Self-employment, entrepreneurship, and the NLSY79

Researchers have used the rich data from the 1979 cohort of the National Longitudinal Survey of Youth to investigate the relationship between self-employment and various job and earnings outcomes; future inquiry may afford valuable insights into other interesting consequences of self-employment.

A relatively small, but growing, body of literature uses microdata from the 1979 National Longitudinal Survey of Youth (NLSY79) to study self-employment and entrepreneurship among young adults. The topics covered in these studies include, but are not limited to, the determinants of entrepreneurship, earnings growth among entrepreneurs, the returns to self-employment, the relationship between criminal activities and self-employment, and job satisfaction among the self-employed.

The NLSY79 is a nationally representative sample of 12,686 men and women who were between the ages of 14 and 22 when they were first interviewed in 1979.1 Survey respondents were interviewed annually from 1979 to 1994 and biannually starting in 1996. Most previous studies using this survey exclude the sample of 1,280 youths designed to represent the population enlisted in the four branches of the military as of September 30, 1978, but retain the supplemental sample of 5,295 civilian black, Hispanic, and economically disadvantaged nonblack, non-Hispanic youth. The NLSY79 contains a wealth of information on the demographic, economic, family background, educational, and psychological characteristics of respondents. Detailed measures of the group’s labor market and life experiences from early adulthood to the mid-forties can also be created for survey respondents.

The NLSY79 is an excellent source of data for conducting research on self-employment and entrepreneurship. The wealth of information available in the survey allows one to build rich empirical models of the entrepreneurial process. Measures of previous wage and salary, self-employment, and unemployment experience can be created, and the NLSY79 contains several uncommon variables, such as those associated with detailed asset categories, family background information, data on criminal activities, Armed Forces Qualification Test (AFQT) scores, and psychological characteristics. Furthermore, a plethora of measures of the dynamics of self-employment may be extracted from the longitudinal data in the survey. For example, measures of transitions to and from self-employment, number of years of self-employment, and whether an individual ever tries self-employment can easily be created. Finally, the returns to self-employment, measured as earnings, job satisfaction, net worth, or other outcomes, can be estimated. Changes over time in labor market status can be used to identify the effects of self-employment, potentially removing biases created by unobserved heterogeneity across individuals. Given these advantages, it is somewhat surprising that more researchers have not used the NLSY79 to study self-employment. In the sections that follow, this article presents estimates of self-employment from the NLSY79, reviews findings from previous studies that used the survey, and discusses some of the merits of the data sets making up the survey.
Self-employment in the NLSY79

In most previous studies using the NLSY79, self-employed workers are defined as those individuals who identify themselves as self-employed in their own business, professional practice, or farm in response to the class-of-worker question relating to the current or most recent job. Unpaid family workers are not counted as self-employed. Individuals who report being enrolled in school and workers who report working fewer than 300 hours in the previous calendar year are often excluded. The hours restriction rules out very small scale business activities.

Self-employment rates increase rapidly as the NLSY79 cohort ages. (See chart 1.) The self-employment rate is defined as the fraction of workers that is self-employed. At age 22, only 5.1 percent of men and 2.6 percent of women are self-employed. By age 42, however, 12.1 percent of men and 9.8 percent of women are self-employed.

The following tabulation shows that self-employment rates also differ substantially by race and its ethnicity:

<table>
<thead>
<tr>
<th>Race or ethnicity</th>
<th>Self-employment rate (percent)</th>
<th>Sample size</th>
<th>Self-employment rate (percent)</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>5.3</td>
<td>14,448</td>
<td>3.2</td>
<td>13,469</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7.4</td>
<td>10,153</td>
<td>4.9</td>
<td>8,404</td>
</tr>
<tr>
<td>White</td>
<td>10.1</td>
<td>31,803</td>
<td>6.9</td>
<td>29,006</td>
</tr>
</tbody>
</table>

As in previous studies, blacks and Hispanics are much less likely to be self-employed than are whites. Only 5.3 percent of black men are self-employed, compared with 10.1 percent of white men. The Hispanic male rate of 7.4 percent is also lower than the white rate, but higher than the black rate. Among women, the black-white and Hispanic-white self-employment rate ratios are similar to those for men. The main difference is that women’s self-employment rates are lower than men’s for all three racial and ethnic groups.

The determinants of self-employment. A few patterns are beginning to emerge in the young and expanding literature on self-employment. The empirical studies in this literature generally find that being male, white, older, married, and an immigrant and having a self-employed parent, more assets, and more education increase self-employment. In contrast, theoretical models of self-employment posit that attitudes toward risk, entrepreneurial ability, and preferences for autonomy are central to the individual’s decision to become self-employed or engage in wage and salary work. Perhaps not surprisingly, there is very little empirical evidence on the importance of these unobservable characteristics in the self-employment decision. One article that uses the NLSY79 offers indirect evidence by examining the relationship between drug dealing and legitimate self-employment: a review of ethnographic studies in the criminology literature indicates that drug dealing may serve as a useful proxy for a low aversion to risk, entrepreneurial ability, and a preference for autonomy.

The 1980 wave of the NLSY contains a special section on participation in illegal activities, including questions on selling marijuana and other “hard” drugs. The answers to these questions, together with data from subsequent years of the survey, are used to examine the relationship between drug dealing as a youth and legitimate self-employment in later years. Using various definitions of drug dealing and specifications of the econometric model, the survey finds drug dealers to be 11 percent to 21 percent more likely than those who are not drug dealers to choose self-employment, all else being equal. After ruling out a few alternative explanations, this article interprets these results as providing indirect evidence that aversion to risk, entrepreneurial ability, and preferences for autonomy are important determinants of self-employment.

In addition to offering detailed information on criminal activities in the 1980 wave, the NLSY79 includes information on whether respondents were interviewed in jail or prison in each year. This information is useful because convictions and incarcerations may have different effects on current and future wage and salary and self-employment earnings. In particular, ex-offenders who choose self-employment do not face discrimination, either pure or statistical, by employers in the labor market, but may face other forms of discrimination, such as that by consumers or lending institutions. Using the NLSY79, the aforementioned study by Robert Fairlie provides evidence on the relationship between incarceration and self-employment. Estimates from probit regressions indicate that having a previous incarceration increases the probability of self-employment by 0.36 percentage point to 0.39 percentage point, or 5.2 percent to 5.9 percent. Thus, self-employment may provide an important alternative to wage and salary work for at least some ex-convicts.

Another finding reported in Fairlie is that AFQT scores have a small and insignificant effect in probit regressions for the probability of self-employment.

Interestingly, previous research using the NLSY79 finds that AFQT scores have a large positive effect on earnings. The general argument is that the scores represent a measure of basic skills that help predict job performance. Although youths who have low levels of these basic skills may have limited opportunities in the wage and salary sector, that barrier does not translate into higher probabilities of self-employment.
The longitudinal nature of the NLSY79 also allows one to explore the effects of previous labor market experience on current self-employment. Ellen Rissman analyzes one aspect of the dynamic relationship between unemployment and self-employment among men.

She finds that the probability of being self-employed in the current year increases significantly if the person was unemployed in the previous year. Stratifying her sample by race, she also finds a positive and significant effect for whites, but not for nonwhites.

The dynamics of self-employment. Previous research on self-employment generally takes a point-in-time focus; longitudinal data in the NLSY79, however, allow for numerous dynamic measures of the concept. For example, Rissman finds that 3.4 percent of wage and salary workers in any given year become self-employed the following year and, conversely, 36.9 percent of the self-employed during a given year make the transition to wage and salary work the next year.

Also, Marianne A. Ferber and Jane Waldfogel find that 24.8 percent of men and 16.5 percent of women in their sample from the NLSY79 report ever being self-employed. By contrast, they find current self-employment rates of 8.8 percent and 5.5 percent for men and women, respectively. Finally, Donald Williams finds that, by 1987, just 3.1 percent of NLSY79 respondents had 2 or more years of self-employment experience and only 1.2 percent of respondents had 3 or more years of self-employment experience.

Although numerous possibilities exist for measuring self-employment dynamics, most previous research has focused on annual transitions to and from self-employment. Estimates for transition matrices that include wage and salary employment, self-employment, and nonemployment are reported in table 1. One-year transition matrices are reported for the 1979–94 period, and 2-year transitions are reported for the 1994–2002 period. Estimates from 1979–94 indicate that 3.4 percent of young men who were wage and salary workers became self-employed the following year. The entry rate for the nonemployed is 2.2 percent. Estimates for 2-year transitions from 1994–2002 indicate a lower self-employment rate from wage and salary work and a higher entry rate from nonemployment. For men, the exit rates from self-employment are 31.6 percent and 24.9 percent for the earlier and later periods, respectively. Self-employment entry rates are generally lower, and exit rates higher, for young women. Overall, the estimates indicate that substantial mobility exists between sectors and into and out of employment among young workers.

The importance of assets has taken center stage in the literature on the determinants of self-employment. Several
recent studies explore this issue by modeling the decision of wage and salary workers or other nonbusiness owners to switch into self-employment over a fixed period.\(^1\) The focus on transitions to self-employment attempts to avoid the endogeneity problem of including assets in a static model of self-employment. The problem is that a positive relationship found in a cross-sectional analysis may simply reflect the possibility that business owners accumulate more wealth instead of wealth increasing the likelihood of owning a business. Although individuals may save in anticipation of becoming self-employed, a measure of assets in year \(t - 1\) should be more exogenous to the entrepreneurial decision than a contemporaneous measure of assets.

Fairlie follows this approach, using net worth data from the NLSY79.\(^2\) Specifically, he estimates probit regressions for the probability of entry into self-employment from wage and salary work that include a measure of net worth. The NLSY did not collect information on assets prior to 1985 and in 1991. For other years, a measure of net worth can be created from the detailed NLSY questions on assets.\(^3\) The coefficients on net worth and its square are statistically significant and indicate a concave relationship. Evaluated at the mean level of net worth (which equals $36,900), the coefficients imply that increasing net worth by $10,000 raises the probability of a transition into self-employment by 0.00044. This percentage represents only 1.5 percent of the sample’s entry rate into self-employment. Thus, the estimates of the coefficients provide some evidence that young men face liquidity constraints, but these constraints do not appear to be overly restrictive.

Creating detailed measures of previous work experience to include in her regressions, Hiromi Taniguchi examines the determinants of transitions from nonemployment to self-employment and to wage and salary employment among women in the NLSY79.\(^4\) Her results indicate that both cumulative work experience and the number of jobs ever held increase the rate of entry into self-employment and wage/salary employment. She also finds that previous self-employment increases the rate of entry into self-employment and has a negative effect on entry into wage and salary employment.

### Table 1. Labor market transition matrices, National Longitudinal Survey of Youth, 1979–2002

[In percent]

<table>
<thead>
<tr>
<th>Gender, category, and year t</th>
<th>Nonemployment,(^1) Wage and salary employment,(^2) Self-employment,(^2) Share of total, year t</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979–94, year t:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonemployment</td>
<td>72.1</td>
<td>25.7</td>
</tr>
<tr>
<td>Wage and salary employment</td>
<td>1.0</td>
<td>95.5</td>
</tr>
<tr>
<td>Self-employment</td>
<td>.7</td>
<td>30.8</td>
</tr>
<tr>
<td>1994–2002, year t:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonemployment</td>
<td>65.6</td>
<td>29.1%</td>
</tr>
<tr>
<td>Wage and salary employment</td>
<td>2.8</td>
<td>94.0</td>
</tr>
<tr>
<td>Self-employment</td>
<td>2.0</td>
<td>22.8</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979–94, year t:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonemployment</td>
<td>75.4</td>
<td>21.4</td>
</tr>
<tr>
<td>Wage and salary employment</td>
<td>3.7</td>
<td>94.1</td>
</tr>
<tr>
<td>Self-employment</td>
<td>4.6</td>
<td>29.0</td>
</tr>
<tr>
<td>1994–2002, year t:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonemployment</td>
<td>68.7</td>
<td>28.1</td>
</tr>
<tr>
<td>Wage and salary employment</td>
<td>6.1</td>
<td>91.5</td>
</tr>
<tr>
<td>Self-employment</td>
<td>10.2</td>
<td>28.6</td>
</tr>
</tbody>
</table>

\(^1\) Those unemployed and not in the labor force.
\(^2\) Measured in year \(t + 1\) for 1979–94 and year \(t + 2\) for 1994–2002.

Note: The sample consists of youths aged 22 to 45 years who are not enrolled in school. All estimates are calculated with the use of annual sample weights provided by the National Longitudinal Survey of Youth.

---

Self-employment earnings in the NLSY79

Although self-employment income can be identified in the NLSY79, earnings among self-employed business owners are typically measured as total annual earnings, which are calculated by summing the responses to questions on military income, wage and salary income, and business or farm income (after expenses) in the previous calendar year. The income from all three sources is added because more than half of the self-employed with positive earnings in the NLSY79 report wage and salary income, but do not report business income.
This is partly due to incorporated business owners reporting their income as wage and salary income, as roughly 50 percent of unincorporated business owners with positive total earnings report zero business income. As suggested by Zagorsky, it may also be due to the ordering of questions on the questionnaire. Respondents were asked, (1) “How much money did you get from the military?”; (2) “Excluding military pay, how much money did you get from wages, salary, commissions, or tips?”; and (3) “Excluding anything you already mentioned, did you receive any business income?” Thus, some of the self-employed may have reported their income in the second question and did not correct their mistake. Another possibility is that the self-employed report only their labor income from the business under wage and salary income.

The following tabulation shows the mean, median, and standard deviation of total annual earnings for self-employed and wage and salary workers:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-employed</td>
<td>Wage and salary</td>
<td>Self-employed</td>
</tr>
<tr>
<td>Mean</td>
<td>$52,300</td>
<td>$28,217</td>
</tr>
<tr>
<td>Median</td>
<td>38,000</td>
<td>20,029</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>48,159</td>
<td>30,040</td>
</tr>
<tr>
<td>Sample size</td>
<td>3,725</td>
<td>1,570</td>
</tr>
</tbody>
</table>

Only full-time workers, defined here as those working at least 1,400 hours in the previous calendar year, are included, in order to control for differences in hours worked. Earnings observations in all years are inflated to 2002 dollars. The responses for each of the three sources of income (the self-employed, wage and salary workers, and the military) are top coded at $75,000 from 1979 to 1984, $100,000 from 1985 to 1994, and the top 2 percent for more recent years. Instead of these top codes, however, the 1994 top code, in 2002 dollars, $121,390, is used for all years in what follows.

As is customary, all top-coded values are set to $150,000.

For men, the self-employed earn substantially more, on average, than do wage and salary workers. Men’s average earnings are $14,042 higher among the self-employed, and median earnings are $4,979 higher. For women, average earnings among the self-employed are slightly higher than average earnings among wage and salary workers, but median earnings are lower. For both men and women, the standard deviation of self-employment income is substantially greater than that of wage and salary income.

Returns to capital. One issue that arises in comparing self-employment earnings with wage and salary earnings from survey data is the treatment of returns to capital. In the NLSY, the question regarding self-employment income asks, “How much did you receive after expenses?” from your farm or business in the past calendar year. Although some uncertainty is involved in answering this question, respondents are likely to interpret the question to include both the returns to labor and the returns to capital. As noted earlier, however, most of the self-employed report their earnings as wage and salary income and not business income. In the case of the respondent who does report self-employment income as business income, it would be preferable to remove the returns to capital before making comparisons with the earnings of wage and salary workers.

Unlike most other data sets, the NLSY79 contains detailed-enough information on assets to enable researchers to conduct a careful analysis of the issue of how returns to capital are treated. The NLSY79 contains data on the market value of the individual’s farm, business, and other real estate and the total amount of debt owed on those assets. This information can be used to calculate the opportunity cost of capital and remove it from business income. With Standard & Poor’s 500 as the alternative investment, adjusted self-employment earnings are 5.2 percent lower than unadjusted self-employment earnings for white men and 4.0 percent lower for white women. Simple adjustments for the opportunity cost of capital also have a small effect on self-employment earnings for blacks and Latinos. Overall, estimates from the NLSY79 suggest that unadjusted self-employment earnings from survey data may provide reasonably accurate measures of the returns to labor.

Earnings regressions. As mentioned earlier, the NLSY79 contains detailed information on individual characteristics such as age, race, education, AFQT scores, and various measures of previous work experience. Earnings regressions that include these observable controls can be used to estimate the between self-employment earnings and wage and salary earnings. Unobserved differences, however, such as entrepreneurial ability and aversion to risk, may also exist between self-employed business owners and wage and salary workers. To address this issue, fixed-effects earnings regressions can be estimated with the longitudinal data in the NLSY79. The individual-level fixed effects control for all observable and unobservable characteristics that do not change over time. Because, over time, individuals make transitions between self-employment, on the one hand, and wage and salary work, on the other, comparisons of self-employment earnings with wage and salary earnings for the same individual in different years contribute to identifying the associated coefficients.

Estimating fixed-effects earnings regressions for young men from disadvantaged families yields some evidence that self-employed business owners earn more than do wage and salary workers.
provide some evidence of lower earnings among self-employed business owners than among wage and salary workers. The results from these earnings comparisons are somewhat sensitive to the use of different measures of income and different econometric models.

In a related study, Justin van der Sluis, Mirjam van Praag and Arjen van Witteloostuijn (2004) estimate the returns to education for entrepreneurs and for wage and salary workers. Using instrumental-variable regressions, they find that the returns to education are 14 percent for the self-employed, much higher than the 10-percent estimated return for wage and salary workers. The detailed data available in the NLSY79 allow these researchers to control for ability and to use family background characteristics, including the mother’s and father’s education, the presence of library cards in the household at age 14, and magazines present in the household at age 14, as instrumental variables for education.

**Earnings profiles.** The longitudinal nature of the NLSY79 enables one to compare earnings profiles for self-employed workers and wage and salary workers. Charts 2 and 3 display earnings-age profiles for full-time self-employed and wage and salary workers. For men (chart 2), average self-employment earnings are always higher and appear to grow at a rate similar to that of wage and salary earnings. For women, average self-employment earnings start out lower than wage and salary earnings, but then grow at a faster rate.

To investigate the question of whether the self-employed experience faster earnings growth than do wage and salary workers, the NLSY79 allows fixed-effects regressions that include interactions between self-employment, on the one hand, and experience, potential experience, or tenure, on the other, to be estimated. Estimating fixed-effects regressions for hourly earnings for a sample of white, non-Hispanic men, Daiji Kawaguchi finds flatter earnings-experience/tenure profiles for self-employed workers than for wage and salary workers. At 10 years of experience and job tenure, self-employed business owners earn 18 percent less than wage and salary workers. An earlier work by Fairlie compares men’s and women’s earnings profiles for whites, blacks, and Hispanics. For white men, the point estimates from these earnings regressions indicate that the self-employed initially experience lower earnings growth than do wage and salary workers. After several years, the trend reverses, and self-employed persons experience faster earnings growth and higher earnings. For Hispanic men, the relative self-employment earnings coefficients suggest that the self-employed start at much lower earnings levels than do wage and salary workers, but experience faster growth rates. For white women, relative self-employment earnings start out positive and then become negative. Relative self-employment earnings coefficients are not statistically significant for black men, black women, or Hispanic women, possibly due to small sample sizes.

**Self-employment and other outcomes.** The detailed information available in the NLSY79 also allows for the analysis of the relationship between self-employment and other outcomes, such as future wage and salary income, job satisfaction, and net worth. One possibility is to examine the relationship between early-career self-employment experience and future labor market outcomes. The NLSY79 is an excellent instrument for this type of analysis because it follows individuals from ages 14 to 22 in 1979 to 37 to 45 in 2002.

A previously mentioned work by Fairlie examines the earnings patterns of less educated individuals who are self-employed early in their careers and makes comparisons with young, less educated wage and salary workers. Self-employment status is determined between ages 22 and 26, and earnings are measured starting at age 27. Estimates from fixed-effects regressions indicate that the self-employed experience faster earnings growth, on average, than do wage and salary workers after a few initial years of slower growth. In a similar vein, Williams examines the relationship between self-employment at ages 16 to 20 and outcomes at ages 25 and 27. He finds that self-employment as a youth is associated with a substantially higher probability of being self-employed in early adulthood (age 27), but also is associated with lower earnings at that age.

Another interesting question that can be answered with the NLSY79 is whether self-employment experience is rewarded in the wage and salary sector. Do self-employment spells limit opportunities for acquiring valuable labor market experience, especially firm- and sector-specific human capital, or do they provide workers with skills that are rewarded in the wage and salary sector? Again using data from the NLSY79, as well as data from the National Longitudinal Survey (NLS) of Young Women, Williams examines the effects of self-employment experience on future wage and salary earnings of men and women. His estimates indicate a negative return for women and little or no return for men.

Most of the focus in the self-employment literature is on earnings, but other outcomes also are of interest. In particular, lower hourly earnings among the self-employed with high levels of tenure may be explained by nonpecuniary factors of the job, such as being one’s own boss. Kawaguchi uses the NLSY79 to investigate whether self-employment is associated with higher levels of job satisfaction. He finds that 65 percent of the self-employed report liking their job “very much,” whereas only 45 percent of wage and salary workers report that level of job satisfaction. Estimates from regression models which control for individual heterogeneity confirm that the self-employed have higher levels of job satisfaction than wage and salary workers have.

Chart 3. Average annual earnings by age for women, NLSY (1979–2002)
RESEARCH USING THE NLSY79 has undoubtedly improved our understanding of the determinants of entrepreneurship, the dynamic process of self-employment, and self-employent earnings patterns. Although the relationship between self-employment and a few outcomes, such as future wage and salary earnings and job satisfaction, has been explored with the exceptionally rich data available in the NLSY79, more research may provide valuable insights into the consequences of self-employment. For example, the detailed data available in the survey allow one to explore the causal relationship between self-employment and several outcomes of interest, such as net worth, business net worth, health insurance and other fringe benefits, and public assistance programs.

Notes

1 See Center for Human Resource Research, NLSY79 Users’ Guide (Columbus, OH, The Ohio State University, 1999), for a detailed description of the NLSY79.

2 Estimates from the NLSY79 are comparable to those from 1990 census microdata using a similar age group. See Robert W. Fairlie, “Does Business Ownership Provide a Source of Upward Mobility for Blacks and Hispanics?” in Douglas Holtz-Eakin, ed., Entrepreneurship and Public Policy (Cambridge, MA, MIT Press, 2004), pp. 634–59. The census shows slightly lower rates, but the relative differences between the races are similar.


5 Ibid.

6 Ibid.


8 Ibid.


11 All estimates are calculated with annual sample weights provided by the NLSY. “Nonemployment” denotes those not in the labor force.


14 The variable having to do with assets is not available in the public-use data, but can be obtained from Jay L. Zagorsky at the Center for Human Resource Research. See Jay L. Zagorsky, “Young Baby Boomers’ Wealth,” working paper (Columbus, OH, Center for Human Resource Research, 1998), for more details on the construction of this variable.


16 Fairlie, “Does Business Ownership Provide a Source of Upward Mobility for Blacks and Hispanics?”

17 Telephone conversation, August 1999.

18 Especially problematic is the fact that 36 individuals have top-coded wage and salary income of more than $4 million each in 1996.

19 Fairlie, “Earnings Growth.”


21 Justin Van der Sluis, Mirjam van Praag, and Arjen van Witteloostuijn, “Comparing the Returns to Education for Entrepreneurs and Employees,” working paper (Amsterdam, University of Amsterdam, 2004).


23 Fairlie, “Does Business Ownership Provide a Source of Upward Mobility for Blacks and Hispanics?”

24 Fairlie, “Earnings Growth.”

25 Williams, “Youth Self-Employment.”
