experiences over the entire individual's work life span. This section considers studies that focused on self-employment as a process, including the length of self-employment, age at which an individual either joins or rejoins the ranks of the self-employed and the step-wise nature of employment process.

The extant literature suggests that there are distinctive paths to self-employment according to the individuals' preference and circumstances. For example, previous empirical studies have indicated that, contrary to the often-held assumption about opportunity-seeking entrepreneurs, individuals go into self-employment for a variety of reasons. Rissman (2003, 2006) argued that self-employment might be largely a temporary situation sought by discouraged wage-sector workers, a situation further explored by Bates and Servon (2000). On the other hand, Kepler and Shane (2007) showed that women differ from men in such factors as motivation, preferences, expectations, and reasons for starting a business, and the differences lead to differential outcomes in venture performance. Carr (1996) argued that women primarily choose self-employment to balance work and family, showing that family characteristics significantly predict self-employment status. Budig (2006) expanded on Carr and showed that reasons for selecting into self-employment differ by type of occupation, and concluded that family factors explained entry into nonprofessional self-employment, but not professional self-employment. These studies suggest that individuals follow a particular path into and out of self-employment over their work life.

Viewing self-employment in terms of a career trajectory, many previous studies incorporated a model in which an individual faced a range of choices with respect to self-employment, and in which self-employment was primarily viewed as a process rather than a static status. For instance, Wennberg, et al. (2007) highlighted that individuals may choose part-time entry to self-employment as a strategy to manage the uncertainty of entrepreneurship, and argue that part-time and full-time employment should be distinguished and that their interrelation needs to be studied in the context of self-employment dynamics. Self-employment status has also been shown to contain a strong genuine state-dependence or inertia effect; using BHPS data, Henley (2004) found that respondents were more likely to be self-employed in a given year if they had been self-employed the previous year, even controlling for observable and unobservable influences, than respondents who had been in paid employment the prior year. Prior job characteristics were

also shown to influence selection into self-employment by Blanchflower and Meyer (1991), who identified previous firm size, previous union status and previous earnings as important determinants of transitioning to self-employment.

Other studies focused on the relationships between the employment choices made earlier on in the life cycle and later career outcomes. For instance, Williams (2004) identified differences in later career outcomes based on early experiences with self-employment, finding that NLSY79 respondents who had been self-employed as young adults were less likely to complete higher education and had lower earnings over time than their non-self-employed peers. Another study by Karoly and Zissimopoulos (2004), which examined older workers using the Health and Retirement Study (HRS), reported higher rates of retirement out of wage and salary work than out of self-employment, in addition to workers shifting from wage work to self-employment as they age. The same authors found in a subsequent study that older workers with better working conditions, including pensions and health insurance, were less likely to transition into self-employment (Karoly and Zissimopoulos 2007).

Women and Minorities

Empirical data indicate that women and minorities, particularly African Americans and Hispanics, lag behind White men in self-employment rates and other small business performance rates (SBA 2007 and 2006). Past research has contributed to elucidating why minorities are less likely to be self-employed or why they might perform less favorably in small business compared with their White counterparts. Overall, research points to the racial difference in individuals' access to resources (e.g., financial capital, business experience) and other types of discrimination as leading explanations. For instance, Fairlie and Robb (2007) showed that the relative lack of prior experience in a family business negatively affected outcomes (sales, profits, and survival) of African-American-owned businesses. Others suggested that minorities, especially African Americans, face discrimination when applying for small business loans. This suggestion is based on findings indicating that a significant difference in loan denial rates remains even after controlling for creditworthiness, personal wealth, and other factors (Blanchflower, et al. 2003 and Cavalluzzo 2005). Lowrey (2007) reported that, to start or acquire a business, African-American-owned firms used loans from private banks less frequently than White-owned firms. Such findings, combined with other research results demonstrating liquidity constraints as a barrier to starting a business (e.g., Evans and Javnovic 1989), suggest that African Americans would be less likely to be self-employed than their White counterparts. Similarly Borjas and Bronars (1989) presented a self-employment selection model in which incomplete price information and consumer discrimination would lead to lower gains for African Americans than Whites, thus discouraging African Americans from pursuing self-employment. Additionally, discrimination once a business is established can impact minority business success. For example, Lowrey (2007) reported that minority-owned firms lag behind White-owned firms in business performance, showing that for every dollar earned by a White-owned employer firm, Hispanic-, Native American-, and Asian-owned businesses earned 56 cents, and Black-owned employer businesses earned just 43 cents.

Research also suggests differences within minority groups' experiences of self-employment in addition to the above-cited works on differences between Whites and minorities. For example,

Lunn and Steen (2005) note that heterogeneity of self-employment rates is most apparent within groups among Asian-Americans, with Koreans self-employed at nearly eight times the rate of Laotians; this pattern is also observed for Hispanics, with a large difference in the self-employment rates of Mexicans and Cubans. These differences are also observed between minority groups; Boyd (1990) observed that while marital status and having children did not affect the probability of self-employment for Blacks, both being married and having children had a positive effect on self-employment probability for Asians.

The existence of a racial gap in start-up and performance among the self-employed (or small business entities) is accepted largely as a stylized fact and, as noted above, a body of research focuses on explaining the observed gap. In contrast, past research on gender and self-employment in the United States has been largely limited to establishing how females and males differ in their characteristics as entrepreneurs or to describing the self-employment selection process for females. An emerging stylized fact about self-employed females is that they are significantly different from male counterparts not only in their financial or human capital but, more importantly, in their reasons for entering self-employment, their attitudes, and their preferences. For instance, Kepler and Shane (2007) used the Panel Study of Entrepreneurial Dynamics to find that there are significant differences between female and male entrepreneurs regarding pecuniary motivations, expectations as to their future income from the business, risk preferences, technological intensity of their business, and their approach to identifying opportunities. Other studies found that the determinants of selection to self-employment for women are different from men; for example, Schiller and Crewson (1997) showed marriage increased the probability of self-employment for women while it decreased the probability for men. Dunn and Holtz-Eakin (1995) found that having one's own financial assets modestly increased the probability of self-employment for men but did not have any effect for women. Other studies that focused exclusively on female self-employment also showed that family structures (marriage and children) mattered in their entry to self-employment (Budig 2006; Carr 1996 Taniguchi 2002). Self-employed women also have significantly different time-use patterns from their male self-employed and female non-self-employed counterparts; Gurley-Calvez et al (2009) found that self-employed women spent less time in work-related activities and more time providing child care. The authors also identified interlocking barriers for African-American women, whose self-employment rates were the lowest relative to men.

Few studies have attempted to explain the gender gap in the United States by taking into account both the difference in the process (that, is, the coefficients) and the difference in background factors (that is, the independent variables). In the United Kingdom, however, Burke and Nolan (2002) used the National Child Development Study to study the effects of regressor variables on outcomes through disaggregating their sample by gender, finding that determinants of self-employment varied between men and women: while completion of post-compulsory schooling is negatively associated with self-employment for men, there was no relationship for women; there was a positive relationship between self-employment and having children for males and less-qualified females, but not for highly qualified females; and, most interestingly, only for women does non-full-time self-employment have a significant negative effect on job creation in their businesses.

Existing literature thus offers substantial background on the determinants of self-employment and the demographics of the self-employed population, as well as on pathways to self-

employment. Our study expands on these findings, as described in subsequent sections, by analyzing the career paths of self-employed individuals from a life cycle perspective and comparing self-employed workers from two generations.

Bibliography

- Bates, Timothy. 1997. *Race, Self-Employment, and Upward Mobility: An Illusive American Dream*. Washington, D.C.: Woodrow Wilson Center Press; Baltimore: John Hopkins University Press.
- Bates, Timothy. 1990. "Entrepreneur Human Capital Inputs and Small Business Longevity." *Review of Economics and Statistics* 72(4), 551-559
- Bates, Timothy and Lisa Servon. 2000. "Viewing Self Employment as a Response to Lack of Suitable Opportunities for Wage Work." *National Journal of Sociology* 12(2), 23-55
- Blanchflower, David G. and Bruce D. Meyer. 1991. "A Longitudinal Analysis of Young Entrepreneurs in Australia and the United States." NBER Working Paper #3746.
- Blanchflower, David G. and Andrew J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1), 26-60
- Blanchflower, David G., Phillip B. Levine, and David J. Zimmerman. 2003. "Discrimination in the Small-Business Credit Market." *The Review of Economics and Statistics* 85(4), 930-943
- Blanchflower, David G. and Andrew J. Oswald. 2007. "What Makes a Young Entrepreneur?" IZA Discussion Paper #3139.
- Boden, Richard J. and Alfred R. Nucci. 1997. "Counting the Self-Employed Using Household and Business Sample Data." *Small Business Economics* 9(5), 427-436
- Borjas, George J. and Stephen G. Bronars. 1989. "Consumer Discrimination and Self-Employment." NBER Working Papers #2627.
- Boyd, Robert L. 1990. "Black and Asian Self-Employment in Large Metropolitan Areas: A Comparative Analysis." *Social Problems* 37(2), 258-274
- Brzinsky-Fay, Christian, Ulrich Kohler, and Magdalena Luniak. 2006. "Sequence Analysis with Stata." *The Stata Journal* 6(4), 435-460
- Budig, Michelle. 2006. "Intersections on the Road to Self-Employment: Gender, Family, and Occupational Class." *Social Forces* 84(4), 2223-2239
- Burke, Andrew E., Felix R. FitzRoy, and Michael A. Nolan. 2002. "Self-Employment Wealth and Job Creation: The Roles of Gender, Nonpecuniary Motivation and Entrepreneurial Ability." *Small Business Economics* 19(3), 255-270
- Carr, Deborah. 1996. "Two Paths to Self-Employment?" *Work and Occupations* 23(1), 26-53
- Cavalluzzo, Ken and John Wolken. 2005. "Small Business Loan Turndowns, Personal Wealth, and Discrimination." *Journal of Business* 78(6), 2153-2178
- Colombier, Nathalie and David Masclet. 2008. "Intergenerational Correlation in Self Employment: Some Further Evidence from French ECHP Data." *Small Business Economics* 30(4), 423-437
- Dickson, Pat H., George T. Solomon, and K. Mark Weaver. 2008. "Entrepreneurial Selection and Success: Does Education Matter?" *Journal of Small Business and Enterprise Development* 15(2), 239-258
- Dunn, Thomas and Douglas Holtz-Eakin. 1995. "Capital Market Constraints, Parental Wealth, and the Transition to Self-Employment Among Men and Women." Center for Policy Research, Syracuse University.

- Dunn, Thomas and Douglas Holtz-Eakin. 2000. "Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links." *Journal of Labor Economics* 18(2), 282-305
- Evans, David S. and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy* 97(4), 808-827
- Evans, David S. and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79(3), 519-535
- Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment." *Journal of Labor Economics* 17(1), 80-108
- Fairlie, Robert W. 2005A. "Technology and Entrepreneurship: A Cross-Industry Analysis of Access to Computers and Self-Employment." The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Fairlie, Robert W. 2005B. "Self-Employment, Entrepreneurship, and the NLSY79." *Monthly Labor Review* February 2005, 40-47
- Fairlie, Robert W. and Alicia M. Robb. 2007A. "Why Are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital." *Journal of Labor Economics* 25(2), 289-323
- Fairlie, Robert W., and Alicia Robb. 2007B. "Families, Human Capital, and Small Business: Evidence from the Characteristics of Business Owners Survey." *Industrial and Labor Relations Review* 60(2), 225-245
- Ferber, Marianne A. and Jane Waldfogel. 1998. "The Long-Term Consequences of Nontraditional Employment." *Monthly Labor Review* 121(5), 3-12
- Georgellis, Yannis, John G. Sessions, and Nikolaos Tsitsianis. 2005. "Windfalls, Wealth, and the Transition to Self-Employment." *Small Business Economics* 25(5), 407-428
- Giannetti, Mariassunta and Simonov, Andre. 2004. "On the Determinants of Entrepreneurial Activity: Social Norms, Economic Environment and Individual Characteristics." *Swedish Economic Policy Review* 11(2), 269-313
- Gurley-Calvez, Tami, Katherine Harper, and Amelia Biehl. 2009. "Self-Employed Women and Time Use." U.S. Small Business Administration, Office of Advocacy.
- Headd, Brian and Bruce Kirchoff. 2007. "Small Business Growth: Searching for Styled Facts." U.S. Small Business Administration, Office of Advocacy.
- Headd, Brian and Radwan Saade. 2008. "Do Business Definition Decisions Distort Small Business Research Results?" U.S. Small Business Administration, Office of Advocacy.
- Henley, Andrew. 2004. "Self-Employment Status: The Role of State Dependence and Initial Circumstances." *Small Business Economics* 22(1), 67-82
- Henley, Andrew. 2005. "Job Creation by the Self-Employed: The Roles of Entrepreneurial and Financial Capital." *Small Business Economics* 25(2), 175-196
- Holtz-Eakin, Douglas, David Joulfaian, and Harvey Rosen. 1994A. "Entrepreneurial Decisions and Liquidity Constraints." *Rand Journal of Economics* 23(2), 334-347
- Holtz-Eakin, Douglas. 1994B. "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." *Journal of Political Economy* 102(1), 53-75

- Holtz-Eakin, Douglas. 2000. "Public Policy Toward Entrepreneurship." *Small Business Economics* 15(4), 283-291
- Hundley, Greg. 2006. "Family Background and the Propensity for Self-Employment." *Industrial Relations: A Journal of Economy and Society* 45(3), 377-392
- Karoly, Lynn A. and Julie Zissimopoulos. 2004. "Self-Employment and the 50+ Population." Washington, DC: AARP Public Policy Institute.
- Karoly, Lynn A. and Julie Zissimopoulos. 2007. "Work and Well-Being Among the Self-Employed at Older Ages." Washington, DC: AARP Public Policy Institute.
- Kepler, Erin and Scott Shane. 2007. "Are Male and Female Entrepreneurs Really That Different?" The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Kim, Phillip, Howard Aldrich, and Lisa Keister. 2006. "Access (Not) Denied: The Impact of Financial, Human, and Cultural Capital on Entrepreneurial Entry in the United States." *Small Business Economics* 27(1), 5-22
- Kobe, Kathryn. 2007. "The Small Business Share of GDP, 1998-2004." U.S. Small Business Administration, Office of Advocacy.
- Lowrey, Ying. 2007. "Minorities in Business: A Demographic Review of Minority Business Ownership." U.S. Small Business Administration, Office of Advocacy.
- Lunn, John and Todd Steen. 2005. "The Heterogeneity of Self-Employment: The Example of Asians in the United States." *Small Business Economics* 24(2), 143-158
- Meager, Nigel. 1992. "Does Unemployment Lead to Self-employment?" *Small Business Economics* 4(2), 87-103
- Moutray, Chad. 2007. "Educational Attainment and Other Characteristics of the Self-Employed: An Examination Using Data from the Panel Study of Income Dynamics." U.S. Small Business Administration, Office of Advocacy.
- Parker, Simon C. 2004. *The Economics of Self-Employment and Entrepreneurship*. New York, NY: Cambridge University Press.
- Rissman, Ellen R. 2003. "Self-Employment As an Alternative to Unemployment." Working Paper Series WP-03-34, Federal Reserve Bank of Chicago.
- Rissman, Ellen R. 2006. "The Self-Employment Duration of Younger Men over the Business Cycle." *Economic Perspectives* Q III, 14-27
- Salazar, Maritza. 2007. "The Effect of Wealth and Race on Start-up Rates." U.S. Small Business Administration, Office of Advocacy.
- Schiller, Bradley R. and Philip E Crewson. 1997. "Entrepreneurial Origins: A Longitudinal Inquiry." *Economic Inquiry* 35(3), 523-531
- Taniguchi, Hiromi. 2002. "Determinants of Women's Entry into Self-Employment." *Social Science Quarterly* 83(3), 875-893
- U.S. Small Business Administration, Office of Advocacy. 2007. "The Small Business Economy for Data Year 2006: A Report to the President."
- U.S. Small Business Administration, Office of Advocacy. 2006. "The Small Business Economy for Data Year 2005: A Report to the President."
- Wennberg, Karl J., Timothy Folta, and Frederic Delmar. 2007. "A Real Options Model of Stepwise Entry into Self-Employment." The Office of Advocacy Small Business Working Papers, U.S. Small Business Administration, Office of Advocacy.
- Williams, Donald R. 2004. "Youth Self Employment: Its Nature and Consequences." *Small Business Economics* 23(4), 323-336

Appendix C

Multivariate Analyses

In this study, we estimate three different models to investigate our research questions. First, we examine the effects of family background and early exposure to self-employment before age 22 on total self-employed years over the next 20-year period. In this first set of estimations, we specifically explore the question of how the extent of self-employment varies by individual and family backgrounds and by early labor-market experiences. Second, instead of focusing on total self-employment years over time, we model the decision to engage in self-employment in each given year. This framework is useful in analyzing how changes in life events and circumstances—such as marital status and family assets—relate to the individual’s choice to be self-employed. Third, we model the financial outcomes of respondents over the 20-year observed period, and investigate whether these outcomes varied by the extent of self-employment, controlling for baseline individual and family characteristics. As dependent variables, we look at several alternative measures, including: income level at the end of the 20-year period, an average year-to-year income growth rate over time, and the standard deviation of the income growth rate as a variability indicator over time.¹³

Multivariate Analysis of Total Self-Employment Years

In this subsection, we examine how baseline factors affect total years of self-employment over the 20-year period from ages 22 to 41. To investigate how the effects of baseline characteristics on years of self-employment might have also changed over time, we also separately examine total years of self-employment in the first ten years of the reference period and total years of self-employment in the last ten years. In modeling years of self-employment, we treat the observed measure of self-employment as being censored at zero and apply the tobit method.

In the tobit model, we assume that there is an unobservable process that explains the level of an individual’s engagement in self-employment on a continuum. Specifically, we describe the relationship between the observed measure of self-employment and the underlying process for person i as: $Y_i = y_i = \beta'X_i + u_i$, if $y_i > 0$; and $Y_i = 0$, if $y_i \leq 0$, where Y is the total number of years of self-employment that can be observed; y is an unobserved continuum measure of self-employment; X is a set of covariates measured prior to or at the beginning of the 20-year period over which self-employment years are measured; and u is an error term. The model is estimated using the maximum likelihood estimation method. The estimated coefficient β measures marginal effects of the covariate X on the underlying self-employment measure y (i.e., $\partial E(y | X)/\partial \beta$), not on the observed self-employment years Y . For the purpose of our analysis, the effect on the unobserved variable y , which can be also interpreted as a desired level of self-employment over the reference period, is an appropriate indicator for assessing how background factors are associated with the level of self-employment.

For the benchmark tobit model, we included as baseline covariates: individual demographic characteristics (gender, race, marital status at age 22, number of children at age 22, highest grade

¹³ We also examine family net assets as a wealth accumulation measure, but due to missing observations, we examined the asset variables for specific survey years only.

completed at age 22, AFQT test score administered in 1980, and an indicator for having a health problem affecting work at age 22); family background covariates (parent's foreign-born status, parents' college completion status, residence on a farm at age 14, residence in the country [nonfarm] at age 14, speaking a language other than English while growing up, and family poverty status of the previous year at age 22); and an indicator for self-employment between ages 20 and 22 as a measure of early relevant labor market experience. As an alternative model, we included additional variables with the view to refining the benchmark specification. These added variables include: parents' total years of schooling, residence in a standard metropolitan statistical area (SMSA) at age 22, and region of residence at age 22. Besides these covariates, we also included in all of our tobit models a set of dummy variables indicating the calendar year at age 22 to control for cohort-specific effects and a set of dummy adjustment variables to address missing covariates.

In cases where observations are missing due to nonresponse, we imputed missing values based on data available from adjacent survey years, if such imputation was considered reasonable. Otherwise, we applied the dummy variable adjustment method to address missing data, filling in missing values using the weighted estimation sample means for continuous variables or the value zero for binary variables. The extent of missing data problems were limited in most cases. Most of the covariates we used had less than five percent of observations missing (after adjacent year imputation), and all variables were missing less than 10 percent of their records, except for the early labor market indicator. Because the NLSY79 did not survey everyone before age 22, we were not able to compute the early labor market indicator of self-employment for about a quarter of the estimation sample (1,912 of 7,482 respondents). We applied both the dummy variable adjustment method and the listwise deletion method with regard to the early labor market indicator, and we present the estimation results from both methods. Since the early labor market indicator is missing at random (i.e., the missing pattern is independent of employment decisions or other covariates), the use of listwise deletion is not expected to lead biased estimators but does compromise the estimation sample size.

The results from the tobit estimations of self-employment years over the entire reference period, over the first decade, and over the last decade are summarized in Exhibit C.1, C.2, and C.3 respectively. For each exhibit, we present the estimation results from the benchmark tobit model, the alternative and expanded model, and the alternative model with deletion of missing early labor market data. As reported in Exhibit C.1, the estimation results for total years of self-employment over the 20-year period are very similar across the three models. The results indicate that females and minorities reported significantly lower years of self-employment than males and Whites. For example, according to the benchmark model presented (see the first column), total years of self-employment decrease by 1.7 years for being female, 2.3 years for being Black, and 1.5 years for being Hispanic. Early engagement in self-employment is estimated to have a sizable and significant effect (9.2 years) on total years of self-employment. The estimation results also suggest that parents' education and residence type at age 14 influenced total years of self-employment. According to the benchmark model, having a college-educated parent, for example, increases total years of self-employment by 1.6 years. Total years of self-employment also increased by 1.6 years if the respondent had lived on a farm at age 14. Residence on a farm while young may thus facilitate later self-employment by offering opportunities that support the pursuit of self-employment. As noted above, our reference to self-

employment years here corresponds to the unobserved measure of years of self-employment. It should be noted, therefore, that we do not expect these estimated marginal effects on observed years of self-employment.

Exhibits C.2 and C.3 suggest that there are some important differences in how baseline characteristics affect the level of engagement in self-employment. For total self-employment years in the first decade from age 22 to 31, the number of children among females at age 22 had significantly positive effects; however, it had significantly negative effects in the second decade from age 32 to 41. (As a result, the number of children among females was not found to have significant effects when the model was estimated for total self-employment years over the entire 20-year period.) These seemingly conflicting results indicate a complex decision-making process that females face in determining engagement in self-employment when they have a child.¹⁴ The results here suggest that having a child by age 22 increases total years of self-employment over the employment life cycle, possibly because younger females with (young) children have the need to work as caretakers, but have more barriers to traditional employer-based jobs, thus leading them to be more likely to pursue self-employment than males with a child or childless females. By the second decade (when children are older), females who had a child by age 22 may have become less motivated or constrained to pursue self-employment per se, while still facing the general challenges associated with being a working mother. The negative effects then may reflect the prevalent tendency for females with children to work less (either in self- or employer-based employment). We also found that being married at 22 and growing up speaking a foreign language at home were positively associated with years of self-employment between ages 22 and 31. These factors, however, did not matter in predicting years of self-employment between ages 32 and 41.

In this subsection, we focused on understanding the relationship between baseline factors and an eventual self-employment outcome measured in terms of total years. The estimation results suggest that some of the variation in self-employment years over time can be explained by baseline characteristics. We also found that the effects of background characteristics on years of self-employment over the observed period may change with time. This finding indicates that examining the effects of background factors at each time period, controlling for other concurrent factors, would provide additional insight into how these factors relate to self-employment outcomes. This relationship is explored in the next subsection.

¹⁴ We also estimated the same model for total years the person was exclusively self-employed as well as total years that the person was exclusively employer-employed. The results from the former estimation model confirm the findings reported in Exhibits C.2-C.3 with respect to the number of children among females at age 22. The results from the latter model, on the other hand, show that the effects were strongly negative for total years in both the first and second decades.

Exhibit C.1
Tobit Estimation of Total Number of Years Self-Employed Between Age 22 and 41
(Weighted)

	(1)			(2)			(3)		
	Benchmark Model Same as MNL model in Exhibit F.1			Alternative Model with additional covariates			Alternative Model except missing early self-employment data		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Constant	-1.629	1.125		-2.989	1.301	**	-1.770	1.468	
Female	-1.748	0.296	***	-1.798	0.298	***	-1.565	0.337	***
Black	-2.261	0.334	***	-2.084	0.347	***	-2.149	0.383	***
Hispanic	-1.503	0.482	***	-1.499	0.509	***	-1.169	0.576	**
Married at age 22	0.387	0.327		0.445	0.329		0.601	0.378	
Number of children at age 22	0.141	0.376		0.157	0.377		0.373	0.443	
Female * Number of children at age 22	-0.082	0.420		-0.102	0.420		-0.536	0.485	
Self-employed at ages 20-22	9.206	0.500	***	9.166	0.502	***	9.104	0.506	***
Highest grade completed at 22	-0.099	0.094		-0.132	0.095		-0.156	0.109	
AFQT score	-0.003	0.006		-0.004	0.006		-0.001	0.007	
Family in poverty last yr (at 22)	-0.105	0.362		0.000	0.363		0.082	0.404	
Health problem affecting wk at 22	0.757	0.508		0.632	0.510		0.468	0.563	
Parent born outside US	-0.564	0.500		-0.569	0.507		-0.775	0.564	
Parent graduated college	1.551	0.357	***	1.189	0.435	***	1.268	0.494	**
Spoke non-English growing up	0.473	0.492		0.531	0.490		0.064	0.559	
Lived in rural area (nonfarm) at 14	0.170	0.349		0.329	0.356		0.116	0.402	
Lived on farm at age 14	1.625	0.588	***	1.692	0.592	***	1.160	0.664	*
Highest grade completed by mother				0.158	0.061	***	0.097	0.069	
Highest grade completed by father				-0.034	0.050		-0.067	0.057	
Lived in an SMSA at age 22				0.642	0.325	**	0.944	0.368	**
Lived in Northeast region at age 22				-0.427	0.403		-0.576	0.460	
Lived in South region at age 22				-0.264	0.343		-0.334	0.383	
Lived in West region at age 22				1.027	0.398	**	0.816	0.449	*
Test of joint significance of repressors	F(28,7452)=19.8***			F(36,7444)=16.0***			F(34,5536)=15.0***		
Number of observations(a)	7,480			7,480			5,570		

*** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent

Note: The observed total number of years self-employed ranged from 0 to 20. The number of self-employed years are regarded as censored at zero. In addition to covariates listed in the table, each estimation model includes a set of dummy adjustment variables for missing covariates and cohort dummy variables (the calendar year indicator at age 22). Fewer (2 less) control variables were included in the model (3) as they become collinear.

(a) The number of observations used in the estimation is reported. The total analysis sample is 7,482, but two individuals are given zero weight and excluded from the estimation with the use of weights.

Exhibit C.2
Tobit Estimation of Number of Years Self-Employed Between Age 22 and 31
(Weighted)

	(1)			(2)			(3)		
	Benchmark Model Same as MNL model in Exhibit F.1			Alternative Model with additional covariates			Alternative Model except missing early self-employment data		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Constant	-2.941	0.81	4 ***	4.387	0.93	9 ***	3.160	1.03	5 ***
Female	-1.315	0.205	***	-1.358	0.206	***	-1.194	0.229	***
Black	-1.468	0.238	***	-1.403	0.248	***	-1.248	0.267	***
Hispanic	-1.216	0.340	***	-1.147	0.357	***	-1.033	0.392	***
Married at age 22	0.457	0.229	**	0.519	0.230	**	0.587	0.259	**
Number of children at age 22	-0.322	0.265		-0.304	0.265		-0.192	0.302	
Female * Number of children at age 22	0.621	0.298	**	0.602	0.297	**	0.314	0.336	
Self-employed at ages 20-22	7.034	0.304	***	7.004	0.305	***	6.851	0.304	***
Highest grade completed at 22	-0.013	0.067		-0.046	0.069		-0.089	0.077	
AFQT score	-0.004	0.004		-0.005	0.005		-0.002	0.005	
Family in poverty last year (at age 22)	-0.213	0.263		-0.122	0.262		-0.036	0.290	
Health problem affecting work at age 22	0.791	0.362		0.714	0.360	**	0.531	0.403	
Parent born outside US	-0.577	0.351	*	-0.577	0.354		-0.530	0.389	
Parent graduated college	0.774	0.250	***	0.415	0.304		0.458	0.336	
Spoke non-English growing up	0.739	0.329	**	0.774	0.327	**	0.602	0.356	*
Lived in rural area (nonfarm) at age 14	0.066	0.245		0.186	0.248		-0.001	0.277	
Lived on farm at age 14	0.890	0.387	**	1.020	0.390	***	0.781	0.414	*
Highest grade completed by mother				0.139	0.043	***	0.090	0.048	*
Highest grade completed by father				-0.021	0.036		-0.031	0.040	
Lived in an SMSA at age 22				0.698	0.227	***	0.756	0.253	***
Lived in Northeast region at age 22				-0.138	0.279		-0.388	0.315	
Lived in South region at age 22				0.034	0.239		-0.002	0.264	
Lived in West region at age 22				0.671	0.273	**	0.482	0.301	
Test of joint significance of repressors	F(28,7452)=26.9***			F(36,7444)=21.5***			F(34,5536)=21.2***		
Number of observations(a)	7,480			7,480			5,570		

*** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent

Note: The observed total number of years self-employed ranged from 0 to 10. The number of self-employed years are regarded as censored at zero. In addition to covariates listed in the table, each estimation model includes a set of dummy adjustment variables for missing covariates and cohort dummy variables (the calendar year indicator at age 22). Fewer (2 less) control variables were included in the model (3) as they become collinear.

(a) The number of observations used in the estimation is reported. The total analysis sample is 7,482, but two individuals are given zero weight and excluded from the estimation with the use of weights.

Exhibit C.3
Tobit Estimation of Total Number of Years Self-Employed Between Age 32 and 41
(Weighted)

	(1)			(2)			(3)		
	Benchmark Model Same as MNL model in Exhibit F.1			Alternative Model with additional covariates			Alternative Model except missing early self-employment data		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Constant	-3.468	1.088	***	-4.257	1.267	***	-4.233	1.478	***
Female	-1.210	0.285	***	-1.228	0.288	***	-1.161	0.336	***
Black	-1.791	0.327	***	-1.614	0.338	***	-1.782	0.388	***
Hispanic	-0.717	0.468		-0.737	0.498		-0.252	0.595	
Married at age 22	0.112	0.320		0.144	0.322		0.375	0.380	
Number of children at age 22	0.490	0.354		0.497	0.354		0.784	0.419	*
Female * Number of children at age 22	-0.991	0.411	**	-0.993	0.412	**	-1.347	0.482	***
Self-employed at ages 20-22	4.415	0.499	***	4.404	0.501	***	4.474	0.501	***
Highest grade completed at 22	-0.092	0.090		-0.113	0.092		-0.059	0.108	
AFQT score	0.002	0.006		0.001	0.006		0.000	0.007	
Family in poverty last year (at age 22)	0.258	0.350		0.322	0.351		0.329	0.399	
Health problem affecting work at age 22	0.247	0.493		0.167	0.495		-0.011	0.581	
Parent born outside US	-0.324	0.492		-0.329	0.496		-0.725	0.578	
Parent graduated college	1.269	0.341	***	0.985	0.418	**	1.162	0.492	**
Spoke non-English growing up	0.013	0.475		0.058	0.472		-0.481	0.585	
Lived in rural area (nonfarm) at age 14	-0.012	0.345		0.108	0.353		0.007	0.409	
Lived on farm at age 14	1.398	0.528	***	1.407	0.536	***	0.974	0.630	
Highest grade completed by mother				0.081	0.059		0.046	0.071	
Highest grade completed by father				0.005	0.049		-0.034	0.058	
Lived in an SMSA at age 22				0.236	0.317		0.526	0.371	
Lived in Northeast region at age 22				-0.333	0.387		-0.247	0.452	
Lived in South region at age 22				-0.339	0.333		-0.379	0.384	
Lived in West region at age 22				0.801	0.381	**	0.761	0.447	*
Test of joint significance of repressors	F(27,7453)=9.22***			F(35,7445)=7.56***			F(33,5537)=6.51***		
Number of observations(a)	7,480			7,480			5,570		

*** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent

Note: The observed total number of years self-employed ranged from 0 to 10. The number of self-employed years are regarded as censored at zero. In addition to covariates listed in the table, each estimation model includes a set of dummy adjustment variables for missing covariates and cohort dummy variables (the calendar year indicator at age 22). Fewer (2 less) control variables were included in the model (3) as they become collinear. For this estimation (the dependent variable=the total years from the last decade), an additional control variable was dropped due to perfect prediction.

(a) The number of observations used in the estimation is reported. The total analysis sample is 7,482, but two individuals are given zero weight and excluded from the estimation with the use of weights.

Panel Data Analysis of the Probability of Self-Employment

In order to assess further how selection into self-employment varies by individual and family background and other key factors suspected to be affecting self-employment (such as assets and health care coverage), we modeled the probability of self-employment utilizing the panel of individual data. We applied the generalized estimating equations (GEE) method to take into account within-individual correlation in the panel data.¹⁵ Specifically, we estimated the model as follows:

$$\log\{E(Y_{it})/(1 - E(Y_{it}))\} = \beta'X_{it}$$

where $Y = 1$ if the person is self-employed at age t and $Y=0$ otherwise. X is a set of covariates including both time-variant and time-invariant demographic characteristics, as well as parent characteristics measured at baseline. As family background variables, we included: parent's foreign-born status, parent's education, residence on a farm at age 14, and family poverty status at age 22. As a measure of early relevant labor market experience, we included an indicator for self-employment between age 20 and 22. For individual characteristics, we controlled for: age (which is the time variable for the model), gender, race, education, marital status, number of children, health status, and current residence in standard metropolitan statistical areas (SMSA). In addition we included employment status and income and asset levels at the beginning of the age year, which is approximated by the measures from the previous year (i.e. we included the lagged values). We also included a set of dummy variables to control for any calendar-year specific effects.

As in the case for the estimation models presented earlier, we imputed missing values based on data available from adjacent survey years. Where such data were not available, we included dummy indicators for missing covariate observations; in such cases, we replaced missing values using the weighted estimation sample means for continuous variables or the value zero for binary variables. As discussed before, while missing data are limited in most cases, the early labor market indicator of self-employment is missing for a quarter of the sample (1,912 out of 7,482 respondents) because not everyone was surveyed between ages 20 and 22. We applied both the

¹⁵ The GEE method was first proposed by Liang and Zeger (1986) and is now widely used by researchers with correlated discrete response data. Assuming that each person i ($i = 1, 2, \dots, N$) is observed at time $t = 1, 2, \dots, T_i$, with an outcome indicator $Y_i = (y_{i1}, y_{i2}, \dots, y_{iT_i})$ and a set of explanatory variables $X_i = (x_{i1}, x_{i2}, \dots, x_{iT_i})$, the GEE model is expressed as: (5) $G\{E(Y_{it})\} = \beta'X_{it}$ and $\text{Var}(Y_{it}) = \phi \text{Var}\{E(Y_{it})\}$, where β is a vector of coefficients to be estimated and usually the primary interest of the research and $G\{\}$ is a link function which describes the relationship between the linear predictor and the mean of the distribution function of the dependent variable. The variance of Y_{it} is described as a function of the mean, with an unknown scalar ϕ . In our case, Y is a self-employment status indicator, and the time indicator is age. We assume that $G\{\}$ is the logit function, Y is a distributed binominal, and $\phi = 1$. An important feature of GEE is that it provides consistent and asymptotically normal estimates of β without requiring the correlation structure between Y_{is} and Y_{it} to be specified correctly. An important feature of GEE is that it does not require the within-group correlation structure (correlation between Y_{is} and Y_{it} for given i) to be specified correctly. With some working correlation structure, the GEE method provides consistent and asymptotically normal estimates of β . The GEE model provides a so-called population-average estimator, which focuses on the population-averaged marginal expectation of the dependent variable (i.e., the average outcome over the population). This contrasts with a so-called subject-specific model, which focuses on the changes in a person's outcomes. In our application, the population-average model shows the average probability of being self-employed (among the population) and how this population-average probability of self-employment might differ based on exploratory factors.

dummy variable adjustment method and the listwise deletion method. The estimation results from both methods were very similar (both in terms of point estimates and significance) and do not alter the findings. For the purpose of discussion, we present here the results from the dummy variable adjustment method. The base model sample included 7,482 individuals over the period from age 23 to 41 (19 periods). Because of the inclusion of lagged values in the model, the estimation sample starts from age 23 instead of age 22. As mentioned above, we examined the effects of health plan coverage in separate estimations, because the survey questions regarding health insurance coverage were not asked prior to 1989. We have therefore estimated the same model with a smaller and more imbalanced sample (in terms of observed ages) rather than using the full panel.

The estimation results are reported in Exhibits C.4 – F.10. Exhibit C.4 shows GEE estimation results based on three different within-group correlation structures. With a view towards providing a more intuitive interpretation of the results, in addition to the estimated coefficient β_j , we present an estimate of the odds ratio, or the exponential of the estimated coefficients, $\exp(\beta_j)$, for each covariate j . The odds ratio for j approximates a change in the odds, which is the probability of being self-employed divided by the probability of not being self-employed, due to a one unit change in the covariate j .¹⁶ The first column of Exhibit C.4 shows the results with the identity matrix as a working correlation matrix. This model specification is the equivalent of computing heteroskedasticity-corrected (“robust”) standard errors for logistic regression. The second column shows the results assuming equal correlations (the nondiagonal element of the matrix is assumed to be a constant), and the third column makes the least restrictive assumption about the correlation structure, with the only constraint being that the diagonal element is assumed to be fixed at one. GEE estimates are asymptotically consistent regardless of how the correlation structures are specified. The estimation results show that, indeed, estimates are very similar to each other with minor differences. The overall findings drawn from the estimation are the same. In addition to the three correlation matrix assumptions presented in Exhibit C.4, we also used autoregressive correlation matrix assumptions (of orders 1 and 10) to estimate the model, and found very similar results. Thus, all correlation structures tested provided empirically consistent findings. In representing our results below we refer to the estimates based on an unstructured correlation matrix because it is least restrictive.¹⁷

Exhibit C.4 indicates that the probability of self-employment varies significantly across early-life experience measures as well as current characteristics. Not surprisingly, the most important predictor of self-employment is previous employment status. Those who were self-employed (either exclusively self-employed or dual-employed) in the previous year are more likely to be self-employed in the current year than those who were employer-employed. For example, for the unstructured correlation matrix model (the third model), we estimated that the odds of self-employment among those who were previously exclusively self-employed were 37 times higher than the odds of self-employment among those who were employer-employed. The exceptionally

¹⁶ We use the term “odds” to refer to this relative probability concept (the ratio of the probability of success over the probability of failure). However, general points about the chance of self-employment may be made by referring to either the odds or probability.

¹⁷ We also inspected a sample within-group correlation matrix of the estimation sample. The empirical correlation matrix does not seem to support the assumption of constant α . A higher order autoregressive or unstructured working correlation matrix produced the estimated correlation matrix which was more similar to the empirical matrix. The coefficient estimates are, however, largely consistent across all matrix assumptions tested.

large odds ratio underscores that it is the strong predictor of the outcome. It is thus important to control it in the model in order to estimate other effects more accurately, independent of the previous self-employment status. The exhibit also shows that early exposure to self-employment is another important factor. The unstructured matrix model (column 3), for example, indicates that the odds of current year self-employment among those who were self-employed between ages 20 and 21 is about three times higher than the odds of current year self-employment among those who did not have this early exposure to self-employment.

Exhibit C.4 also indicates that the probability of being self-employed was significantly higher among those who had at least one college-educated parent and those who lived on a farm when they were young. In terms of the odds ratio, the odds of self-employment for those with a college-educated parent were 14 percent higher than for those with parents who did not complete college; similarly, the odds of self-employment for those who lived on farm at age 14 were 46 percent higher than the odds for those who did not. The odds of self-employment were also significantly (14 percent) higher for those who were married than for those who were not; 13 percent higher for those who currently live in an SMSA than for those who do not (odds ratio of 1.13); and 20 percent higher for a person who is one year older.

The exhibit also indicates that the likelihood of being self-employed was significantly lower among females, minorities, and those who report a health problem that affects the ability to work. In terms of the odds ratio, the odds for females being self-employed is 62 percent of that for males; the odds for Blacks being self-employed is 63 percent of that for Whites; and the odds of Hispanics being self-employed is 70 percent of that for Whites. The odds of self-employment among those who report a health problem is 83 percent of those who do not report such a problem. Completing a four-year college degree is negatively related to self-employment probability (odds ratio of 0.82), suggesting that college graduates are less likely to be self-employed; however, controlling for completion of a college degree, an additional year of education increases the odds of being self-employed by 5 percent (odds ratio of 1.05). One interpretation of this result is that additional years of schooling help equip individuals with skills needed to be self-employed, but higher degrees do not necessarily increase the likelihood of self-employment. Another way to look at the results is that college graduates are less likely to be self-employed than noncollege graduates, but among college graduates, those with more schooling are more likely to be self-employed.

The effects of financial resources were examined using family income and net worth measures. As noted, these variables are measured in the previous year to minimize the endogeneity problem arising from self-employment statuses and income. The estimation results show that the effects of financial resources are positive and significant, but the effect is so small that an increase of \$1,000 or even \$100,000 does not make any consequential difference in the odds of being self-employed.

Exhibit C.4
Probably of Self Employment
Generalized Equating Equations (GEE) Estimation Results (Weighted)

	(1)			(2)				(3)		
	<u>Independent Correlation Matrix</u>			<u>Equal Correlation Matrix</u>				<u>Unstructured Correlation Matrix</u>		
	Odds Ratio	Coef.	Std. Err.	Odds Ratio	Coef.	Std. Err.		Odds Ratio	Coef.	Std. Err.
Constant		-7.779	1.067 ***		-7.365	0.955 ***			-7.264	0.952 ***
Background/History										
Parent born outside U.S.	0.947	-0.055	0.066	0.949	-0.052	0.062		0.968	-0.032	0.072
Parent completed college	1.124	0.116	0.053 **	1.118	0.111	0.049 **		1.141	0.132	0.056 **
Lived on farm at age 14	1.256	0.228	0.076 ***	1.327	0.283	0.072 ***		1.459	0.378	0.083 ***
Family in poverty at age 22	0.937	-0.065	0.056	0.938	-0.064	0.052		0.888	-0.119	0.063 *
Self-employed ages 20-22	1.601	0.471	0.073 ***	1.938	0.662	0.066 ***		3.026	1.107	0.069 ***
Current year characteristics										
Age	1.211	0.191	0.068 ***	1.189	0.173	0.061 ***		1.202	0.184	0.061 ***
Age squared	0.997	-0.003	0.001 **	0.998	-0.002	0.001 ***		0.998	-0.002	0.001 ***
Female	0.703	-0.353	0.058 ***	0.696	-0.363	0.053 ***		0.622	-0.475	0.061 ***
Black	0.707	-0.346	0.049 ***	0.693	-0.367	0.045 ***		0.625	-0.471	0.053 ***
Hispanic	0.829	-0.187	0.056 ***	0.798	-0.226	0.052 ***		0.701	-0.355	0.063 ***
Highest grade completed	1.042	0.042	0.017 **	1.046	0.045	0.016 ***		1.048	0.047	0.018 ***
Did not complete 12th grade	1.149	0.139	0.078 *	1.155	0.144	0.073 **		1.115	0.108	0.083
Completed four-year college	0.846	-0.167	0.082 **	0.845	-0.168	0.074 **		0.820	-0.198	0.080 **
Number of children	1.018	0.018	0.021	1.005	0.005	0.020		1.004	0.004	0.021
Married	1.055	0.054	0.041	1.087	0.083	0.037 **		1.114	0.108	0.036 ***
Number of children x Female	0.950	-0.051	0.030 *	0.959	-0.042	0.027		0.954	-0.047	0.031
Health problem affecting work	0.859	-0.152	0.081 *	0.841	-0.173	0.074 **		0.830	-0.186	0.069 ***
Live in an SMSA	1.109	0.103	0.051 **	1.111	0.105	0.046 **		1.128	0.121	0.047 **
Vbls from Previous Year (t-1)										
Family income (\$'000)	1.000	4.0E-04	2.3E-04 *	1.000	3.7E-04	1.9E-04 **		1.000	3.3E-04	1.4E-04 **
Family net worth (\$'000)	1.000	2.5E-04	8.9E-05 ***	1.000	2.0E-04	7.7E-05 ***		1.000	2.6E-04	6.7E-05 ***
Exclusively self-employed	351.748	5.863	0.065 ***	127.716	4.850	0.048 ***		37.006	3.611	0.046 ***
Dual-employed	76.302	4.335	0.046 ***	37.925	3.636	0.042 ***		10.943	2.393	0.047 ***
Not employed	1.841	0.610	0.061 ***	1.692	0.526	0.057 ***		1.658	0.506	0.055 ***

Data: NLSY79, 1979-2005 surveys. ** Statistically significant at 1 percent; * at 5 percent; * at 10 percent.

The sample includes 7,482 individual each with 19 periods observed from age 22 to 41. In addition to the covariates reported in the table, each estimation model included a set of dummy variables for calendar year indicators and dummy indicators for missing observations. Missing values are replaced by the sample means for continuous variables and 0 for binary variables. Number of observations included in the estimation are 142,120, consisting of 19 periods and 7,480 persons.

Regression Analysis of Economic Outcomes

Having analyzed how baseline individual and family background characteristics and early self-employment experiences are linked with self-employment trajectory type outcomes and with the extent of self-employment over the early adult work life, we now examine how respondents' self-employment trajectories might affect economic outcomes. Descriptive analyses in Appendices D and E suggested that the average family income was likely to be higher among those who were extensively engaged in self-employment. In this appendix, we further investigate economic outcomes by years of self-employment by examining alternative income measures and controlling for background characteristics.

We use annual net family income, as well as the respondent's own reported income, as our economic measures. We examine family income, assuming that an individual's employment decision is motivated to maximize economic gains for his/her family unit. As such, family income can be viewed as one measure for an individual's success as an economic agent. The family income variable used here includes both earned income (e.g. wages from employer-based work and earnings from an individual's own business) as well as unearned income (e.g. interest, capital gains). We also examined how individual self-employment patterns are related to the respondent's own income. The own income measure used in our study includes wages and salaries as well as income (earnings) from the respondent's own business or farm. The use of own income allows us to assess a more direct link between financial outcomes and respondents' own career pathway choices.¹⁸ In addition to income variables, we looked at reported family net worth as a wealth accumulation measure; however, due to a high percentage of missing observations, we regard the net worth variable only as a secondary measure.

We model the financial outcomes of respondents as a function of their extent of self-employment as follows:

$$Y_i = \beta' X_i + \eta' Z_i + u_i, \text{ for } i = 1, 2, \dots, N,$$

N represents all respondents. Y is an economic outcome measure; X represents demographic characteristics that are expected to explain the respondent's economic outcomes; Z is a measure for the extent of self-employment; and u is an error term. We estimate the model using OLS, making the usual assumption that the error term is independent and identically distributed. Our primary interest here is to estimate and test the significance of the coefficient η . For the covariates X, we include: gender, race, AFQT score (a proxy for the respondent's general skill level), and the highest grade completed by the respondent's parent. As a measure of the extent of self-employment, we used total years of self-employment over the 20-year reference period. We estimated models using several alternative measures for financial outcome Y, including:

- income at the end of the 20-year period (at age 41)
- average income over the 20-year period
- average year to year income growth rate over the 20-year period
- standard deviation and coefficient of variation of the income growth rate

¹⁸ However, we cannot make a precise correlation between a particular job and its portion of total own income received due to the limitations of the data.

- family net worth at the end of the 20-year period (at age 41) and at the survey year 2000 (ages 37-41).

The annual income level at age 41 was examined to assess how the respondents' self-employment experience over 20 years informed their economic outcomes at the end of the period studied. Since this measure is sensitive to idiosyncratic disturbances of the particular period, we also examined the respondents' income, averaged over the 20-year period. The average income is used to measure typical economic performance experienced along the employment pathway, rather than the end result. In addition to the level of income, we also use the average year-to-year change as an outcome measure to assess whether and how employment experiences over time may explain how quickly income grew over time. To assess how stable an economic pattern had been for the family and individual, we also examined the standard deviation and the coefficient of variance (i.e., the standard deviation divided by the mean) of the income growth rate. As noted above, we examined both annual net family income and own income measures as measures of respondents' income. As a secondary measure of economic outcomes, we used the family net worth variable; specifically, we estimated the net worth at the end of the 20-year period using imputed values for missing observations as well as for a specific year (2000) without imputed observations (the year was selected based on the size of nonmissing observations).

Income data are not available for every round of the survey. For example, net family income was reported for all 22 rounds of the survey for only a small fraction (about 8 percent) of the sample, and the data were not missing for five or more rounds for half of the sample. In computing the means of income over the 20-year period, the missing observations were excluded. We then estimated the model in three different ways: (1) including all individuals for whom the means could be calculated (i.e., income was reported for at least one round), (2) including those missing data for eight or fewer rounds of the survey, and (3) including those missing five or fewer rounds. We found that the findings from the different estimation samples produced consistent results. We therefore present the results based on a subsample who were missing income data for eight or fewer rounds.

We first examine the relationship between economic outcome measures and self-employment trajectory types among those who had ever been self-employed. In these analyses, our aim is to make inferences about individuals who choose to engage in self-employment. For more generalizable analyses, we examine the effects of the number of self-employment years on the economic outcome measures using the full sample.

The results from the family and own income estimations are presented in Exhibit C.5. Overall, we found that the level of self-employment did not explain either the income growth rates or the variability of the growth rates. Background characteristics also did not predict the income growth rate measures. On the other hand, we found that employment patterns (and background variables) explained the variation in the *level* of income.

Exhibit C.5
OLS Estimation of Family and Individual Income Measures (Weighted)

	(1) Family Income at Age 41 (in '000)			(2) Average Family Income (in '000)			(3) Average Family Income Growth Rate (percent)		(4) Coefficient of Variation for Growth Rate (percent)	
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	<u>Coef.</u>	<u>Std. Err.</u>
Number of years self-employed	1.10	0.31	***	1.39	0.21	***	-0.07	0.05	-0.10	0.22
Number of years employer-employed	0.92	0.19	***	0.84	0.14	***	-0.11	0.07	-0.30	0.34
Female	-9.23	1.65	***	-18.28	1.25	***	-0.44	0.55	-0.25	1.86
Black	-4.84	1.41	***	4.93	0.95	***	-0.57	0.37	1.31	3.61
Hispanic	1.84	1.72		5.92	1.16	***	-0.59	0.52	-3.81	2.62
AFQT	0.54	0.04	***	0.20	0.02	***	0.01	0.01	-0.01	0.03
Parent's highest grade completed	1.63	0.33	***	1.50	0.22	***	0.07	0.10	-0.12	0.35
Spouse income (in '000)	0.87	0.05	***	2.05	0.10	***	-0.04	0.03	-0.12	0.11
Constant	-14.97	4.93	***	-12.51	3.30	***	2.91	2.00	12.01	10.05
Number of observations	6,233			5,293			5,293		5,293	
	(5) Own Income at Age 41 (in '000)			(6) Average Own Income (in '000)			(7) Average Own Income Growth Rate (percent)		(8) Coefficient of Variation for Growth Rate (percent)	
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	<u>Coef.</u>	<u>Std. Err.</u>
Number of years self-employed	1.17	0.26	***	0.63	0.08	***	0.35	0.31	0.40	0.41
Number of years employer-employed	1.54	0.14	***	0.58	0.09	***	0.45	0.50	0.67	0.81
Female	-20.51	1.49	***	-9.04	0.39	***	2.80	2.44	-1.90	1.52
Black	1.49	1.32		-0.55	0.35		-0.01	0.35	1.67	1.36
Hispanic	2.93	1.56	*	1.01	0.49	**	0.02	0.50	1.45	1.22
AFQT	0.37	0.03	***	0.14	0.01	***	0.02	0.03	0.00	0.03
Parent's highest grade completed	0.82	0.29	***	0.44	0.08	***	0.23	0.24	0.07	0.08
Constant	-8.90	3.86	**	1.32	1.98		-12.33	14.41	-11.89	17.62
Number of observations	3,523			5,510			4,378		4,378	

Source: NLSY79. Notes: *** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent. For (1) and (5), we include only those for whom income was reported or could be imputed. For (2)-(4) and (6)-(8), we include only those for whom the data were reported or imputed for 12 or more out of 20 years. Family and individual incomes are constructed from different sources and not always consistent. For family income, nonsurvey year incomes were imputed; for individual income, they were not imputed; consequently, the number of observations available for a year-to-year estimate of family income was fewer than the number available for yearly individual income. Each equation contains dummy indicators for missing variables for AFQT, parents' highest grade completed, and spouse's income (if applicable). For (1) and (5), spouse's income reported at age 41 is included. For other equations, average spousal income over the 20 year period is included. The table reports robust (Huber-White) standard errors.

Exhibit C.6
OLS Estimation of Family Net Worth at Age 41 and In 2000 (Weighted)

	(1)			(2)		
	Family Net Worth at Age 41			Family Net Worth in 2000		
	<u>Coef.</u>	<u>Std. Err.</u>		<u>Coef.</u>	<u>Std. Err.</u>	
Years self-employed	5.73	1.92	***	4.79	2.05	**
Years employer-employed	-2.32	1.30	*	-2.08	1.64	
Female	-36.97	12.04	***	-68.54	12.72	***
Black	-62.84	7.55	***	-37.32	8.47	***
Hispanic	5.70	11.21		8.95	12.23	
AFQT	2.34	0.23	***	2.18	0.28	***
Parent's highest grade completed	16.08	2.33	***	12.33	2.72	***
Spouse income (in '000)	3.40	0.47	***	4.38	0.48	***
Constant	-171.73	34.65	***	-174.22	38.99	***
Number of observations	6,787			4,852		

*Source: NLSY79. Notes: *** Significant at 1 percent; ** Significant at 5 percent; * Significant at 10 percent. The models are estimated using the imputed net worth values for missing observations for (1) and (3), and using only those for whom net worth was reported for (2) and (4). Each equation contains dummy indicators for missing variables for AFQT, parents' highest grade completed, and spouse's income (if applicable). The table reports robust (Huber-White) standard errors.*

Exhibit C.5 presents the estimation results on the effects of self-employment years on income measures based on the full estimation sample of respondents. For these estimation models, we also included in the estimation the number of years that individuals were engaged employer-based jobs. Similarly to the estimations of the trajectory type effects above, all covariates, including years of self-employment, do not predict growth rate measures at all. However, the number of self-employment years does appear to have significant effects on respondents' level of income. Specifically, each additional year of self-employment, holding years of employer-based employment constant, increases the income level. The findings also suggest that the marginal effects of an additional year of self-employment are slightly higher than those for an additional year of employer-based employment. For instance, the results suggest that an additional year in self-employment increased the level of the family income by \$1,390 and own income by \$630, while an additional year of employer-based employment increased the family income by \$840 and own income by \$580. The results from the family net worth estimations are presented in Exhibit C.6. These results are consistent with those from the income estimations presented in Exhibit C.5. We found that an additional year of self-employment increased family net worth significantly, while an additional year in an employer-based job did not.