
# The long term returns of attempting self-employment with regular employment as a fall back option

Moira Daly

PII: S0927-5371(15)00039-1
DOI: doi: 10.1016/j.labeco.2015.03.013
Reference: LABECO 1391

To appear in: *Labour Economics*

Received date: 9 July 2013
Revised date: 22 March 2015
Accepted date: 25 March 2015

Please cite this article as: Daly, Moira, The long term returns of attempting self-employment with regular employment as a fall back option, *Labour Economics* (2015), doi: 10.1016/j.labeco.2015.03.013

# The long term returns of attempting self-employment with regular employment as a fall back option

Moira Daly

*Copenhagen Business School, Department of Economics, Porcelaenshaven 16A, DK-2000 Frederiksberg, Denmark*

*moda.eco@cbs.dk*

*+45 26 20 24 88*

**Abstract**

There is a substantial body of research investigating the returns to self-employment. Relatively little attention has been paid to the returns from attempting self-employment while acknowledging that the decision to try self-employment is reversible. But this is the option considered by the worker deciding whether to become self-employed, as are any resulting positive or negative changes to the worker's wages, hours worked, or likelihood of job or business termination and resulting unemployment. The full consequences of attempting self-employment are determined by comparing the actual income streams of individuals who do and who do not attempt self-employment. Selection on observables is controlled by employing nearest-neighbor matching with bias correction. The main result is that there is no significant evidence that individuals who attempt self-employment are punished for doing so: after 15 years, those who attempt self-employment receive an (insignificant) 8% and a (significant) 22% premium in labor income and in labor and asset income, respectively. The consequences of attempting self-employment vary by occupation: individuals in technical and professional occupations realize significant gains, of 45% to 62% after 15 years, whereas craftsmen see no significant differences in income. Quantile treatment effects on the treated are estimated and reveal that the average positive premiums are driven by the upper tail of the treatment effect distribution.

*Key words:* Self-employment, Entry, Matching
*JEL:* J23, J24

## 1. Introduction

In 2009, 15.3 million individuals, or about 1 in 9 workers in the U.S., were self-employed, according to the Bureau of Labor Statistics (Hipple, 2010). More than 1 in 5 American men are estimated to have tried self-employment over the period from 1979 to 1993, although the majority of these spells were brief: between two-thirds and three-quarters of them lasted a year or less (Williams, 2000).

The significant portion of the labor force that either is self-employed or has tried self-employment is driven in part by government programs designed to promote exactly this. Tax incentives are given to sole proprietors, and small businesses are exempted from certain regulations.[1] The Self-Employment Assistance program (SEA) encourages the unemployed to become self-employed by waiving the work-search requirement when the applicant is spending their time establishing a business. In 2012, the U.S. Labor Department announced that an additional $35 million would be directed toward improving the SEA.[2] The Small Business Administration (SBA), whose mission is "to maintain and strengthen the nation's economy by enabling the establishment and vitality of small businesses," was recently given the go-ahead, via the Small Business Jobs Act of 2010, to increase the lending support available to small business owners to more than $12 billion and to provide up to $50 million in grants to Small Business Development Centers across the country. This is in addition to the counseling the SBA already provided. The new programs are designed to increase self-employment by easing informational and financial barriers in an effort to spur economic growth and provide individuals with a route out of poverty.

Although self-employment is often used as a proxy for entrepreneurship, the choice to work for oneself is important in its own right, as are its consequences.[3] The criterion for eligibility for many of the programs above is self-employment, where no distinction is made between what Baumol (2011) calls "replicative" entrepreneurs, who make up the majority of the self-employed, and "innovative" entrepreneurs.[4] A substantial portion of the labor force attempts self-employment, which motivates its study, according to Hamilton (2000), who tries to determine which labor-market

[1] See Blanchflower & Oswald (1998), Bruce & Schuetze (2004)

[2] See US Labor Department press release number: 12-1073-NAT

[3] See Andersen & Nielsen (2012) for a survey of the large literature that uses self-employment as a proxy for entrepreneurship. Several recent papers have documented the fact that self-employment is not a good proxy for entrepreneurial activity. Hurst & Pugsley (2010) find that the majority of small businesses do not bring new ideas to the market and do not intend to innovate or grow.

[4] Faggio & Silva (2012) provide a nice overview of the self-employment literature as well as the research on the distinction between entrepreneurship and self-employment.

model best reconciles the observed returns from attempting self-employment. If self-employment is encouraged as a pathway out of poverty, the pecuniary rewards ought to be positive. By using OLS and quantile regression, Hamilton finds that the returns of self-employment are substantially lower than those of remaining an employee, and he concludes that those who try self-employment have strong preferences for such non-pecuniary benefits as "being your own boss."

Bruce and Schuetze (2004) point out that any evaluation of the costs and benefits of the above government programs must examine the consequences to the individual attempting self-employment. If individuals who attempt self-employment are penalized for doing so, then encouraging the unemployed to become self-employed rather than to find a job may not be good policy. By using the PSID, Bruce & Schuetze estimate, via pooled OLS, the effects of short spells of self-employment (between one and four years) on wages among individuals who began in and returned to full-time wage or salary employment. They find that short spells of self-employment reduce average hourly earnings.

The important question asked by both studies can be answered in a more comprehensive manner. The consequences of attempting self-employment are not reflected solely in the difference in wages between the employed and the self-employed, which is the focus of Hamilton (2000). Nor are they sufficiently captured by considering only the effects on the employee's wages after a spell of self-employment has ended, as this ignores the presumably higher wages of the successfully self-employed individual, a problem Bruce and Schuetze (2004) discuss. The reality lies between these two cases: some of those who attempt self-employment remain self-employed, some return to the wage-and-salary sector, and some become unemployed.[5] The decision to try self-employment is reversible: workers can return to paid employment should their enterprises fail. This option is taken into consideration by the worker who is deciding whether to become self-employed, as are any resulting positive or negative effects on subsequent wages. That is, in determining what returns he should expect from attempting self-employment, the worker must consider the probability that he will stay self-employed and the wage he can expect to make while self-employed, as well as the probability that he will return to a paid job and the wage he would receive then.[6]

[5] As shown in the appendix, only 56% of the sample remained self-employed in the year following the attempt. Of the rest, 27% returned to regular employment, 15% started regular jobs in addition to self-employment, and 2% are not working.

[6] Williams (2000) and Bruce & Schuetze (2004) both provide excellent discussions of this, although both focus on subsequent wages in the regular-employment sector.

As pointed out by Kahn and Lang (1992) and by Martinez-Granado (2005), most workers who are not self-employed face binding hours constraints: they would like to work more hours but cannot. Self-employment may provide a means of avoiding these constraints. Differences in the likelihoods of job or business termination, in the likelihood of unemployment resulting from this, and in the length of unemployment that results from an attempt at self-employment are also incorporated into the individual's decision-making process.[7] Finally, as pointed out by Hamilton (2000), Moskowitz and Vissing-Jørgensen (2002), and Rosen and Willen (2002), the decision to try self-employment is based on both the return to human capital and the return to capital. The method employed in this paper incorporates all of these margins into the calculation of the present discounted value of the future income that an individual considers in deciding whether or not to attempt self-employment.

Some individuals may choose self-employment simply in order to be their own bosses, even in the face of monetary loss. Conversely, impatient individuals may opt for regular employment if the returns from self-employment would take years to materialize, despite how large these returns might be. The main objective of this paper is to compare the ex-ante present values of the future incomes of individuals who attempt self-employment (ASE) and individuals who do not (PE), in order to reveal to what extent the desire for self-employment is driven by either the anticipation of pecuniary rewards or of non-pecuniary benefits.

Although the decision to attempt self-employment is driven in part by average returns, the average effect alone does not fully capture either the returns of the typical ASE or the risk of undertaking such an endeavor. For this reason, the distributions of returns for the ASE and the PE are compared. If individuals who attempt self-employment are in fact punished for doing so, the efficacy of programs such as the SEA can be called into question. If, on the other hand, the effect on the future present discounted value of income from attempted self-employment is non-negative, then the SEA's goal of encouraging self-employment as an alternative to traditional paid work is a reasonable one.

The Panel Study of Income Dynamics (PSID) is used to compare, at the worker level, the income streams of those who do and those who do not attempt self-employment. Specifically, the income streams of individuals who try self-employment are compared to appropriate alternative income streams to establish whether or not those who are attempting self-employment are being paid at a premium. The measurement of income streams used allows for sector switching, spells of unemployment,

[7] See for instance Phillips & Kirchoff (1989), who find evidence that the self-employed face a higher probablity of involuntary job termination.

and variable hours, as well as any subsequent effects on wages. Rather than parametrically estimating the income equations and the probability of sectoral switching, the realized income streams of those who have tried self-employment are observed and compared to the income streams of those who have never tried self-employment. Because actual income streams are used, the full consequences of trying self-employment are better captured: all of the extensive margins and wage effects that a worker considers in his self-employment decision are incorporated alongside any changes in the covariances among them.

Two measures of income are considered in the creation of the present values: annual total labor income and a measure that captures both total labor income and net asset income. The effect of attempting self-employment on the present discounted value of the individual's income is further unpacked by considering separately the annual effect on income, wages, and hours worked.

Selection on observables across the ASE and PE groups is controlled using Abadie and Imbens's (2011) method of matching with bias correction. One advantage of any matching approach is that it forces the researcher to acknowledge issues of balance between the treated and control groups—here, the ASE and the PE—that might lead to extrapolation using traditional regression techniques. The method of Abadie and Imbens, in particular, provides a bias correction for imperfect matching. The more comprehensive the list of control variables, the more likely unconfoundedness is to hold: all variables that are correlated either with the decision to try self-employment or with the outcome variable should be included. The basic control variables used by Hamilton (2000) and by Bruce and Schuetze (2004) are augmented with a rich set of (ex-ante) variables that are correlated both with selection into self-employment and with income. These include industry and occupation indicators, wages, hours worked annually, home ownership status, the self-employment status of the individual's father, whether someone in the family owns a business, and the working status of the individual's wife. Because many of these additional variables are imbalanced across the ASE and PE groups, their inclusion and the attention paid to balancing the sample implies that the resultant estimates will be less biased.

The results reveal that on average, those who attempt self-employment are not punished for doing so. After 15 years, individuals who have attempted self-employment realize about 8% higher present discounted values in their total labor income, although this difference is insignificant. However, the gains in the present discounted value of total labor and asset income are 22% higher for the ASE than for the PE, and this value is significant. Estimates of the quantile treatment effects on the treated reveal that these positive returns are being driven by the upper tail of the distribution.

In order to better understand the mechanism behind the differences in the present discounted values that result from attempting self-employment, the determinants of the present value, wages, and annual hours worked, are also considered over the same 15-year period. The clear driver behind the gains is a significant increase in the number of hours worked annually: the ASE individual worked around 300 more hours per year following the attempt at self-employment. Wages are either not significantly affected by the self-employment attempt or are positively affected, depending on the measure used. Keep in mind that the wage effect discussed here is the average effect on subsequent wages of trying self-employment, and is not conditional on remaining self-employed or even employed. Taken together, these results suggest that individuals who attempt self-employment may do so for pecuniary as well as non-pecuniary benefits. Attempting self-employment does not, on average, diminish subsequent wages, it does open up an avenue by which individuals can achieve higher earnings: working more.

The results are not driven by selection. If the assumption of unconfoundedness holds, no significant differences should exist in income, annual hours, or wages between the ASE individuals and their matches prior to the self-employment attempt. And, reassuringly, no significant evidence of trending behavior or selection was found in the two years prior to the self-employment attempt, regardless of the outcome variable used.

As suggested by Williams (2000), among others, the returns on attempted self-employment may vary by occupation, in which case the average effects may mask important differences. Moreover, the goal here is to understand the present-value comparison that the individual makes in deciding on self-employment, and it is reasonable to expect the individual to consider the experiences of others in the same occupation. Ideally, the above analysis would be replicated for each occupational group. However, given the limited sample sizes of the PSID, only the two most prevalent occupations among those attempting self-employment are considered separately: those with technical or professional occupations, and craftsmen. The results suggest that the returns from attempted self-employment are indeed heterogeneous with respect to occupation. Individuals coming from professional and technical fields enjoy substantial and significant returns from self-employment. After 15 years, they realize significant gains, of 45% and 62% in the present discounted values of labor income and of labor and asset income, respectively. Again, an analysis of the distribution of the treatment effects reveals that the upper tail of the distribution drives these gains. The average increase in the present discounted value of income from attempting self-employment is a product of significant increases in both wages and work hours. No evidence for selection is found.

Craftsmen on the other hand see losses, albeit insignificant ones, when labor income is considered. They lose about 6% of the 15-year present discounted value of the labor income that they would have received had they not attempted self-employment. This average loss reflects a treatment effect distribution which lies primarily below 0, at least from the 10th to the 90th percentile. When asset as well as labor income is included, however, craftsmen are rewarded for attempting self-employment, realizing a gain of about 13% after 15 years, though again the effect is insignificant.

The main result of the paper is that there is no significant evidence that those who attempt self-employment are punished for doing so. On the contrary, the rewards to labor and asset income are positive regardless of the occupational sample one considers. However, only individuals in the technical and professional occupations see significantly higher labor income. When all occupations are considered together, the increases in the income of the ASE over those of the PE are driven primarily by increases in the number of hours worked. This suggests that one motivation for seeking self-employment may be to escape binding hours constraints present in the paid employment sector.[8] The ASE individuals whose prior occupations were technical or professional earn substantially more than their PE counterparts because they work longer hours for higher wages. In both cases, pecuniary and non-pecuniary motives may play a role in the decision to try self-employment. On the other hand, those ASE who were previously craftsmen do not receive significant rewards or losses from attempting self-employment. Knowledge of such heterogeneity across occupations may be helpful to policymakers in determining the candidates most likely to benefit from aid, as in the SEA program.

To our knowledge, this is the first study to estimate the longer-term consequences of attempting self-employment on the present discounted value of future income, wages, and hours using the realized paths of these variables with such a comprehensive list of control variables.[9][10]

A number of robustness checks are performed. First, in order to bolster the claim

[8] Parker, Belghitar, & Barmby (2005) find evidence that the self-employed insure themselves in response to greater uncerntainty by working longer hours, which may be another explanation.

[9] Rosen and Willen (2002) develop an elegant structural model to explain the worker's choice between self-employment and non-self-employment income streams, while allowing workers to smooth consumption over time by investing in a risky asset. However, they condition on those who attempt and persist in self-employment for at least five years, i.e. the successfully self-employed.

[10] Berglann, Moen, Røed, and Skogstrøm (2011) use nearest-neighbor matching on a Norwegian employer-employee matched administrative dataset to look at the short-term effects on income of having tried entrepreneurship during the five-year period from 2001 to 2005.

of selection on observables, the local average treatment effect (LATE) of attempting self-employment is estimated using an instrumental variable approach. Second, the radius-matching estimator with regression adjustment proposed by Lechner, Miguel, and Wunsch (2011) is used to investigate the finite properties of Abadie and Imbens's matching estimator. Third, in order to understand how sensitive the results presented are to the omission of a confounding variable, the sensitivity analysis proposed by Ichino, Mealli, and Nannicini (2008) is implemented. Fourth, the importance of the individual's wife's asset income to the labor and asset income measure is investigated.

The paper proceeds as follows: Section 2 provides a review of the methods to be used; Section 3 presents the data, the definitions of self-employment and wage or salary employment, the outcome and match variables to be considered, and summary statistics; Section 4 presents the main results; Section 5 presents the robustness checks; and Section 6 presents our conclusions.

## 2. Method

The first goal is to compare the income streams of individuals who try self-employment to appropriate alternative income streams—those they would have had if they had not tried self-employment—in order to establish whether they are being are paid a premium. The measure of income streams should allow for sectoral switching, spells of unemployment, variable hours, and any effects on subsequent wages. Rather than the income equations and the probability of sectoral switching being parametrically estimated, the realized income streams of those who have tried self-employment are observed and compared with the income streams of those who have not. This way, the worker's optimal (or forced) decision to change sectors incorporates the realized hours worked, the ability to work multiple jobs, any wage consequences or any changes in the covariances among these variables—that is, all the factors that a worker would incorporate into a realistic decision regarding self-employment. Specifically, the interest here is in understanding the difference between the actualized present discounted value of income following the decision to try self-employment and the value the income would have had if the worker had not tried self-employment. Call the potential present discounted value of income resulting from trying and not trying self-employment $Y_i(1)$ and $Y_i(0)$ respectively:

$$Y_i(W_i) = \sum_{t=1}^{J} R^{-t} y_{it}(W_i)$$

where $W_i$ is a binary variable indicating whether one has tried self-employment at time $t = 1$, $R$ is the discount rate, $J$ is the horizon over which we take the present

value, $y_{it}(W_i)$ is the annual income at time $t$ for a worker who chose $W = W_i$. For a worker who attempts self-employment, $y_{i2}(1)$ would correspond to the individual's income one year after having tried self-employment, *regardless of whether or not he chose to remain in the self-employment sector*.[11] The individual level premium is the treatment effect, $Y_i(1) - Y_i(0)$, the difference between the present value of income of the individual who tried self-employment less that of the present value of income of his twin. The goal is to estimate the ATT:

$$ATT = E(Y_i(1) - Y_i(0)|W_i = 1)$$

After investigating the effects on total income, this measure is disaggregated in order to see what drives differences in the total income for those who have and have not tried self-employment. ATTs are also calculated for annual income, hours worked and wage.

The classic problem is that $Y_i(1)$ is observed only if the worker chooses self-employment and $Y_i(0)$ is observed otherwise, necessitating the construction of counterfactuals. If a linear relationship between the outcomes and the observable and unobservable variables is assumed, and the distribution of observable variables is similar for those who have tried self-employment and those who have not, then traditional parametric regression techniques are appropriate as a way to control for selection on observables. However, if that distribution is different between the two groups, the construction of counterfactuals via traditional regression techniques will rely more heavily on extrapolation. Even if the distributions are similar, the assumption of linearity may not be desirable or realistic. In that case, a method based on matching would make sense.[12] Justification for the assumption of unconfoundedness will rely heavily on the quality of the matching variables, a point which will be addressed in the data section.

The matching estimator chosen here is the nonparametric bias-adjusted estimator of Abadie and Imbens (2011). It is particularly robust in comparison to propensity score methods. Each treated observation is matched with one or several control observations according to the Mahalanobis distance of the observables. In practice, this matching may not be exact. Importantly, as opposed to the more commonly used simple-nearest-neighbor matching, which estimates the ATT as the difference in means of the outcome variables in the matched dataset (ignoring the lack of perfect

[11] The online appendix presents results for those who remained self-employed for two, five, and ten years. As expected, those who remain self-employed generally do better than the pool of all individuals who attempt self-employment.

[12] See Imbens (2004) for a nice overview.

matching), the Abadie and Imbens (2011) method uses local linear regression to correct the bias that results from any mismatches.[13] It also allows for the calculation of robust (to heterogeneous treatment effects) standard errors and exact matching.[14]

In order to identify the ATT, the following assumptions must be made:

1. Unconfoundedness for controls (selection on observables), $W$ is independent of $Y(0)$ conditional on observable covariates $X = x$;
2. Overlap, $Pr(W = 1|X = x) < 1 - \eta$, for $\eta > 0$ must hold for all $X$.

Under these two assumptions, the ATT is identified.[15] A propensity score, $Pr(W = 1|X = x)$, will be estimated using a logit model and those observations whose propensity scores are not in the region of common support will be dropped.

To investigate the degree to which the assumption of unconfoundedness is met, the effects of trying self-employment on the outcome variables prior to the date at which self-employment is attempted will be presented along with the ATTs. These effects should be insignificant if no selection on observables is occurring. The robustness section also gives the results of implementing Frölich's (2007) nonparametric IV estimation and Ichino, Mealli, and Nannicini's (2006) sensitivity analysis.

## 3. The Data

The PSID is a nationally representative longitudinal survey that began in 1968 with approximately 4,500 families. The survey questions individuals every year and tracks the various family members as they age. The starting population consists of non-student males who are not part of the survey of economic opportunity (SEO) sample, a part of the core sample that oversamples low-income households. The remaining sample is representative of the U.S. population in 1968 and is often used in the earnings literature for longitudinal studies; see, for example, Hryshko (2012) and Guvenen (2009).[16] The sample is restricted to heads of household, as defined by the PSID, as the detailed information needed to conduct the analys is only available for those individuals. Individuals whose occupation is farming are dropped. Individuals are asked "Do you work for someone else, yourself, or what?" Possible answers to this

[13] The disadvantage of this estimator is that it is not efficient.

[14] For these reasons, this method was chosen over similar estimators, such as radius matching with bias adjustment.

[15] These assumptions are also needed to identify Firpo (2007)'s quantile treatment effect on the treated along with an assumption for the uniqueness of quantiles, see Firpo (2007) for details.

[16] A nice reference to the PSID design is Hill (1992)

question include "someone else," "self," and "both."[17] Like Williams (2000), Bruce and Schuetze (2004), Rosen and Willen (2002), and Hurst, Li, and Pugsley (2010), this question is used to identify the self-employed.[18]

We are interested in the effects of switching from paid employment to self-employment. The initial self-employed sample is defined as those individuals who answered "self" for the first time between 1971 and 1981 and who were then between the ages of 18 and 45, as we would like to follow individuals for at least 15 years after the decision to try self-employment.[19] Although information is available from 1968, we are concerned with the transition into self-employment, so the earliest year in which treatment could occur is 1969. In fact, we set our cutoff at 1970 to allow the use of some matching variables that were not available earlier. Since 1996, income data has been collected only biannually. The control group consists of those who never answered "self" during the sample period. If after becoming self-employed, an individual returns to working for an employer, this individual remains part of the attempted self-employment sample (i.e., the treatment group).

Ideally, the present discounted values used in the calculation of the treatment effects would be taken with respect to lifetime income, but the data limits our ability to do this. Instead, we focus on the individuals who were observed for at least 16 consecutive years. Among those who have attempted self-employment, at least one year prior to the self-employment attempt must also be available, so that matching variables can be constructed. The subsequent years of observations are then used to calculate the present discounted value for the first 5, 10, and 15 years ($J = 5, 10, 15$ in the above notation). Individuals who never attempted self-employment and who were observed for at least 16 years are also included. The present values of income over the next $J$ years are calculated using a 2% discount rate, chosen to align with Hamilton's value.

[17] In the robustness section, the PSID and SIPP from which Hamilton's conclusions are drawn are compared in order to ensure that the different conlcusions drawn here do not stem from a difference in sample construction or in earnings measures.

[18] As opposed to Moskowitz & Vissing-Jorgensen (2002), and to Heaton & Lucas (2000), who focus on business owners. Neither group, self-employed or business owners, is a subset of the other in the PSID. This may be due in part to the fact that the self-employment question is asked at the individual level, but the business-ownership question is asked at the family level.

[19] An earlier draft of this paper also considered 20-year present values, hence the age limit of 45 years. (It was not pursued here because of the small sample sizes.) Including the 45-50 group at the time of SE, or equivalently including those who are 60-65 at the end of the period, does not change the results: only 9 observations are added to the ASE group. Our sample then covers the ages 18-60, in comparison to, for example, Levine & Rubinstein (2013), who use prime-age workers between 25 and 55.

Both annual total labor income and annual total labor and asset income are used in the calculation of the present values. All values are reported in 1984 dollars.[20] Annual total labor income includes the head's wages, bonuses, overtime pay, commissions, professional practice income, and the head's labor part of income from roomers and boarders, farming, and unincorporated businesses. The annual total labor and asset income is the sum of (1) the annual total labor income of the head, (2) the head's income from rent, interest, dividends, and alimony, and (3) the head's and wife's net asset income from farming, from unincorporated businesses, and from roomers and boarders. Ideally, (3) would include only the head's portion of the net asset income.[21] Starting in 1985, the PSID did collect this data separately for the head and wife. In the robustness section, it is shown that the net asset income from the wife is extremely small. All results are presented for both measures.

In order to determine more precisely the mechanism underlying any differences in present values across the PE and ASE groups, annual differences in labor income, labor-plus-net-asset income, annual hours worked, and two wage measures are also considered as outcome variables. The first of these, "wage," is the total labor income divided by the total hours worked annually. The second, "wage plus," is calculated as labor and asset income divided by hours worked annually.

The more comprehensive the list of matching variables, the more likely unconfoundedness is to hold. Matching variables should not be affected by treatment status, and for this reason all matching variables have been defined prior to the year in which self-employment was attempted. Variables that affect the participation decision and the outcome variable should be included; Stuart (2010) advises that variables should be included liberally, on the grounds that the omission of a confounding variable will have large effects on bias, but the inclusion of variables that are not correlated with the outcome would produce only a slight increase in variance. A starting point in the selection of covariates is the set of standard control variables, such as those used by Hamilton (2000): the individual's education, labor market experience, tenure, marital status, and race, and whether the individual is part of a union, is a veteran, or is disabled. This list clearly does not include all

[20] Again, this was chose to allign with Hamilton's figures.

[21] The labor and (net) asset income measure used is the PSID variable "taxable income of the head and wife" less the wife's labor income and certian types of wives' asset incomes. As pointed out by an anonymous referee, gains to the total family income may be due to issues of assortative matching. To avoid such issues, the labor and asset income measure is used. Total family income has been used in the study of self-employment by Rosen & Willen (2002) and by Hurst, Li, & Pugsley (2010), and it could be used to back out a similar measure, but also includes transfer income.

the variables that meet the above criteria, though. A brief discussion of additional relevant covariates is presented below.

Dunn and Holtz-Eakin (2000) find that the self-employment experience of an individual's father increases that individual's probability of self-employment. Levine and Rubinstein (2013) find that the self-employed who own incorporated businesses see significant positive returns. An indicator for the self-employment status of the father is included, along with indicators for whether someone in the family owns an incorporated or unincorporated business.

Several studies find that wealth affects the decision to try self-employment by relaxing capital market constraints; see, for example, Evans and Jovanovic (1989), Quadrini (2000), Blanchflower and Oswald (1998), and Dunn and Holtz-Eakin (2000). Unfortunately, no direct measure of individual wealth is available during the sample period. Bracke, Hilber, and Silva (2013) find that purchasing a home reduces the probability of becoming self-employed in professional, technical, or management fields or in a position with dependent workers. To capture these effects, an indicator for home ownership is included along with information on the value of the house and other family income (i.e., income not received by the head). A variable for the receipt of a lump-sum payment is also included, in the spirit of Blanchflower and Oswald (1998).[22] Variables are included capturing whether the wife works and, if she does, her income.

Lazear (2005) points out that a small business owner must be a "jack of all trades." Indicators for first occupation are therefore included, which, taken together with current occupation, capture the specificity or generality of work experience. A variable is also included indicating whether the individual has had "a number of different kinds of jobs," "different jobs, mostly [in] the same occupation," or is "on first job now," again because it captures generality or specificity of skill acquisition. Folta, Delmar, and Wennberg (2010), using Swedish data, define "hybrid entrepreneurship" as working simultaneously for oneself and as a paid employee. They find that this is a significant pathway into self-employment.[23] If an individual answered "both" to the question "Do you work for someone else, yourself, or what?" they are considered both a self-employed individual and a paid employee, and an indicator for this dual status is set to 1.

Fairlie and Meyer (1994) and Clark and Drinkwater (2000) find that racial, ethnic, and religious background characteristics significantly determine self-employment in

[22] The reason for these payments is unknown. That is, we can't say whether it was from an inheritance, a lottery, a pension payout, etc.

[23] They find that over 21% of all self-employment entries are preceded by hybrid activity.

the U.S. and among British immigrants, respectively. An indicator for nonwhite status is therefore included, as are indicators for religion; answers are Baptist, other Protestant, Catholic, Jewish, and "not available, don't know, or refuse to answer."[24] Faggio and Silva (2012) used British data to find that self-employment is positively related to gross firm creation and incidence of innovation in urban areas but not in rural areas. City size indicators and indicators for region of the U.S. are thus included. An individual's occupation and industry have also been found to affect both the decision to try self-employment and the success of the attempt; see Le (1999) for a nice review of empirical studies of self-employment.[25] Lastly, variables describing attributes of the individual's prior job and their attitude toward it are included: wage, annual working hours, whether the individual has a second job, and whether the individual is looking for another job.[26]

Table 1 provides summary statistics for the variables used in the matching process and for various outcome variables. The first three columns display the summary statistics for the entire sample prior to matching. The first column gives the average value of the variable in the period prior to attempting self-employment for the ASE. The second column gives the corresponding value for the PE. The third column gives the standardized differences and the fourth column the significance from a t-test.[27]

Those who attempt self-employment are significantly more likely to be white, highly educated, and single. They are less likely to be in a union. They make about \$1.25 less per hour but work almost 200 hours more per year. They have approximately two years less tenure at their current job and about one year less experience. The ASE are more than twice as likely to have a father who is self-employed: 9.3% have self-employed fathers, compared to 4.1% of the PE. Forty-five percent of those who attempt self-employment state that they or someone in their family owns a business, compared to 3% of those who do not attempt self-employment and about 44% of those businesses are incorporated. Those who attempt self-employment are also almost twice as likely to have received a lump-sum payment,

[24] More detail was available in some years, but due to changing questions over the years, this was the most specific level of detail that was consistently available. Unfortunately, ethnicity is not available during the sample period, for the same reason.

[25] The 1968-1980 Retrospective Occupation-Industry Files put out by the PSID were used to align industry and occupation defintions over time.

[26] Hamilton (2000) found positive selection into self-employment, although studies by Evans & Jovanovic (1989) and by Blanchflower and Meyer (1992) found evidence of negative selection.

[27] The standardized difference is calculated as $SD = \frac{(\overline{x}_{ASE} - \overline{x}_{PE})}{\sqrt{.5*(v_{ASE}+v_{PR})}}$

**Table 1:** Balance of full and matched samples

| | Full Sample | | | | Matched Sample | | | |
|---|---|---|---|---|---|---|---|---|
| | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* |
| Father self-employed | 0.093 | 0.041 | 0.148 | *** | 0.102 | 0.089 | 0.033 | |
| Other family Income | 10.42 | 8.22 | 0.153 | *** | 10.20 | 8.45 | 0.116 | * |
| Own House | 0.579 | 0.640 | -0.090 | | 0.575 | 0.612 | -0.054 | |
| Own House*house value | 40.60 | 42.19 | -0.023 | | 39.74 | 42.26 | -0.037 | |
| Married | 0.821 | 0.874 | -0.104 | * | 0.827 | 0.890 | -0.128 | * |
| Married*wife works | 0.507 | 0.558 | -0.073 | | 0.512 | 0.589 | -0.109 | |
| Married*wife works*wife's labor income | 4.99 | 5.81 | -0.073 | | 4.83 | 5.46 | -0.058 | |
| Nonwhite | 0.021 | 0.083 | -0.196 | *** | 0.024 | 0.053 | -0.109 | |
| College graduate | 0.314 | 0.238 | 0.121 | ** | 0.307 | 0.295 | 0.018 | |
| High school graduate | 0.329 | 0.349 | -0.030 | | 0.323 | 0.309 | 0.021 | |
| High school dropout | 0.100 | 0.159 | -0.124 | * | 0.102 | 0.089 | 0.033 | |
| Number of children | 1.221 | 1.426 | -0.107 | * | 1.197 | 1.455 | -0.140 | ** |
| Veteran | 0.314 | 0.350 | -0.053 | | 0.315 | 0.327 | -0.018 | |
| Disabled | 0.050 | 0.041 | 0.032 | | 0.055 | 0.018 | 0.141 | ** |
| Received lump sum | 0.114 | 0.065 | 0.123 | ** | 0.118 | 0.077 | 0.099 | |
| **Size of Largest City** | | | | | | | | |
| 500,000+ | 0.264 | 0.241 | 0.038 | | 0.244 | 0.280 | -0.057 | |
| 100,000-499,999 | 0.271 | 0.257 | 0.024 | | 0.268 | 0.215 | 0.088 | |
| 10,000-99,999 | 0.336 | 0.365 | -0.043 | | 0.362 | 0.356 | 0.009 | |
| <10,000, non continental | 0.129 | 0.138 | -0.019 | | 0.126 | 0.150 | -0.048 | |
| **Region of U.S.** | | | | | | | | |
| Northeast | 0.221 | 0.193 | 0.050 | | 0.197 | 0.191 | 0.011 | |
| North central | 0.307 | 0.351 | -0.066 | | 0.323 | 0.309 | 0.021 | |
| South | 0.279 | 0.297 | -0.029 | | 0.283 | 0.297 | -0.021 | |
| West | 0.193 | 0.159 | 0.063 | | 0.197 | 0.203 | -0.010 | |
| **Religion** | | | | | | | | |
| Not available, don't know, refuse, other | 0.164 | 0.118 | 0.094 | * | 0.150 | 0.134 | 0.032 | |
| Baptist | 0.107 | 0.200 | -0.184 | *** | 0.110 | 0.161 | -0.106 | |
| Other Protestant | 0.443 | 0.436 | 0.010 | | 0.472 | 0.453 | 0.028 | |
| Catholic | 0.221 | 0.228 | -0.011 | | 0.220 | 0.215 | 0.010 | |
| Jewish | 0.064 | 0.018 | 0.164 | *** | 0.047 | 0.037 | 0.034 | |
| Union | 0.200 | 0.329 | -0.209 | *** | 0.197 | 0.291 | -0.156 | ** |
| Working both as SE and PE | 0.179 | 0.011 | 0.419 | *** | 0.126 | 0.126 | 0.000 | |
| Working a second job | 0.271 | 0.228 | 0.071 | | 0.276 | 0.346 | -0.108 | |
| Looking for another job | 0.200 | 0.091 | 0.220 | *** | 0.197 | 0.193 | 0.007 | |
| Family incorporated business | 0.200 | 0.012 | 0.453 | *** | 0.157 | 0.152 | 0.012 | |
| Family unincorporated business | 0.250 | 0.025 | 0.487 | *** | 0.236 | 0.244 | -0.013 | |
| Annual hours worked | 2388 | 2203 | 0.218 | *** | 2378 | 2332 | 0.051 | |
| Wage | 12.12 | 13.38 | -0.113 | ** | 12.31 | 13.26 | -0.085 | |
| Tenure | 3.56 | 5.52 | -0.281 | *** | 3.71 | 5.00 | -0.194 | *** |
| Experience | 10.06 | 11.34 | -0.134 | ** | 10.02 | 10.54 | -0.057 | |

**Table 1 Continued:** Balance of full and matched samples

| | Full Sample | | | | Matched Sample | | | |
|---|---|---|---|---|---|---|---|---|
| | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* |
| **First Occupation** | | | | | | | | |
| None, never had one | 0.021 | 0.012 | 0.051 | | 0.016 | 0.014 | 0.012 | |
| Professional, technical and kindred | 0.136 | 0.124 | 0.025 | | 0.134 | 0.146 | -0.024 | |
| Managers, officials and proprietors | 0.036 | 0.033 | 0.011 | | 0.031 | 0.022 | 0.043 | |
| Clerical and sales | 0.179 | 0.118 | 0.120 | ** | 0.181 | 0.104 | 0.156 | ** |
| Craftsmen, foremen, and kindred | 0.171 | 0.107 | 0.132 | ** | 0.173 | 0.096 | 0.159 | ** |
| Operatives and kindred workers | 0.171 | 0.214 | -0.076 | | 0.165 | 0.181 | -0.029 | |
| Laborers and service workers | 0.214 | 0.268 | -0.089 | | 0.220 | 0.311 | -0.145 | ** |
| Miscellaneous | 0.071 | 0.124 | -0.125 | * | 0.079 | 0.126 | -0.110 | |
| **History of Changing Occupation and/or Job** | | | | | | | | |
| Number of different kinds of jobs | 0.357 | 0.409 | -0.076 | | 0.362 | 0.360 | 0.003 | |
| Different jobs, mostly same occ. | 0.064 | 0.097 | -0.085 | | 0.055 | 0.057 | -0.006 | |
| Mostly same occupation | 0.486 | 0.442 | 0.062 | | 0.488 | 0.496 | -0.011 | |
| Not available, inappropriate, on first job now | 0.093 | 0.052 | 0.112 | ** | 0.094 | 0.087 | 0.019 | |
| **Occupation** | | | | | | | | |
| Craftsmen | 0.236 | 0.248 | -0.020 | | 0.244 | 0.283 | -0.063 | |
| Service | 0.036 | 0.040 | -0.016 | | 0.039 | 0.035 | 0.015 | |
| Not available, don't know | 0.007 | 0.012 | -0.037 | | 0.008 | 0.000 | 0.089 | ** |
| Laborers | 0.036 | 0.048 | -0.043 | | 0.039 | 0.043 | -0.014 | |
| Managers/Admisintration | 0.193 | 0.115 | 0.153 | *** | 0.173 | 0.163 | 0.019 | |
| Operatives | 0.050 | 0.151 | -0.240 | *** | 0.055 | 0.079 | -0.067 | |
| Professional/Technical | 0.236 | 0.230 | 0.009 | | 0.236 | 0.242 | -0.010 | |
| Sales/ Clerical | 0.164 | 0.099 | 0.137 | ** | 0.157 | 0.089 | 0.149 | ** |
| **Industry** | | | | | | | | |
| Agriculture | 0.021 | 0.011 | 0.059 | | 0.024 | 0.002 | 0.136 | *** |
| Business and Repair Services | 0.050 | 0.023 | 0.100 | ** | 0.047 | 0.018 | 0.118 | * |
| Construction | 0.179 | 0.063 | 0.254 | *** | 0.165 | 0.126 | 0.079 | |
| Entertainment | 0.007 | 0.006 | 0.011 | | 0.008 | 0.018 | -0.062 | |
| FIRE | 0.100 | 0.031 | 0.197 | *** | 0.102 | 0.037 | 0.181 | *** |
| Inappropriate | 0.007 | 0.009 | -0.014 | | 0.008 | 0.006 | 0.017 | |
| Missing | 0.014 | 0.010 | 0.025 | | 0.016 | 0.016 | 0.000 | |
| Manufacturing | 0.171 | 0.364 | -0.315 | *** | 0.189 | 0.167 | 0.040 | |
| Mining | 0.007 | 0.013 | -0.043 | | 0.008 | 0.004 | 0.036 | |
| Personal Services | 0.014 | 0.006 | 0.056 | | 0.016 | 0.010 | 0.037 | |
| Professional Services | 0.150 | 0.109 | 0.086 | | 0.142 | 0.132 | 0.020 | |
| Public Admistration | 0.021 | 0.127 | -0.291 | *** | 0.024 | 0.106 | -0.240 | *** |
| Transportation | 0.057 | 0.106 | -0.127 | * | 0.063 | 0.148 | -0.196 | ** |
| Wholesale and Retail Trade | 0.200 | 0.120 | 0.155 | *** | 0.189 | 0.211 | -0.038 | |

**Table 1 Continued:** Balance of full and matched samples

| | Full Sample | | | | Matched Sample | | | |
|---|---|---|---|---|---|---|---|---|
| | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* | *Attempt* | *No Attempt* | *Stand. Diff.* | *Significance of T-test* |
| **Year Dummies** | | | | | | | | |
| 1970 | 0.129 | 0.068 | 0.146 | *** | 0.102 | 0.073 | 0.074 | |
| 1971 | 0.043 | 0.070 | -0.083 | | 0.039 | 0.053 | -0.046 | |
| 1972 | 0.086 | 0.072 | 0.036 | | 0.087 | 0.087 | 0.000 | |
| 1973 | 0.086 | 0.075 | 0.027 | | 0.087 | 0.089 | -0.005 | |
| 1974 | 0.093 | 0.081 | 0.029 | | 0.094 | 0.116 | -0.050 | |
| 1975 | 0.114 | 0.083 | 0.075 | | 0.110 | 0.085 | 0.061 | |
| 1976 | 0.107 | 0.086 | 0.051 | | 0.118 | 0.142 | -0.050 | |
| 1977 | 0.071 | 0.090 | -0.048 | | 0.079 | 0.055 | 0.067 | |
| 1978 | 0.086 | 0.090 | -0.010 | | 0.094 | 0.108 | -0.032 | |
| 1979 | 0.107 | 0.092 | 0.035 | | 0.102 | 0.094 | 0.019 | |
| 1980 | 0.050 | 0.096 | -0.126 | * | 0.055 | 0.059 | -0.012 | |
| 1981 | 0.029 | 0.098 | -0.203 | *** | 0.031 | 0.039 | -0.030 | |
| **Outcome Variables** | | | | | | | | |
| Annual hours worked | 2388 | 2203 | 0.218 | *** | 2378 | 2332 | 0.051 | |
| Wage | 12.12 | 13.38 | -0.113 | ** | 12.31 | 13.26 | -0.085 | |
| Wage plus | 13.65 | 13.64 | 0.001 | | 13.79 | 13.85 | -0.005 | |
| Labor income | 27.75 | 28.88 | -0.045 | | 27.87 | 30.27 | -0.093 | |
| Labor plus asset income | 31.42 | 29.45 | 0.071 | | 31.45 | 31.62 | -0.006 | |
| **Observations** | **140** | **5404** | | | **127** | **4077** | | |

*Notes:* All monetary figures are presented in thousands of 1984 dollars, except the wage which is in 1984 dollars. "Working both as SE and PE" is an indicator variable that takes the value 1 if the individual worked both as an employee and as self-employed prior to becoming solely self-employed. Total observations used in estimation shown, though the matched sample contains 4 times the number of treated observations. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

11.4% to 6.5%. The ASE have on average $2,200 more in other family income than the PE.

Approximately 18% of those who attempt self-employment do so through the hybrid mechanism described above, a proportion that is much larger than the 1% among those who never attempt self-employment. The ASE are also much more likely to be looking for another job: 20% compared to 9%, a significant difference. There are significant differences in religious affiliation between the ASE and PE groups, with the ASE being three times as likely to be Jewish, about half as likely to be Baptist, and significantly more likely to have refused to answer or to have answered "other."

There are noticeable differences in the distribution of workers across industries. For instance, a self-employed individual is about three times as likely to be in construction or in finance, insurance, and real estate and almost twice as likely to work in wholesale and retail trade or in business and repair services. On the other hand, the ASE are only half as likely to work in manufacturing or in transportation, and only 2% of the ASE work in public administration, as opposed to about 13% of the PE.

Differences in occupation are also present: those who attempt self-employment are significantly more likely to have been managers or administrators, 19% as opposed to 12%, or in sales, 16% to 10%. At the same time, the ASE are a third less likely to have worked as operatives. The two most common occupations among both the ASE and the PE are technical or professional careers and craftsmen, with almost a quarter of the ASE and PE engaging in each. Nine percent of the ASE are on their first job or don't know how many times they have changed their occupation, compared to 5% of the PE.[28]

The average rate at which individuals who have not tried self-employment attempt to do so (not shown in table) is 2.67% overall from 1971 to 1982 as compared to the 2.0% for newly self-employed individuals from 1973 to 1992 reported by Quadrini (2000). This difference could be due simply to the change in time period: Evans and Leighton (1989) found that self-employment rates peaked in 1983 and declined thereafter.[29]

[28] For both industry and occupation, between 1% and 2% of individuals have missing information or else report not knowing their industry or occupation. These records are not dropped because it could be that the fact of not knowing the industry or occupation, or of having missing information, itself reveals information about the type of worker that could be valuable in the matching process.

[29] Levine & Rubinstein (2013) find 1.4%, but among 25- to 55-year olds during the period 1994-2010, but this is hard to compare given the difference in the time periods considered.

**Table 2:** Propensity score distributions

| | Entire Sample | | Common Support | |
|---|---|---|---|---|
| | *Attempt* | *No Attempt* | *Attempt* | *No Attempt* |
| Smallest | 0.0020 | 0.0002 | 0.0020 | 0.0020 |
| 1st Percentile | 0.0023 | 0.0002 | 0.0023 | 0.0020 |
| 5th Percentile | 0.0047 | 0.0004 | 0.0044 | 0.0023 |
| 10th Percentile | 0.0099 | 0.0008 | 0.0075 | 0.0028 |
| 25th Percentile | 0.0374 | 0.0020 | 0.0266 | 0.0044 |
| 50th Percentile | 0.1315 | 0.0055 | 0.0945 | 0.0086 |
| 75th Percentile | 0.4909 | 0.0153 | 0.3998 | 0.0208 |
| 90th Percentile | 0.7821 | 0.0398 | 0.5964 | 0.0510 |
| 95th Percentile | 0.8897 | 0.0714 | 0.6896 | 0.0904 |
| 99th Percentile | 0.9733 | 0.2507 | 0.7701 | 0.2834 |
| Largest | 0.9771 | 0.7957 | 0.7941 | 0.7957 |
| Observations | 140 | 5,404 | 127 | 4077 |

*Notes:* The columns labeled "Attempt"/"No Attempt" display the propensity score distributions for those that attempt self-employment and those who do not, respectively.

### *3.1. Creating the matched samples*

A logit is used to estimate the probability that an individual will attempt self-employment—the propensity score—the results of which are presented in the online appendix. The intuition gleaned from Table 1 carries through to the logit estimation.

Table 2 presents the propensity score distributions for the entire sample and for the sample trimmed such that the maximum propensity score for the PE group is no less than the maximum propensity score of the ASE group and such that the minimum propensity score of the PE group is no less than the minimum propensity score of the ASE group, i.e. the common support sample. Note that by imposing this common support assumption, we also satisfy Assumption 2 above, overlap, because the maximum propensity in either group is less than 0.80. The resulting trimmed sample contains about 91% of the total ASE group (those with the lowest propensity scores) and about 75% of the PE group (those with the highest propensity scores). The first two columns of Table 2, which display the propensity score distributions for the full sample, clearly show the positive skewness of the distributions for both the ASE and the PE, although the skewness is more extreme for the latter. It is important to note that trimming the sample inherently alters the population of

inference. However, when the main results (the ATTs on income to be presented in Table 3) were calculated without the common support assumption, the estimates produced had the same sign and overall significance.[30]

Abadie and Imbens's (2011) matching method with bias correction is implemented on the common support sample. Following Imbens's (2007) suggestion, we choose four nearest-neighbor matches for both the estimation of the treatment effects and the calculation of robust standard errors.[31] Importantly, the method corrects the bias created by imperfect matching by using local linear regression to adjust the estimates. That is, the method is a combination of both matching and regression adjustment.[32] To achieve better balancing, the method allows more weight to be put on particular variables in the matching process. Following the suggestion of Stuart (2010), when issues of imbalance on a few covariates were being addressed, stronger weight was given to the self-employment status of the father, whether someone in the family owned a business, and whether the individual was both self-employed and regularly employed.[33] A better balance can be achieved by requiring a smaller number of matches, but this improvement is made at the expense of robustness. Using just one match yielded better balance and frequently stronger (more significant) results; nonetheless, we use four matches.

Columns 4 and 5 of Table 1 present the means of the covariates for the ASE and PE groups respectively for the matched sample. As expected, the means of the ASE group have changed very little relative to the full sample, and the small change is due only to the imposition of common support. In order to evaluate the quality of the matching process, Rosenbaum and Rubin (1983) recommend calculating the standardized difference/standardized bias before and after matching. The standardized differences pre- and post-matching are displayed in columns 3 and 6 respectively.

[30] The extreme skewness of the distribution of the propensity scores is worrying because a few control observations can have a large effect on estimates. The main results are also calculated from the sample containing individuals whose propensity score is less than 0.25. These are presented in appendix Table A4. The sign and significance of the estimates remains unchanged.

[31] Throughout the paper, we state that we calculate the ATTs. When looking at the standard errors of the estimates, however, it is important to note that we are estimating the sample average treatment effect on the treated, the SATT. This procedure was implemented using nnmatch in Stata; see Abadie, Drukker, Leber Herr, & Imbens (2004).

[32] See Malmendier & Tate (2009) for an application of the method with a similar number of treated observations.

[33] Additional variables were given higher weight depending on the specification in order to achieve better balance. The year in which self-employment was attempted was always given higher weight. In the robustness section, a radius-matching method with bias correction is used that yields very similar results without putting heavier weights on particular variables.

Absolute values greater than 0.25 should signal caution; see Rubin (2001) and Stuart (2010). Seven of the variables in the full sample exceed this threshold: working as SE and PE, family incorporated business, family unincorporated business, tenure, working in the construction industry, working in the manufacturing industry, and working in public administration. Seven more variables have a standardized difference greater than or equal to 0.20 in absolute value but less than 0.25: nonwhite, union, looking for another job, annual hours worked, working in the FIRE industry, being an operative, and the 1981 year dummy. Interestingly, several of these variables are often omitted from studies of the returns from self-employment experience. Specifically, industry, occupation, annual hours worked, and family ownership of a business are not included in the studies of Hamilton (2000), Bruce and Schuetze (2004), and Williams (2000).[34] The matching process very much improves the balance in the sample: none of the variables has a standardized difference greater than 0.25 in absolute value. This implies that regression adjustment will correct the bias from the lack of balance, which is the advantage of using the Abadie and Imbens (2011) method with bias correction.

T-tests can also be used as a measure of balance; see Caliendo and Kopeinig (2008). As the authors point out, this method masks the change in bias between the full and matched samples. On the other hand, this method does yield significance levels. However, Stuart (2010) argues that t-tests should not be used to evaluate balance for two reasons: first, balance is an in-sample property; second, hypothesis tests confound changes in balance with changes in power. Nonetheless, significance levels are shown in columns 4 and 8 for the sample pre- and post-matching, respectively. The matching process decreases the number of significant differences by about half. Note that given the number of t-tests performed, some significant results would be expected by chance.

Matching was also performed on two sub-samples: a sample consisting of only the individuals who were engaged in technical or professional occupations, and a sample containing those who were craftsmen in the year prior to the self-employment attempt. Generally, the matched samples did not show as much of an improvement in balance relative to the full sample, no doubt due to the imposition of exact matching on occupation.

Table A2 presents the balancing information for the sample consisting of only the individuals who were engaged in technical or professional occupations. Prior to matching, the sample is very imbalanced: 15 of the covariates have a standardized

[34] Williams does include interactions of self-employment experience with measures of occupation that are defined differently from those presented here.

difference greater than 0.25 in absolute value, while 17 display significant differences across the ASE and PE groups. Post-matching, the sample is more balanced: only 6 variables remain with standardized differences greater than 0.25, and only 10 exhibit significant differences. In both the full and matched samples, significantly more of the ASE report that they are looking for another job. In the matched sample, 40% of the ASE report doing so, against only 11.7% of the PE. The resulting standardized difference of 0.477 is large but still considerably lower than the 0.555 standardized difference in the full sample. The ASE also have approximately 2.8 years of tenure compared with the 4.1 years of the PSE, resulting in a standardized difference of –0.268, again markedly better than the –0.464 of the unmatched sample. Ten percent of those who do not attempt self-employment describe themselves as having had different jobs in the same occupations, whereas none of the ASE do, resulting in a standardized difference of –0.332, close in magnitude to that of the unmatched sample. Among the PE, 11.7% report working in public administration, as opposed to none of the ASE, giving a standardized difference of –0.362, down from –0.439 in the full sample. In addition, 6.7% of the PE work in wholesale and retail trades, while none of the ASE do, resulting in a standardized difference of –0.266. Finally, in 1974 and 1980, no individuals attempted self-employment; this contrasts with 9% and 10% of the PE observations that were used as matches, resulting in standardized differences of –0.306 and –0.329 respectively.

Despite the model's inability to balance these variables, large improvements were made in the balance of home ownership, home value, number of children, family ownership of businesses, likelihood that the individual was Jewish, amount of experience, and likelihood that the individual was working in the manufacturing or professional services. Not including these variables would clearly have created an omitted-variable bias in the estimate of the returns to attempting self-employment. Admittedly, the larger percentage of ASE than PE who are looking for a new job is worrisome. However, an alternative matched sample that uses only one match is presented in the online appendix. The "looking for another job" variable is exactly matched, there, and that specification yields only four significant differences among covariates and also yields larger and more significant estimates than the main results presented in the paper. This suggests that the bias correction in Abadie and Imbens's (2011) method is effective in spite of the seemingly worse balance.

Table A3 presents the balancing information for the craftsmen sub-sample. Prior to matching, 11 variables in the craftsmen sub-sample have standardized differences greater than 0.25, and 17 are significantly different across the ASE and PE groups. Post-matching, only 6 variables have a standardized difference greater than 0.25, and only 7 are significantly different across the groups. Specifically, in the matched

craftsmen sample, the ASE are three times as likely to be Catholic, and they are significantly less likely to be in the transportation industry or to report their first occupation as "miscellaneous." Only 32.3% of the ASE are unionized, as opposed to 50% of the PE. Finally, the ASE have more family income.

Regardless of the sample considered, the severe imbalance that was present among the covariates that are, arguably, the most important for selection into the ASE group: whether or not the individual was both self-employed and regularly employed prior to becoming ASE, and whether or not a family member of the individual owned a business, has been removed. As discussed above, "better" balancing is possible with fewer matches, and even delivers more significant estimates, but comes at the cost of bias. It is important to recognize that a simple matching procedure would be inappropriate with these balance characteristics, which is why the nearest-neighbor-with-bias adjustment is used instead.

## 4. Results

### *4.1. Average treatment effect on the treated*

#### *4.1.1. Discounted value of future income*

Panel B of Table 3 presents the estimates of the ATT as calculated from the present discounted value of both labor income alone, in columns 1 through 3, and labor and asset income together, in columns 4 through 6. The average treatment effects on the treated calculated from all occupations are reported in columns 1 and 4. The ATTs calculated from individuals who were in technical and professional occupations prior to attempting self-employment are shown in columns 2 and 5, and the ATTs from those who were craftsmen prior to attempting self-employment are shown in columns 3 and 6.

Consider first column 1, "All," of Panel B. The row labeled "5 years" presents the ATT calculated from the present discounted value of labor income for the 5 years following the attempt at self-employment (including the year in which that occurred) for all occupations. After 5 years, those who attempt self-employment can expect to make \$5,200 more than their twins, or a 3.4% relative gain over what they would have earned had they never attempted self-employment.[35] This is an insignificant difference. The row labeled "10 years" shows that those who attempt self-employment can expect the difference in the total present discounted value of labor income after 10 years to be \$14,700, a relative gain of 4.9%. After 15 years, that gain increases to \$33,400 a 7.6% relative gain. These effects are not significant.

[35] The percentage gain or loss shown in Table 5 is calculated as $\frac{\widehat{ATT}}{\overline{Y}(0)}$

Column 4 displays the ATTs calculated from the present discounted value of both labor and asset income for all occupations. After 5 years, the ASE can expect to make $28,500, or 17.3% more than what they would have made as an employee, which is a significant difference. The ASE continue to fare significantly better as the time horizon extends to 10 and 15 years, at which point the ASE are making 19.8% more and 21.9% more, respectively, than they would have as paid employees.

To evaluate the degree to which the assumption of unconfoundedness is met, Panel A of Table 3 presents the differences in annual income measures in the year prior to the attempt at self-employment and in the year prior to that. These are in the rows labeled "$t = 0$", "$t = -1$". Column 1 displays the differences in labor income and column 4 displays the differences in labor and asset income. There is no evidence of trending behavior or positive selection, regardless of which income measure is used. If anything, the estimates suggest negative selection in the year prior to attempting self-employment, though they are not significant. These results agree with the negative selection (measured by the effect on prior wages) found by Bruce and Schuetze (2004), and they contradict the positive selection into self-employment found by Hamilton (2000).

This first look across all occupations reveals that the labor income of those who attempt self-employment is not diminished as a result. In fact, if both asset and labor income are considered, the ASE are unambiguously rewarded for trying self-employment.[36] Furthermore, the rewards for attempting self-employment continue to grow with the time horizon, ruling out a transient boost. These effects are not a by-product of selection.

Substantial heterogeneity may underlie the ATTs calculated from all occupations. To investigate this possibility, the ATT is calculated for those whose occupation in the period prior to attempting self-employment is technical or professional and for those employed as craftsmen. These two occupations were chosen because they are the two most prevalent ASE occupations, each representing about 24% of the ASE group.[37]

Column 2 of Panel B in Table 3 displays the ATTs calculated from the present discounted value of labor income for those whose occupation is technical or profes-

[36] Hamilton (2000) also finds that his equity-adjusted draw measure is larger than the draw measure alone and that it overtakes PE wages, although the net result on the present value of earnings of being in business is still 5% less than what the worker could have earned had they remained in PE.

[37] In a world with larger sample sizes, the manager-administrator occupation group would be interesting to explore, but unfortunately this is not feasible with this sample.

**Table 3:** Average treatment effects on the treated, present discounted value of future income

| | | Labor Income | | | Labor and Asset Income | | |
|---|---|---|---|---|---|---|---|
| | | *All* | *Technical, Professional* | *Craftsmen* | *All* | *Technical, Professional* | *Craftsmen* |
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| | | *(A) Difference in annual measures before attempting self-employment* | | | | | |
| *t=-1* | Difference | 0·8 | 4·5 | −0·7 | 1·2 | 4·3 | 0·4 |
| | Standard Error | (1·5) | (3·3) | (2·4) | (1·7) | (3·5) | (2·5) |
| | % Gain or Loss | 2·6 | 13·8 | −2·8 | 3·9 | 12·5 | 1·6 |
| | Number treated | 115 | 27 | 28 | 115 | 27 | 28 |
| *t=0* | Difference | −2·4 | −2·9 | −2·3 | −0·2 | 0·7 | 0·0 |
| | Standard Error | (1·5) | (3·2) | (2·1) | (1·7) | (3·5) | (2·2) |
| | % Gain or Loss | −7·9 | −8·7 | −9·3 | −0·5 | 2·1 | 0·0 |
| | Number treated | 127 | 30 | 31 | 127 | 30 | 31 |
| | | *(B) Difference in present values after attempting self-employment* | | | | | |
| *5 years* | Difference | 5·2 | 48·8 | −16·0 | 28·5** | 78·2** | 11·9 |
| | Standard Error | (11·7) | (35·3) | (10·2) | (13·0) | (37·4) | (13·1) |
| | % Gain or Loss | 3·4 | 27·9 | −11·6 | 17·3 | 43·6 | 8·5 |
| *10 years* | Difference | 14·7 | 135·9* | −25·4 | 63·4** | 210·0*** | 20·2 |
| | Standard Error | (22·5) | (72·8) | (17·7) | (25·9) | (81·9) | (23·7) |
| | % Gain or Loss | 4·9 | 38·4 | −9·5 | 19·8 | 57·2 | 7·5 |
| *15 years* | Difference | 33·4 | 238·8* | −24·2 | 102·6** | 342·0** | 49·0 |
| | Standard Error | (36·2) | (127·7) | (23·9) | (41·5) | (138·4) | (33·6) |
| | % Gain or Loss | 7·6 | 45·1 | −6·4 | 21·9 | 61·9 | 12·8 |
| | Treated | 127 | 30 | 31 | 127 | 30 | 31 |
| | Total | 4,204 | 1,034 | 1,082 | 4,204 | 1,034 | 1,082 |

*Notes:* Robust standard errors shown in parenthesis, 4 matches used. All monetary figures presented in thousands of 1984 dollars. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

sional. After 5 years, those who attempt self-employment are rewarded for doing so, realizing gains of about 28% relative to their twins, although this effect is not significant. After 10 years, the ASE realize a large and significant gain of 38%, and this increases to 45% at the 15-year horizon. Column 5 in the same panel reveals that these gains are even more substantial if asset income is included: the ASE make relative gains of about 44%, 57%, and 70% at the 5-, 10-, and 15-year horizons respectively, all of which are statistically significant. These results suggest that the effects of attempting self-employment for this group are not temporary but persist and grow at least until the 15-year horizon. Columns 2 and 5 of Panel A verify that these gains are not being created by positive selection.

Craftsmen, on the other hand, see losses (although insignificant ones) when labor income is considered. Column 3 of Panel B shows that craftsmen who attempt self-employment are punished for doing so when labor income alone is considered. They lose about 12%, 10%, and 6% of the present discounted value of the labor income they would have received at the 5-, 10-, and 15-year horizons had they not attempted self-employment, although the effects are never significant. Column 3 of Panel A shows that these craftsmen were not doing well prior to trying self-employment: the ASE were making about 9% less than their twins. The matching process did not effectively get rid of all selection. However, when asset income is included, as shown in column 6 of Panel B, craftsmen are rewarded for attempting self-employment, realizing a reward of about 13% after 15 years, although the effect is again insignificant. Column 6 of Panel A provides evidence that no selection is at play when the labor–plus-asset-income measure is used.

Taken together, the results in the "15 years" row in Panel B of Table 3 suggest that there is no significant evidence that those who attempt self-employment are punished for doing so. On the contrary, the rewards to labor and asset income are positive regardless of what occupational sample one considers: after 15 years, the ASE make a significant, 21.9% premium over what they would have earned without self-employment, and those ASE who were in technical or professional occupations realize gains of more than 60%. After 15 years, only the individuals in the technical and professional occupations see significantly higher labor income, realizing gains of approximately 45%; the group of craftsmen alone and the group of all ASE pooled together each see premiums that are not significantly different from 0.

Panel A of Table 3 shows that the matching method succeeded, for the most part, in removing any issues of selection. One exception is the case of craftsmen when labor income is considered. This explains, to some extent, the negative point estimates attained for this group. It remains to be determined whether these effects are driven by changes in wages or in hours. In the next section we attempt to answer

this question.

*4.1.2. Wages, Hours and Annual Income*

The total effect of attempting self-employment on the present value of income is driven by changes in wages, in the number of hours worked annually and in the covariances among these. To determine the importance of changes in hours and wages, the ATTs are calculated for wages (= labor income/annual hours worked), wages plus (= labor and asset income/annual hours), labor income, labor-and-asset income, and annual hours worked, for the two years prior to the attempt at self-employment ($t = 0; -1$) and for each year thereafter. Note that not all individuals had a positive number of hours worked every year, hence the wage was not available for all years. The number of such individuals, however, was rather small, and was effectively 0 for the ASE group. Figure 1 presents these ATTs for all occupations along with their 90% confidence intervals. The same information is given in tabular format (along with sample sizes for the wage variables) in the online appendix.

Figure 1 reveals that, when all occupations are considered together, the most significant driver behind the gains from attempting self-employment is a significant increase in annual hours worked. This is shown in the bottom-most graph. Prior to undertaking self-employment, the ASE do not work significantly more hours than their matches. Immediately afterward, however, the ASE work almost 300 hours more per year, an average of 9% more than their PE counterparts, which is a statistically significant difference.[38] This positive and significant difference persists at more or less the same level for the entire 14 years after the self-employment attempt. The wage, shown in the top left graph of Figure 1, is not significantly affected by the self-employment attempt.[39] Including asset income in the calculation of the wage, as shown in the "wage plus" graph at the top right of Figure 1, shifts the effect of trying self-employment in a positive direction and introduces a slightly positive trend. Over this period, the ASE make about $2.50 more per hour, or about 15% more than those who never attempted self-employment. However, the increase is significant in only 5 of the 15 periods, and it is interrupted by a large drop in period 15.

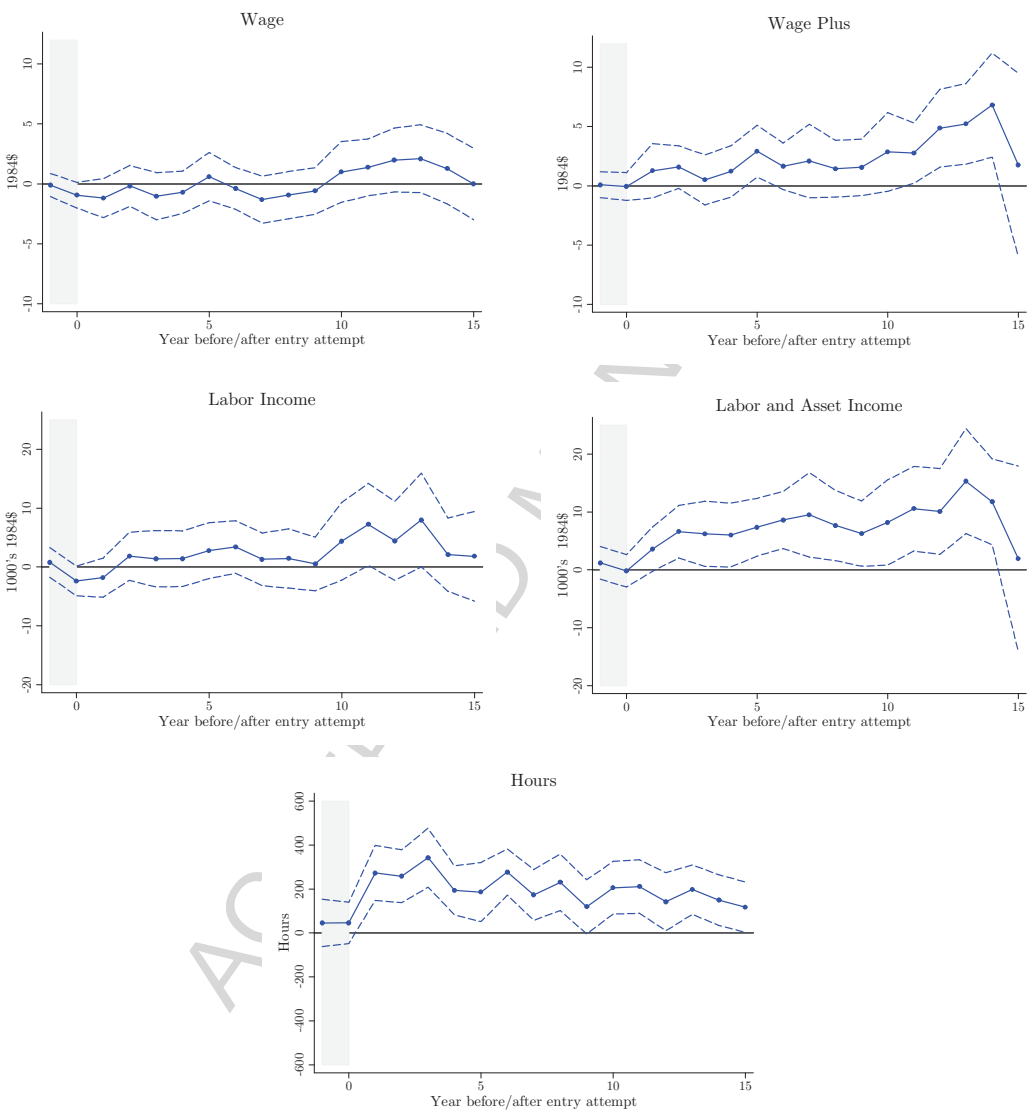
The net effect on annual labor income and on annual labor and asset income is shown in the middle two graphs of Figure 1. As expected, the pattern of average

[38] The percentages used in this section are calculated as the ATT at time $t$ divided by the average value at time $t$ for the PE group. If an average percentage is described, for instance over 15 years, it is the average of the per-period percentages over those 15 years.

[39] Note that the ATTs for the wage are estimates of $\Delta(Income/hours)$ due to the self-employment attempt, not $\Delta Income/\Delta hours$

treatment effects on the treated over the 15 year period mirrors that presented in the corresponding wage graphs, although they display larger gains due to the large increase in annual hours worked. The ATTs calculated from the labor income are for the most part positive, but are never significant. On the other hand, the ATTs calculated from labor and asset income are always positive and are for the most part statistically significant. These ATTs average about $8000 per year over the period, which is a gain of about 20% over those who did not attempt self-employment. Note that aside from the negative returns in the first period, when labor income is considered, the non-negative pattern of returns suggests that even impatient individuals may opt for self-employment. The ATTs of the annual measures add up approximately to the ATTs calculated from the 5-, 10-, and 15-year present values of income, as presented in columns 1 and 4 of Table 3. The differences can be attributed to discounting and to the bias correction used in the matching procedure.

When individuals of all occupations are considered simultaneously, the gain to labor income from attempting self-employment comes predominantly from the increase in hours worked. However, when the labor and asset income measure is considered, higher wages also become a contributing factor.




Average treatment effects over time, the solid line, shown with 90% confidence intervals, the dashed lines.

**Figure 1:** Average treatment effects on the treated, all occupations

Figure A2 in the appendix presents the results for those who attempt self-employment and who are in technical or professional occupations. It reveals that this group realizes substantial and increasing gains to their wage relative to their PE counterparts. The average ATT during the first 5 years is almost \$4, 20% more than that of individuals who did not try self-employment. During years 5 through 10, the premium increases to almost \$7, a wage 35% higher than the ASE would have earned had they remained paid employees. In the final five years of observation, the ASE make, on average, slightly more than \$8.50, which is 40% more than their PE counterparts, although the effect is significant in only three of the periods. When the wage-plus measure is considered, the rewards are more pronounced: the majority of the positive average treatment effects on the treated are significant, as is shown in the top right graph of Figure A2. These effects are also increasing over time, growing from an average of approximately \$6.50 in the first five years to more than \$11 in the following 10 years. These are relative gains of almost 40% and more than 50%, respectively. The ASE engaged in a technical or professional occupations are working about 10% more hours, or about 200 hours more per year, with no clear trend over the period. Due to the underlying volatility in hours, the ATT on hours is only significant in three of the periods.

As expected, the total effect of attempted self-employment that results from higher wages and working hours is positive and produces generally increasing ATTs on annual income measures. As with wages, these rewards are more pronounced with the labor and asset income measure than with the labor income measure. The effects in 14 of the 15 years are significant with the former, whereas only 6 are significant with the latter. On average, the technical and professional ASE earn 40% to 60% more per year than their PE counterparts, depending upon whether labor income or labor and asset income is considered. This agrees with the average treatment effects on the treated calculated from the present discounted values presented in Table 3. Overall, those whose occupations are technical and professional see large rewards for attempting self-employment stemming from higher wages and more work hours. Again, the insignificant effects shown in periods prior to the self-employment attempt reveal that selection is not at play.

Craftsmen do not fare as well, as can be seen in Figure A3. The ATTs on wage average –\$0.35 over the 15-year period, display no clear trends and rarely differ significantly from 0. Incorporating the hourly asset income shifts the wage series in a positive direction. The treatment effects calculated from the wage-plus measure are predominantly positive and average \$2 over the time period, which is about 15% higher than what the ASE would have made had they never attempted self-employment. These are rarely significant. Craftsmen who have attempted self-

employment tend to work about 10% less, or 250 hours less, in the five years following the attempt, although they come to work more hours as time passes. By the sixth year, the ASE tend to work more hours than their PE counterparts, and they work on average about 7% more hours over the rest of the observation period. The combination of similar wages and fewer hours worked during the period immediately following the ASE attempt leads to negative, but rarely significant, effects on labor income. In later years, the effect on labor income is even smaller in magnitude, alternating in sign and never significant. When asset income is included, the ATTs are generally positive: on average the ASE make about 13% more per year than their PE counterparts throughout the time period. The magnitude of the reward is substantial in the last five years of the observation period and is significant in three of those five years, averaging about seven thousand dollars, a 25% gain relative to the PE.

*4.1.3. Quantile treatment effects on the treated*

To investigate the distribution of the treatment effects, the quantile treatment effects on the treated, QTT, are estimated following Firpo (2007).[40] The uppermost two graphs of Figure 2 present the quantile treatment effects on the treated as calculated from the present discounted value of total labor income and of total labor and asset income at the 5-, 10-, and 15-year horizons for all occupations. Regardless of the income measure used, it is clear that the rewards from attempting self-employment are being driven by the upper tail of the distribution. Individuals in the 75th and 90th percentiles see increasing rewards from attempting self-employment. When the labor plus asset income measure is used, these rewards are quite large and statistically significant. The median treatment effects on the treated are both negative and significant if labor income is considered, but are effectively 0 if instead the labor plus asset income is used. The QTTs at the 10th and 25th percentiles are also significantly negative and are generally decreasing over time. The variance of the distribution of treatments effects is increasing with respect to the time horizon.

The middle two graphs of Figure 2 display the quantile treatment effects on the treated as calculated from the present value of labor income and from labor and asset income for the technical and professional ASE. The shapes of the distributions are similar to those shown for all occupations but are more positively skewed. The QTTs at the 90th percentile are very large in magnitude, reaching approximately

[40] The Stata command ivqte was used to generate the propensity score suggested by Firpo (2007). See Frölich and Melly (2010) for a nice overview. These scores were then used to calculate the weights defined in equation (3) of Firpo.

1 and 1.3 million dollars respectively at the 15-year horizon for the labor income measure and the labor and asset income measure. The quantile treatment effect on the treated at the 75th percentile is also large, positive, and significant. The median treatment effects on the treated of total labor income are insignificant and relatively close to zero. However the median treatment effects on the treated of total labor and asset income are significant and positive, about $120,000 after 10 years. The 10th and 25th percentile QTTs are negative and are relatively constant in comparison to those of the 75th and 90th percentiles. In absolute value, they are much smaller in magnitude than the QTTs found at the upper tail of the distribution, which is to be expected given the ATTs presented in Table 3.

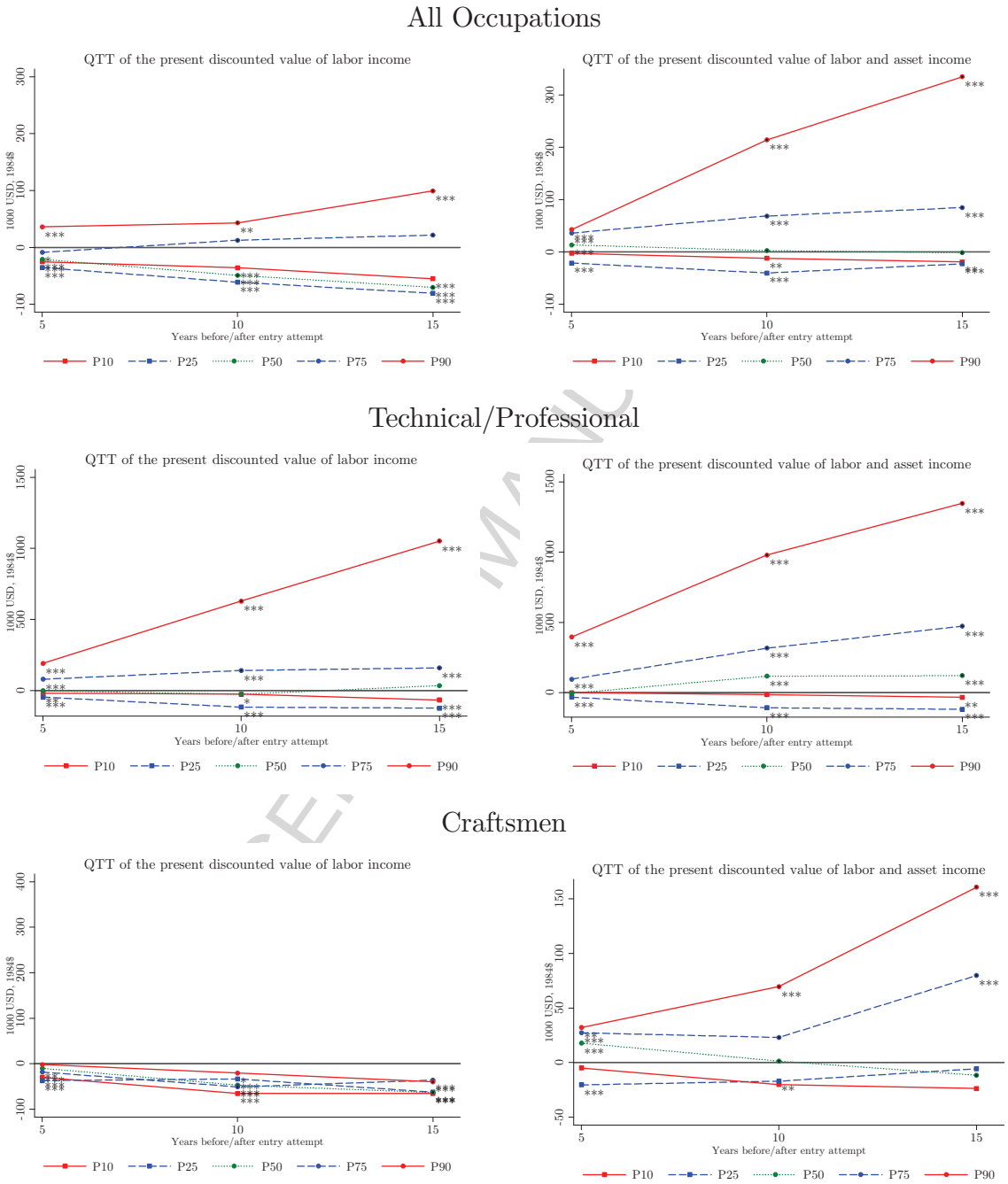
The bottom two graphs of Figure 2 display the QTTs for the ASE who were craftsmen prior to their attempt. The negative average treatment effects shown in column 3 of Table 3 are the result of a distribution of treatment effects that, at least between the 10th and the 90th percentiles, lies entirely below 0. However, if the labor and asset income measure is used, the picture is less bleak. Those ASE at the 90th and 75th percentiles earn significant rewards, approximately $160,000 and $80,000 respectively, relative to their PE partners.

The quantile treatment effects on the treated calculated from wages and hours are presented in the appendix, where it is shown that average gains (if any) in wages are driven by the upper tail of the distribution and that non-negative hour effects are generally present throughout the distribution for all samples.

## 5. Robustness Checks

### *5.1. IV estimation of LATE*

So far, identification has relied on the assumption of selection on observables. To bolster this premise, the local average treatment effect (LATE) of attempting self-employment is estimated. The estimation is performed separately for two instruments: the self-employment status of the father, and whether the individual received a lump sum payment in the year prior to the attempt. Although the father's self-employment status has been used as an IV, its exogeneity is questionable. Specifically, if, as Dunn and Holtz-Eakin (2000) suggest, parents actually pass entrepreneurial skills on to their sons, rather than or in addition to a simple taste for self-employment, then future income ought to be higher, violating exogeneity. Blanchflower and Oswald (1998), using British data, find a positive and significant correlation between receiving an inheritance or a monetary gift and the probability of becoming self-employed. This suggests that such a measure is a viable instrument. As was mentioned, the PSID contains a variable for whether or not an individual has




*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

**Figure 2:** Quantile treatment effects on the treated, present discounted value of future income

received a lump sum payment. Although neither IV candidate is without problems, the resulting estimates serve as a reality check on the validity of the assumption of selection on observables.

The LATE is estimated rather than the ATE because neither of the proposed instruments justifies the assumption that all individuals could be induced to attempt self-employment simply through a change in the instrument, and no sufficiently strong instrument is available. This exercise can be reasonably performed only on the sample containing all occupations; the sample sizes in the occupation-specific sub-samples are too small to rely on the asymptotic properties of the estimator. The nonparametric estimator proposed by Frölich (2007) is used. The covariates included in the estimation are the same as those used in the matching estimation.[41] The resulting estimates are presented in Table 4.

For ease of comparison, columns 1 and 4 of Table 4 present the main results, the estimated ATT of labor income and labor and asset income, respectively, that were obtained using Abadie and Imbens's (2011) matching estimator with bias correction and presented in Table 3. Columns 2 and 5 present the estimated LATEs for labor income and for labor and asset income, respectively, when the lump sum payment is used as an instrument for attempting self-employment. Analogously, columns 3 and 6 present the LATE estimates when the self-employment status of father is used as an instrument. The estimated LATEs on labor income obtained using lump sum payments as an instrument, shown in column 2, are larger than the ATTs obtained via matching, shown in column 1, but are insignificant in both cases. Columns 4 and 5 reveal that both estimation techniques yield remarkably close and highly significant values for the effects on labor and asset income, despite being estimates of different parameters (the LATE and the ATT respectively). The LATEs for labor income attained when the father's self-employment status is used as an instrument, shown in column 3, are substantially larger than the ATTs attained via matching and are significant. The LATEs for labor and asset income are still large and significant, as shown in column 6, but are larger than the ATTs attained via matching, in column 4, only for the 15-year time horizon. The fact that all estimates were of the same sign (and in the case of the lump sum IV, the same magnitude) regardless of the estimation strategy lends credibility to the assumption of selection on observables.

*5.2. Finite Sample Properties*

Given the small sample sizes in this paper, the finite sample properties of the matching estimator used are important. Huber, Lechner, and Wunsch (2013) survey

[41] The nplate command in stata was used with a bandwidth of infinity and lamda=1

**Table 4:** IV estimation of LATE

| | Labor income | | | Labor and asset income | | |
|---|---|---|---|---|---|---|
| | *Abadie Imbens* | *Lump Sum* | *Father SE* | *Abadie Imbens* | *Lump Sum* | *Father SE* |
| *5 years* | 5·2 | 9·2 | 25·7*** | 28·5** | 32·2* | 19·7*** |
| | (11·7) | (13·9) | (3·6) | (13·0) | (16·5) | (3·6) |
| *10 years* | 14·7 | 21·7 | 43·9*** | 63·4** | 64·0* | 30·2*** |
| | (22·5) | (27·7) | (7·5) | (25·9) | (32·8) | (6·7) |
| *15 years* | 33·4 | 44·4 | 167·5*** | 102·6** | 107·5** | 190·2*** |
| | (36·2) | (43·8) | (19·6) | (41·5) | (50·8) | (19·6) |
| Proportion of compliers | −− | 0·149 | 0·449 | −− | 0·149 | 0·449 |
| Observations | 4,204 | 4,204 | 4,204 | 4,204 | 4,204 | 4,204 |

*Notes:* The stata command nplate was used. All monetary figures presented in thousands of 1984 dollars. The columns labeled "Lump Sum" present the IV estimates when an indicator for lump sum payment is used as an instrument. The columns labeled "Father SE" present the IV estimates when the self-employment status of the father is used as the instrument. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

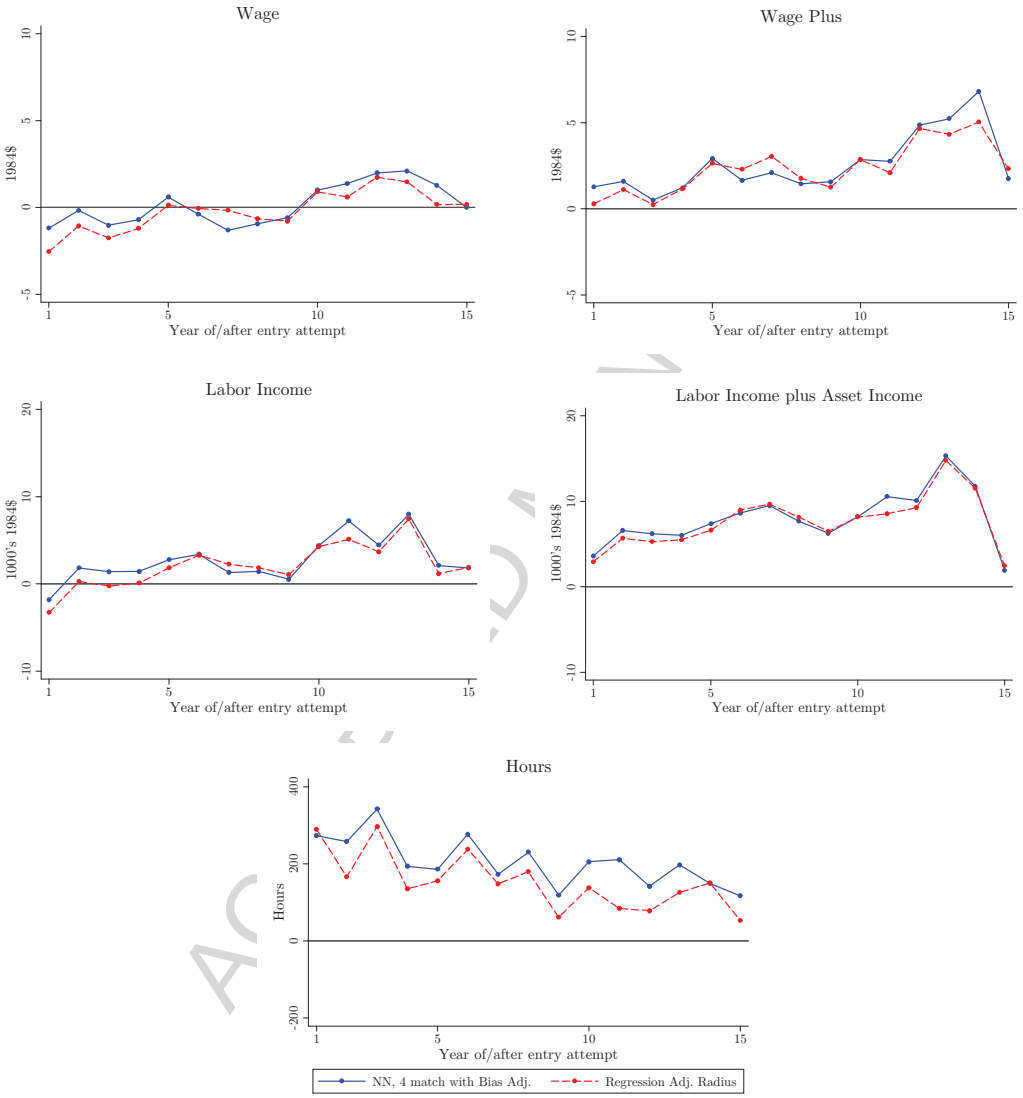
the finite sample properties of a large number of estimators based on the propensity score. Although they do not explicitly cover Abadie and Imbens's nearest neighbor matching estimator with bias correction, they do investigate several similar matching estimators in the same class (direct matching as opposed to inverse probability weighting, kernel matching, or parametric methods). Both within its class and across classes, the radius matching with regression adjustment proposed by Lechner, Miguel, and Wunsch (2011) dominates on the basis of the RMSE when its performance is evaluated along several dimensions: sample size (300, 1200, and 4800), strengths of selection, specifications of the propensity score, and proportion of treated observations. The Lechner, Miguel, and Wunsch estimator will be used to verify that the resulting estimates are similar to those of Abadie and Imbens's estimator. This comparison serves to verify that the Abadie and Imbens estimator is reasonable with the sample sizes used and also provides evidence that the main results presented earlier are not a function of the particular matching method used.

Table 5 presents the ATTs calculated with the Abadie and Imbens method, in columns 1 and 3, and the ATTs calculated with the Lechner, Miguel, and Wunsch estimator, in columns 2 and 4. These two sets of estimates lead to the same conclusions, although the Abadie and Imbens estimates are slightly higher. Figure 3 plots the ATTs for wages, annual income, and hours both for the nearest-neighbor matching estimator with bias correction and for radius matching with bias correction. These estimators produce similar results, no doubt because they are from the same class. The similarity of the Huber, Lechner, and Wunsch estimates to the Abadie and Imbens estimates suggests that the latter too possess reasonable finite sample properties and that the main results are not a function of the particular matching method used.

**Table 5:** Regression Adjusted Radius Matching

| | | Labor Income | | Labor and Asset Income | |
|---|---|---|---|---|---|
| | | *AI* | *LMW* | *AI* | *LMW* |
| | | (1) | (2) | (4) | (5) |
| *5 years* | Difference | 5·2 | −1·5 | 28·5∗ | 24·8∗ |
| | Standard Error | (11·7) | (11·5) | (13·0) | (13·0) |
| | % Gain or Loss | 154·0 | −1·0 | 164·0 | 15·6 |
| *10 years* | Difference | 14·7 | 9·5 | 63·4∗∗ | 60·9∗∗ |
| | Standard Error | (22·5) | (24·1) | (25·9) | (27·7) |
| | % Gain or Loss | 302·9 | 3·1 | 320·6 | 19·2 |
| *15 years* | Difference | 33·4 | 24·7 | 102·6∗∗ | 97·6∗∗ |
| | Standard Error | (36·2) | (41·8) | (41·5) | (47·3) |
| | % Gain or Loss | 441·7 | 5·5 | 468·1 | 20·8 |
| | Treated | 127 | 127 | 127 | 127 |
| | Total | 4,204 | 4,204 | 4,204 | 4,204 |

*Notes:* The stata command radiusmatch was used to calculate the Lechner, Miquel and Wunsch estimates, shown in the column labeled "LMW". The Abadie and Imbens estimates are shown in the columns labeled "AI". All monetary figures presented in thousands of 1984 dollars. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.




The solid line plots the estimates resulting from Abadie and Imbens' nearest neighbor matching with bias adjustment. The dashed line plots the Lechner, Miguel and Wunsch radius matching with regression adjustment estimates.

**Figure 3:** Comparison of Matching for All Occupations

*5.3. Sensitivity Analysis*

To understand how sensitive the results presented are to the omission of a confounding variable, the method proposed by Ichino, Mealli, and Nannicini (2008) is implemented. The procedure entails $r$ simulations of a binary variable whose omission would lead to (in this case) positive bias in the ATT. The ATT including this simulated variable is computed at each iteration. They give the following example: assume that a binary skill variable, $U$, is omitted from the matching process. If this skill leads to both a higher probability of attempting self-employment and a higher income among those who don't attempt self-employment, its omission from the matching process would lead to a positive bias. The properties of such an omitted variable can be pinned down with four probabilities:

$$\begin{aligned} \Pr(U &= 1|T=i, Y=j) \equiv p_{ij} \\ i &= 0,1 \; j = 0,1 \end{aligned}$$

where $T$ is an indicator of whether the individual has attempted self-employment and $Y$ is an indicator that takes the value 1 if the outcome is above its mean. A positive bias will then occur if (1) there is a positive effect into treatment, $p_{1.} - p_{o.} > 0$, and (2) there is a positive effect on the untreated outcome, $p_{01} - p_{00.}$. In each iteration, two logit models are estimated. The first logit estimates $P(Y = 1|T = 0, U, W)$, from which the effect of U on the relative probability of a positive outcome in the case of no treatment, the "outcome effect," is calculated as the estimated odds ratio of the variable U. The second logit estimates $P(T = 1|U, W)$, from which the effect of $U$ on the relative probability of being treated, the "selection effect," is also calculated as the odds ratio of U. Under conditions (1) and (2) above, the output and selection effects must be greater than 1. Following Ichino, Mealli, and Nannicini (2008), we calibrate the omitted variable to mimic the properties of several observed variables that meet that criteria described above, and we then calculate the resulting ATTs. "Killer" confounders, the outcome and selection effects of which are chosen to push the ATT toward 0, are also created and are used to understand how plausible such confounders are. For the details of this method, see Ichino, Mealli, and Nannicini (2008).

The value of $U$ was simulated 500 times for each calibration. It was not feasible to perform this exercise without assuming homogenous effects in the implementation of the Abadie and Imbens matching estimator.[42] The simulation exercise was performed using the 10-year present discounted value of the labor plus asset income outcome

[42] This assumption will not change the point estimates but will affect the standard errors.

variable because its ATT is significant in Table 3, although the conclusions do not vary if other outcome variables are considered instead.[43]

Table 6 presents the results. Row 1 and column 7 of the table show the ATT calculated from the 10-year present discounted value of the labor plus asset income, as reported in Table 3. The standard error under the assumption of homoskedasticity is given in column 8. The second row shows the impact on the ATT of omitting a neutral confounder—one whose outcome and selection effects are each 1. If we compare row 1 to row 2, we see that adding a neutral confounder has very little effect. The second panel of the table, labeled "confounder-like", displays the effects of omitting variables that have the same properties as important matching variables. The effects on the ATTs are small even for a confounder like incorporated business ownership, which has substantial outcome and selection effects. The third panel of the table, "killer confounders" demonstrates that even with very large outcome and selection effects, of approximately 400 and 60 respectively, the ATT retains its significance. The ATTs' insensitivity to the omission of potential confounders may in part be due to the large number of observed variables used in the matching process in conjunction with the local linear regression for bias correction used by the Abadie and Imbens method.

*5.4. Magnitude of wife asset income in husband's labor and asset income measure*

It may be of concern that the annual total labor and asset income measure includes not only the head's but also the wife's asset income from farming, unincorporated businesses, and roomers and boarders. Unfortunately, the PSID did not collect these data separately until 1985, and the main sample used here considers only individuals who attempted self-employment for the first time between 1971 and 1981. Nonetheless, some sense of the magnitude of the wives' asset incomes can be gleaned from the available data. The average treatment effects on the treated of wife asset income were calculated for the period in which self-employment is attempted by their husbands using the sample of husbands whose first attempt occurred in the period 1985–1991 and their matches. The results are presented in the online appendix. The ATT estimates are small in magnitude and insignificant; furthermore, the quantile treatment effects at the 10th, 25th, 50th, and 75th percentiles are 0 regardless of occupational categorization. Only at the 90th percentile, and only for the entire sample

[43] The Stata program written by Tommaso Nannicini, sensatt, was modified so that Abadie and Imbens's nearest-neigbor matching with adjustment could be used to calculate the ATT, which is not an option in sensatt. The confounding variabe was also included in the logit that determines the trimmed sample. The probabilities used for the killer confounders were calculated using Nannicini's matlab code, which is available on his website.

**Table 6:** Effects of omitting a confounding variable

| | $p_{11}$ | $p_{10}$ | $p_{01}$ | $p_{00}$ | Outcome Effect | Selection Effect | ATT | SE |
|---|---|---|---|---|---|---|---|---|
| No confounder | | | | | | | 63·4 | (20·4) |
| Neutral confounder | 0.50 | 0·50 | 0·50 | 0·50 | 1·0 | 1·0 | 63·6 | (20·5) |
| Confounder-like: | | | | | | | | |
| Father self-employed | 0.09 | 0·09 | 0·04 | 0·04 | 1·2 | 2·6 | 62·8 | (20·5) |
| Lump sum | 0.14 | 0·09 | 0·08 | 0·06 | 1·3 | 1·9 | 62·6 | (20·5) |
| Worked both as SE and PE | 0.17 | 0·19 | 0·01 | 0·01 | 0·8 | 23·0 | 62·8 | (21·4) |
| College degree | 0.42 | 0·22 | 0·42 | 0·11 | 6·2 | 1·5 | 62·9 | (20·4) |
| Business ownership | 0.45 | 0·45 | 0·06 | 0·02 | 3·0 | 28·3 | 61·0 | (21·9) |
| Incorporated business | 0.24 | 0·16 | 0·02 | 0·01 | 3·6 | 27·0 | 60·9 | (21·4) |
| Killer Confounders: | | | | | | | | |
| *s=.1 d=.1* | 0.40 | 0·40 | 0·39 | 0·29 | 1·6 | 1·3 | 63·3 | (20·5) |
| *s=.2 d=.2* | 0.49 | 0·49 | 0·49 | 0·29 | 2·4 | 1·7 | 62·0 | (20·4) |
| *s=.3 d=.3* | 0.59 | 0·59 | 0·58 | 0·28 | 3·6 | 2·2 | 61·2 | (20·4) |
| *s=.4 d=.4* | 0.69 | 0·69 | 0·68 | 0·28 | 5·6 | 3·0 | 60·9 | (20·4) |
| *s=.5 d=.5* | 0.78 | 0·78 | 0·77 | 0·27 | 9·6 | 4·6 | 60·4 | (20·3) |
| *s=.6 d=.6* | 0.88 | 0·88 | 0·86 | 0·26 | 20·1 | 9·3 | 59·6 | (20·2) |
| *s=.7 d=.7* | 0.98 | 0·98 | 0·96 | 0·26 | 433·1 | 59·9 | 55·5 | (21·0) |

*Notes:* 4 matches used in matching, homogenous treatment effects assumed. Note that in order to pin down the confounders, p11=p10 is assumed to hold as in Ichino, Mealli, and Nannicini (2008). All monetary figures presented in thousands of 1984 dollars. "Worked both as SE and PE" is an indicator variable that takes the value 1 if the individual worked both as an employee and as self-employed prior to becoming solely self-employed. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

and the technical and professional sub-sample, do the effects differ significantly from 0: they are –$300 and –$1200 respectively. Assuming that the same patterns hold throughout the time frame considered in the main sample, the inclusion of the wife's asset income in the labor and asset income of the husband should either not affect the rewards flowing to the husband or should lead to a downward bias.

## 6. Conclusion

On average, those who attempt self-employment are not punished for doing so. In fact, the gains in the present discounted value of total labor and asset income for the ASE are 22% higher than those of the PE and are significant. Estimates of the quantile treatment effects on the treated reveal that these positive returns are being driven by the upper tail of the distribution. The driver behind the gains from attempting self-employment is the significant increase in annual hours worked. Wages are either not significantly affected by the attempt at self-employment or are positively affected, depending upon the measure used. Taken together, these results suggest that individuals who attempt self-employment may do so for pecuniary as well as non-pecuniary benefits. Attempting self-employment, on average, does not diminish subsequent wages; rather, it opens up an avenue by which individuals can realize higher incomes, namely by working more.

The returns from attempting self-employment are heterogeneous with respect to occupation. Individuals in professional or technical occupations prior to attempting self-employment enjoy substantial and significant returns. After 15 years, they realize significant gains, of 45% and 62% respectively in the present discounted value of labor income and of labor and asset income. The average increase in the present discounted value of income from attempting self-employment is a product of both significantly higher wages and longer work hours. Craftsmen, on the other hand, see losses, albeit insignificant ones, when labor income is considered. When asset as well as labor income is included, craftsmen are rewarded for attempting self-employment, realizing a gain of about 13% after 15 years, though again the effect is insignificant.

The positive effects of attempting self-employment are driven by the upper tail of the distribution when income and wages are considered; whereas the non-negative effects are present throughout the entire treatment effect distribution when the number of hours worked is considered.

These results suggest that there is no significant evidence that those who attempt self-employment are punished for doing so. On the contrary, the rewards in labor and asset income are positive regardless of the individual's occupation, significantly so when all occupations are pooled together and when the technical or professional

occupation is considered in isolation. However, only individuals in technical and professional occupations see significantly higher labor income. These results suggest that the desire to try self-employment could be driven by the anticipation of pecuniary rewards as well as of non-pecuniary benefits, and could justify policies whose purpose is to encourage self-employment as an escape from poverty.

To our knowledge, this is the first study to estimate the longer-term consequences of attempting self-employment on the present value of future income and to do so in a nonparametric way with such a comprehensive list of control variables.

A number of robustness checks are performed in order to verify the assumption of unconfoundedness. The findings are in one sense a lower bound: Hurst, Li, and Pugsley (2010) used PSID data from the Consumer Expenditure Survey to show that the self-employed underreport their income by about 30%. On the other hand, the magnitude of the premium for trying self-employment may be overestimated: the PSID data used here does not include other non-wage aspects of job compensation, such as health insurance.

Only the individual effects of attempting self-employment have been explored. It is quite plausible that a husband's self-employment attempt could also have effects on his wife. The online appendix presents evidence that the self-employment rate of ASE wives jumps after their husbands' self-employment attempts, in accord with Bruce (1999). No significant differences were detected in income, although the power is low due to the small sample sizes. The income dynamics of the wives of the self-employed would be very interesting to study in the future.

## Acknowledgements

I would like to thank Kevin Lang for the encouragement and direction he gave me on an earlier version of this paper. I would also like to thank Randy Ellis, Michael Manove, and Hsueh-Ling Huynh, seminar participants at Boston University and Copenhagen Business School. Finally, I thank the editor and an anonymous referee, both of whose suggestions led to substantial improvements to this paper.

## References


[1] Abadie, A, Drukker D,Leber Herr, J, Imbens, G.W. Implementing matching estimators for average treatment effects in Stata. Stata journal 4 (2004): 290-311.

[2] Abadie, A, Imbens, Guido W. Bias-corrected matching estimators for average treatment effects. Journal of Business & Economic Statistics 29, no. 1 (2011): 1-11.

[3] Abadie, Alberto, Imbens, Guido W. Matching on the Estimated Propensity Score, NBER working paper 15301, December 2012

[4] Andersen, Steffen & Nielsen,Kasper Meisner. Ability or Finances as Constraints on Entrepreneurship? Evidence from Survival Rates in a Natural Experiment. Review of Financial Studies, Society for Financial Studies, vol. 25(12), pages 3684-3710 , 2012

[5] Baumol, W. Toward Prosperity and Growth. extract from Kauffman Thought-book 2011

[6] Becker, S., Ichino, A Estimation of Average Treatment Effects Based on Propensity Scores. The Stata Journal 2002; 2(4): 358–377.

[7] Berglann, Helge, Moen, Espen R., Røed, Knut and Skogstrøm, Hens Frederik Entrepreneurship: Origins and Returns. Labour Economics, 18, (2011) 180-193

[8] Bracke, Philippe, Hilber, Christian, Silva, Olmo Homeownership and Entrepreneurship: The Role of Commitment and Mortgage Debt. IZA Discussion Papers 7417, Institute for the Study of Labor (IZA) May 2013.

[9] Blanchflower, David G & Meyer, Bruce D A Longitudinal Analysis of the Young Self-Employed in Australia and the United States. Small Business Economics, Springer, vol. 6(1), pages 1-19, February.1994

[10] Blanchflower, David G & Oswald, Andrew J What Makes an Entrepreneur? Journal of Labor Economics, University of Chicago Press, vol. 16(1), pages 26-60, January.1998

[11] Bruce, Donald Do Husbands Matter? Married Women Entering Self-Employment. Small Business Economics, Vol. 13, No. 4 (Dec., 1999/2000), pp. 317-329

[12] Bruce, Donald and Schuetze, Herbert J. The Labor Market Consequences of Experience in Self-Employment. Labour Economics 2004;11: 575-598.

[13] Caliendo M , Kopeinig S Some practical guidance for the implementation of propensity score matching. Journal of economic surveys 22 (1), 31-72, 2008

[14] Carroll, Glenn R. & Mosakowski, Elaine The Career Dynamics of Self-Employment. Administrative Science Quarterly, Vol. 32, No. 4 (Dec., 1987), pp. 570-589

[15] Clark, Kenneth & Drinkwater, Stephen Pushed out or pulled in? Self-employment among ethnic minorities in England and Wales. Labour Economics, Volume 7, Issue 5, September 2000, Pages 603-628,

[16] Davis, Steven and Willen, Paul. Occupation-level Income Shocks and Asset Returns: Their Covariance and Implications for Portfolio Choice. NBER Working Paper, No. 7905.

[17] De Wit, Gerrit ; Van Winden,Frans A.A.M. An empirical analysis of self-employment in the Netherlands. Economics Letters, Volume 32, Issue 1, January 1990, Pages 97-100

[18] Dunn, Thomas and Holtz-Eakin, Douglas Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links. Journal of Labor Economics April 2000; 18(2): 282-305.

[19] Evans, David S & Jovanovic, Boyan An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. Journal of Political Economy, University of Chicago Press, vol. 97(4), pages 808-27, August.1989.

[20] Evans, David S & Leighton, Linda S. Some Empirical Aspects of Entrepreneurship. American Economic Review, American Economic Association, vol. 79(3), pages 519-35, June.1989

[21] Faggio, Giulia & Silva,Olmo Does Self-Employment Measure Entrepreneurship? Evidence from Great Britain. SERC Discussion Papers 0109, Spatial Economics Research Centre, LSE. 2012.

[22] Fairlie, Robert W. & Meyer,Bruce D. The Ethnic and Racial Character of Self-Employment. NBER Working Papers *4791*, National Bureau of Economic Research, Inc.1994

[23] Firpo, Sergio. Efficient Semiparametric Estimation of Quantile Treatment Effects. Econometrica 2007; 75(1): 259-276.

[24] Folta, Timothy B., Delmar, Frédéric and Wennberg, Karl Hybrid entrepreneurship. Management Science 56, no. 2 (2010): 253-269.

[25] Frölich, Markus Nonparametric IV estimation of local average treatment effects with covariates. Journal of Econometrics 139 (2007) 35-37

[26] Frölich, Markus and Melly, Blaise Estimation of quantile treatement effects with Stata. The Stata Journal (2010), 10, Number 3, pp. 423-457

[27] Guvenen, F. An empirical investigation of labor income processes. Review of Economic Dynamics, 12 (1), 58–79, 2009

[28] Hamilton, Barton. Does Entrepreneurship Pay? An empirical Analysis of the Returns of Self-Employment. Journal of Political Economy 2000; 108(3): 604-631.

[29] Heaton, John & Lucas,Deborah. Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk. Journal of Finance, American Finance Association, vol. 55(3), pages 1163-1198, 06. 2000.

[30] Hill, Martha S. The Panel Survey of Income Dynamics, A User's Guide. Sage Publications, 1992

[31] Hipple, Steven F. Self-employment in the United States. Monthly Labor Review 133.9 (2010): 17-32.

[32] Hryshko,Dmytro Labor income profiles are not heterogeneous: Evidence from income growth rates. Quantitative Economics, Volume 3, Issue 2, Article first published online: 25 JUL 2012

[33] Huber,Martin , Lechner,Michael & Wunsch,Conny The performance of estimators based on the propensity score. Journal of Econometrics, Volume 175, Issue 1, July 2013, Pages 1-21

[34] Hurst, Erik & Li,Geng & Pugsley, Benjamin Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed NBER Working Papers 16527, National Bureau of Economic Research, Inc. 2010.

[35] Ichino, Andrea, Mealli, Fabrizia, and Nannicini, Tommaso From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity? Journal of Applied Econometrics, 2008, 23(3), 305-327

[36] Imbens, GW. The role of the propensity score in estimating dose-response functions. Biometrika 2000; 87(3): 706-710.

[37] Imbens, GW. Nonparametric estimation of average treatment effects under exogeneity: A review. Review of Economics and Statistics 2004; 86(1): 4-29

[38] Imbens, GW. Implementing Matching Estimators for Average Treatement Effects in Stata. West Coast Stata Users Group meeting, Los Angeles October, 2007

[39] Jovanovic, Boyan Selection and the Evolution of Industry. Econometrica, Econometric Society, vol. 50(3), pages 649-70, May.1982

[40] Kahn, S. and Lang, K. Constraints on the Choice of Work Hours: Agency vs. Specific Capital. Journal of Human Resources, 27 (Fall 1992), 661-678.

[41] Koenker, Roger and Hallock, Kevin. Quantile Regression. Journal of Economic Perspectives 2001; 14(4):143-156.

[42] Lazear, Edward P. Entrepreneurship. Journal of Labor Economics, University of Chicago Press, vol. 23(4), pages 649-680, October 2005

[43] Le, Anh T. Empirical Studies of Self-Employment. Journal of Economic Surveys, Volume 13, Issue 4, pages 381–416, September 1999

[44] Lechner, Michael , Miquel, Ruth and Wunsch, Conny Long-run effects of public sector sponsored training in West Germany. Journal of the European Economic Association, Volume 9, Issue 4, pages 742–784, August 2011

[45] Levine, Ross and Rubinstein,Yona Smart and Illicit: Who Becomes an Entrepreneur and Does it Pay? Working Paper 19276, National Bureau of Economic Research August 2013.

[46] Malmendier, Ulrike and Tate, Geoffrey Superstar CEOs. The Quarterly Journal of Economics (2009) 124 (4): 1593-1638

[47] Martinez-Granado, Maite Testing labour supply and hours constraints. Labour Economics 12.3 (2005): 321-343.

[48] Moskowitz, Tobias J., and Annette Vissing-Jorgensen. The returns to entrepreneurial investment: A private equity premium puzzle?. No. w8876. National Bureau of Economic Research, 2002.

[49] Nanda, Ramana. Cost of External Finance and Selecion into Entrepreneurship, HBS Working Paper 2008; Number 08-047.

[50] Parker, Simon C., Belghitar, Yacine and Barmby, Tim. Wage Uncertainty and the Labour Supply of Self-Employed Workers. The Economic Journal 2005; 115

[51] Phillips, Bruce D., and Kirchhoff, Bruce A. Formation, Growth, and Survival: Small Firm Dynamics in the U.S. Economy. Small Business Economics 1989; 1: 65-74.

[52] Quadrini, Vincenzo Entrepreneurship, Saving and Social Mobility. Review of Economic Dynamics Elsevier for the Society for Economic Dynamics, vol. 3(1), pages 1-40, January 2000.

[53] Rissman, Ellen. Self-Employment as an Alternative to Unemployment, FRB of Chicago Working Paper 2003; No. 2003-34. Available at SSRN: http://ssrn.com/abstract=486988

[54] Rosen, Harvey and Willen, Paul Risk Return and Self-Employment. Mimeo 2002; Princeton and Chicago University.

[55] Rosen, Sherwin. The Economics of Superstars. American Economic Review December 1981; 71: 845-58.

[56] Rosenbaum, Paul R., and Donald B. Rubin. "The central role of the propensity score in observational studies for causal effects." Biometrika 70.1 (1983): 41-55.

[57] Rubin, Donald B. Using propensity scores to help design observational studies: application to the tobacco litigation. Health Services and Outcomes Research Methodology 2.3-4 (2001): 169-188.

[58] Stuart, Elizabeth A. Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics 25.1 (2010): 1.

[59] Williams, Donald R. Consequences of self-employment for women and men in the United States. *Labour Economics*, Elsevier, vol. 7(5), pages 665-687, September 2000

[60] Wooldridge, Jeffrey M. Econometric Analysis of Cross Section and Panel Data. MIT Press, Second Edition, see Chapter 21, page 929

## Highlights

Another look at the choice to try self-employment: a matching approach

- On average those who attempt self-employment (ASE) are not punished for doing so
- After 15 years, the ASE realize 8-20% higher present discounted values of income
- This premium is driven by the ASE's higher working hours and similar wages
- The ASE who are technical/professional earn large wage premiums and work more hours
- Positive effects in income and wages are driven by the upper tail