

Erasmus Universiteit Rotterdam

Erasmus School of Economics

Bachelorscriptie Algemene Economie

The effect of children on the self-employment probability.

Naam: Rutger van Noort

Studentnummer: 415785

Begeleider: Max van Lent

Tweede beoordelaar: Josse Delfgaauw

Datum definitieve versie: 23-8-2017

Abstract

This paper tests the effect of children on the chance to be self-employed. This is done through numerous regressions, making a distinction between having children and the amount of children. I find that both having children and the amount of children do not affect the chance of self-employment, but they do affect it indirectly through mechanisms. I have tested three of these mechanisms to try to explain how this works.

Introduction

Sometimes when people get children they need to be more flexible or just want to have a little more time. Changing from a wage job to a self-employed job does just that. Of course, there are also other reasons for people to be self-employed, but what I want to research in this paper is what kind of effect children have on self-employment. Thus the research question will be:

What is the effect of children on the chance to be self-employed?

To answer this question I will make use of the data from the LISS panel that surveyed 7000 people in the Netherlands on different subjects of which 'work and schooling', 'family and household', 'income', 'social integration and leisure' and 'background variables' are the ones that I will be using in this paper.

The definition of self-employment differs depending on who (or what) you ask. Basically it is any form of employment where the person in employment is not working for a boss and is independent of most people, the exception being a business partnership (Investopedia, n.d. para. 2). A self-employed person can thus be an independent contractor, a hair salon owner or anything of the like. They can set their own hours and thus work whenever they want. However, a self-employed person does not earn a salary or wage and has to make money by making profits through their business operations. They are also responsible for paying their own taxes as opposed to a wage job where the employer usually does that. Furthermore, there are no employee benefits and the risk of self-employment is higher than that of a wage job.

A self-employed person is thus different from a business owner. As an article on Investopedia states it "A business owner is someone who owns a company but does not work with the day-to-day operation of the company. In contrast, a person who is self-employed owns his own business, of which he is also the primary or sole operator." To come back to the previously mentioned example of the hair salon owner, that means that that owner also works in their own salon and does not just own it. This is an important distinction to make, as I will be working with the definition of self-employment as described and not with the definition of business owners.

There is, of course, already literature available that describes the effects of children on the chance to be self-employed. However, most (or all) of these works treat that effect as a secondary thing and the authors are not necessarily interested in that effect. As far as I know, a paper that focused on this as a main subject and has more or less the same research question, has not been published as of yet. Other papers that do contain children as a variable in their calculations make use of datasets from different countries. For example Rees and Shah (1986) and Burke et al. (2002) use a dataset from the UK and Aronson (1991) uses data from the USA, whereas I use data from the Netherlands. Thus, there could be differences in the outcomes. The reason why I am interested in the effect of children on self-employment, is because self-employment is an important part of the economy as a whole and the effect of children has not been researched a lot yet and that effect differs in the few papers that even discussed it at all. As Goetz et al. (2012) state it "recent studies suggest that self-employment has tangible positive economic impacts not only on wage and salary employment, but also on per capita income growth and poverty reduction."

Burke et al. (2002) have done some research on the determinants of self-employment and found that having children has a positive effect on the chance to be self-employed for males and what they call "less qualified females". On the other hand Rees and Shah (1986) have also done research on the determinants of self-employment. But, their research was not solely focused on this. Their main focus was on a self-employment/paid-employment choice problem and they used a probit model to predict the probability of self-employment. I will be doing this a little bit differently, but that will be discussed in the next section. One of their variables was also children and they found that it was not a significant determinant of self-employment, although they do not refer to it anywhere in the text other than the descriptive

statistics. However, things can change. Their research was done in the UK in 1986 , more than 30 years ago, and I want to know if this is still the case or that people have some reason to become self-employed more now because of their children. This seems to be the case as the results from Burke et al. (2002) seem to contradict the results from Rees and Shah (1986).

Since the paper from Burke et al. (2002) is more recent than that of Rees and Shah (1986) and it is more focused on the effect of children on self-employment, I will formulate a hypothesis that is based on their results. The first hypothesis is therefore:

H1: the probability to be self-employed increases when people get children

I also think that, perhaps at a certain point, people will want to be a little more flexible when they get a lot of children and maybe become self-employed. Aronson(1991) found that the chance to be self-employed is higher with more children in the household. Based on that my second hypothesis will be:

H2: the chance to be self-employed for a person increases when their number of children rises.

I also want to know what exactly causes the effect of children on self-employment, if there are any causes at all. More specifically, I think there exist certain mechanisms that affect the chances to be self-employed. I will test the effect of three possible mechanisms which will be more elaborately described in the methodology section. Basically, I will test the effect of the amount of leisure time, how that leisure time is spent and how many hours a person works on self-employment and how, and if, they are influenced by having children and the amount of children.

Based on the paper of Hyttinen and Ruuskanen (2007), who find that the self-employed have less leisure time, I think that the less leisure time someone has the more chance they have to be self-employed. The variable in the dataset that is used for this is described as the satisfaction of the amount of leisure time. Thus my next hypothesis is:

H3: The satisfaction of the amount of leisure time is a mechanism and if it goes down, the probability to be self-employed increases.

Then there is the matter of flexibility which, I think, is the most important reason why someone would want to become self-employed. Hyytinen and Ruuskanen (2007) also find that the self-employed who have small children are more likely to work after 5 p.m.. This is a sign of flexibility for the self-employed and thus the fourth hypothesis will be:

H4: The satisfaction of the way in which leisure time is spent (flexibility) is a mechanism and if it goes up, the probability to be self-employed increases

The amount of hours worked should increase if you have less leisure time available. Furthermore, Hyytinen and Ruuskanen (2007) also find that the self-employed work longer effective hours. Therefore my last hypothesis will be:

H5: The amount of hours worked is a mechanism and if it goes up, the probability to be self-employed increases.

How I will test these hypotheses will be handled in the next section.

Methodology

As I have mentioned, I will use data from the LISS panel that tracks a number of individuals from 2008 till 2015, asking them to fill in surveys on different subjects. These subjects include the variables that I will be using in this paper. I want to find out what the effect of children is on the chance to be self-employed and for that I will, of course, need to make a regression analysis with self-employment as the dependent variable and children as the independent variable, but there might be other factors that influence the probability of a person to be self-employed. Thus I will use some control variables of which I think might affect the chance of a person to be self-employed. The control variables I will use are gender, education, income, amount of working hours, satisfaction of the amount of leisure time and flexibility (satisfaction of the way in which leisure time is spent). All these variables could have implications for being self-employed. Mothers might be more inclined to stay home and care for their children than fathers are, making gender a possible control variable. The type of education could also affect whether someone becomes self-employed or might make more money in a wage job. A person with a good and steady income might choose self-employment earlier (or not) than someone who is just scraping by. The fact that

Hyytinen and Ruuskanen (2007) find that the self-employed have less leisure time, means that this could also affect the chance to be self-employed, just like flexibility and amount of hours worked which are closely tied to each other.

To add to that, Blanchflower (2000) found that self-employment is higher among men than among women and Boden (1999) finds that men and women generally have different reasons to become self-employed. This justifies the use of gender as a control variable. Rees and Shah (1986) also found that education and age are significant determinants of self-employment, justifying the use of education as a control variable. However, I will not use age as a control variable, since that is correlated with income and will be included in the individual fixed effects regression I will also be using. Unfortunately, there is no literature available on the effect of income on self-employment, but I think if a person has a good income with the job they have now, they might be less inclined to give up that extra income and start working for themselves. Based on this I will thus also use income as a control variable.

Thus the first regression will look like this: $self\ employment = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * education + \beta_4 * netinc + \beta_5 * workhours + \beta_6 * leisure + \beta_7 * Wleisure + \varepsilon$

Where α is the constant and ε the error term. So, self-employment is the dependent variable and is defined in the dataset as an autonomous professional, freelancer, or self-employed. It is a value of 1 when self-employed and 0 when not self-employed. This will stay the same in every following regression that uses self-employment as the dependent variable. The variable *wchildren* means whether a person has children or not and is the independent variable. The rest are control variables as described. Gender is a value of 1 for males and 2 for females, *netinc* means the amount of net income per month, but education is a bit more elaborate. So, to elaborate, the types of education have been given numbers in the dataset. The higher this number is, the higher the education. For example, an education value of 4 is an MBO (middelbaar beroepsonderwijs) and a value of 6 is a WO (wetenschappelijke opleiding). *Workhours* is the amount of hours worked, *leisure* is the satisfaction of the amount of leisure time and *Wleisure* is the flexibility, described in the dataset as the satisfaction of the way in which leisure time is spent. Also, I will use a p-value of 0.05 to determine the significance of a variable.

I will also use a second regression to test the effect of the amount of children on the chance to be self-employed. Thus, the second linear regression I will be using is:

$$\text{self employment} = \alpha + \beta_1 * \text{amchildren} + \beta_2 * \text{gender} + \beta_3 * \text{education} + \beta_4 * \text{netinc} + \beta_5 * \text{workhours} + \beta_6 * \text{leisure} + \beta_7 * \text{Wleisure} + \varepsilon$$

In this regression the variable amchildren is the amount of children a person has. This regression will be used to (partly) test the second hypothesis. The rest is the same as in the previous regression.

The shortcomings of this model are that there could always be omitted variables, such as the simple want of someone to be self-employed, because it seems like fun to that person or other factors like these that cannot be taken into account in this regression. It could also be that some variables change over time, of which the most important are wchildren and amchildren. In an attempt to fix this I will also be running two individual fixed effects regressions. This way I can observe whether there is a relationship between the amount of children (or having children) and self-employment over time for an individual, since an individual fixed effects model compares an individual with their selves in different time periods. Thus I would be able to establish whether the change in wchildren or amchildren increases or decreases the chances to be self-employed for an individual.

The individual fixed effects regression for hypothesis 1 will look like this:

$$\text{self employment} = \alpha_i + \beta_1 * \text{wchildren}_i + \beta_2 * \text{netinc}_i + \beta_3 * \text{workhours}_i + \beta_4 * \text{leisure}_i + \beta_5 * \text{Wleisure}_i + \text{year}_t + \varepsilon_{it}$$

Where again α is the constant and ε is the error term. Also, i is the index for the individual and t is the index for the time period. Again, wchildren is the variable for whether a person has children or not. Netinc is again the net income per month and is included because this variable is time-variant. Workhours is the amount of hours a person works per week. This variable is also controlled for, because it could affect the outcome and the amount of hours worked also tends to change over time and individual fixed effects regressions only account for time-invariant variables (such as gender, which does not change over time). Leisure is the satisfaction of the amount of leisure time a person has and is measured on a scale of 0 to 10 where 0 is "not satisfied at all" and 10 is "completely satisfied". This variable is included for the same reason as workhours; it is time-variant. Wleisure is the satisfaction of the way in

which leisure time is spent (flexibility) and is also measured on a scale of 0 to 10 and included for the same reason as the other variables described above. Also, I will use a p-value of 0.05 to determine whether the variable is significant or not in this regression as well.

The individual fixed effects regression for hypothesis 2 will look like this:

$$\text{self employment} = \alpha_i + \beta_1 * \text{amchildren}_i + \beta_2 * \text{netinc}_i + \beta_3 * \text{workhours}_i + \beta_4 * \text{leisure}_i + \beta_5 * \text{Wleisure}_i + \text{year}_t + \varepsilon_{it}$$

Where all variables mean the same as in the previous regression, except for *amchildren* which is the amount of children.

The reason why I have not included gender and education in the IFE regressions is because gender is time-invariant; it does not change over the course of time and is thus already captured in the IFE. Education, however, can change, but in the dataset I use it does not change for most people, only for some. Because of this, including it in the IFE regression could cause problems or inaccurate results. So, to avoid this I will not use education either in the IFE regressions.

With these regressions I should be able to get a good idea of the effects that the amount of children and having any children at all have on the chance to be self-employed and whether it has a significant effect at all.

There is also a chance that certain mechanisms exist. One of which I argued would be the amount of spare time a person has. For instance, when someone has children they most likely have less spare time and might want to make up for it by working less or spending their time differently.

So, I will check whether children increase or decrease the satisfaction of the amount of leisure time that a person has, and whether this has an effect on the chance to be self-employed. For this I will be running three linear regressions.

$$\text{leisure} = \alpha + \beta_1 * \text{wchildren} + \beta_2 * \text{gender} + \beta_3 * \text{netinc} + \beta_4 * \text{education} + \varepsilon$$

$$\text{leisure} = \alpha + \beta_1 * \text{amchildren} + \beta_2 * \text{gender} + \beta_3 * \text{netinc} + \beta_4 * \text{education} + \varepsilon$$

Where leisure is the dependent variable and describes the satisfaction of the amount of leisure time that a person has. It has been given a value between 0 and 10 (0 and 10 included) as explained earlier. Amchildren is again the amount of children a person has and wchildren is whether a person has children or not. These two variables about children will be the independent variables in each respective regression and the rest are control variables as described earlier. With this regression I will test whether the amount of children and having any children at all have an effect on the satisfaction of the amount of leisure time that a person has.

After that I will use the regression: $self\ employment = \alpha + \beta_1 * leisure + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education$

This regression will be used to determine whether the satisfaction of the amount of leisure time a person has, has an effect on the chance to be self-employed. Leisure will in this case be the independent variable and self-employment will be the dependent variable, the rest will again be control variables.

Another mechanism I will be testing is the satisfaction of the way in which leisure time is spent (flexibility). I think that not only the satisfaction of the amount of time will be influenced by children and influences self-employment probability, but also the way in which that leisure time is spent or divided.

Again, I will use three linear regressions to test whether children have an effect on the satisfaction of the way in which leisure time is spent, and whether this has an effect on the chance to be self-employed. The first regression will be used to test whether having children has an effect on the satisfaction of the way in which leisure time is spent and will look like this: $Wleisure = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

Where all the variables mean the same as in the previous regressions for the first mechanism, with the exception being the dependent variable which is now Wleisure. This variable describes the satisfaction of the way in which leisure time is spent.

The second regression will then look like this: $Wleisure = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

This regression tests the effect of the amount of children on the satisfaction of the way in which leisure time is spent. All variables are the same as in the last regression with *wchildren* being swapped out for *amchildren* which describes the amount of children.

The last regression of this mechanism which is used to test the effect of *Wleisure* on self-employment looks like this: $self\ employment = \alpha + \beta_1 * Wleisure + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education$

All variables have already been described and mean the same as in the previous regressions.

The third and last mechanism I will test is the effect of the amount of hours worked on the chance to be self-employed and how the amount of hours worked is influenced by children.

For this I will again be running three standard linear regressions. The first one looks like this:

$$workhours = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$$

It tests the effect of having children on the amount of work hours where *workhours* is the amount of hours a person works per week and the dependent variable. The rest of the variables are the same as in the previous regressions and *wchildren* is the independent variable.

The second regression for the third mechanism looks like this: $workhours = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

It tests the effect of the amount of children on the amount of work hours per week. Again, the other variables are the same as in the previous regressions and *amchildren* is the independent variable.

The last regression will be used to test the effect of the amount of work hours on the chance to be self-employed. The regression for that looks like this: $self\ employment = \alpha + \beta_1 * workhours + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education + \varepsilon$

Where all the variables mean the same as already described and *workhours* is the independent variable.

The reason why I will use loose standard linear regressions is because no other single regression can check for mechanisms in my case. One might think that an instrumental variable regression would be good for this, but as Dippel et al. (2017) state it "using an

Instrumental Variable (IV) identifies the causal effect of the endogenous treatment on both the intermediate and the final outcome variable, but not the extent to which the intermediate variable affects the final outcome” and I am interested in the extent to which the intermediate variable (leisure, for example) affects the final outcome (self-employment). Thus, an IV regression cannot be used in my case.

An issue with the data, however, could be that there is self-selection of people into self-employment. However, Hamilton (2000) finds that the earnings differential between paid-employment and self-employment cannot be explained by the selection of low-ability workers into self-employment. This means that at least the skills and abilities of self-employed workers should be more or less random, although this cannot be said for other characteristics. For instance, Boden (1996) finds that there is a relatively strong, positive influence of fertility upon women’s selection into self-employment. He also finds, in a later paper that he wrote in 1999, that women’s and men’s reasons for getting into self-employment are different and that especially women with young children give flexibility and family related things as a reason to get into self-employment. This could mean that one of the two genders is more likely to be self-employed than the other.

Data

In this paper I make use of data of the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands). The LISS panel is a representative sample of Dutch individuals who participate in monthly internet surveys. The panel is based on a true probability sample of households drawn from the population register. Households that could not otherwise participate are provided with a computer and Internet connection. A longitudinal survey is fielded in the panel every year, covering a large variety of domains including work, education, income, housing, time use, political views, values and personality.

The total amount of observations in my dataset is 95.576 over 8 years; from 2008 till 2015. Although, there are some missing values. Not all participants filled in the survey every year and some other participants did not fill in everything even though they did participate in the survey each year. The amount of missing values thus differs for each variable.

The amount of people that switch from self-employment to paid employment in the dataset totals exactly a hundred. What is interesting though, is that nobody in the dataset switches from paid-employment to self-employment. This is not a problem though, as long as they do actually switch jobs, but it is noteworthy.

The most important parts of the data are described in the table below.

	Amount (description... whatever...)	Mean	Standard deviation
Observations	95576		
Years	From 2008 till 2015 (min 2008, max 2015)		
Switched from self-employment to paid-employment	100		
Amount of children		1.270	1.270
Average amount of hours worked by the self-employed (per week)		33.83	18.79
Average amount of hours worked by the paid-employed (per week)		31.78	14.01
Average satisfaction of the amount of leisure time by the self-employed		6.790	2.267
Average satisfaction of the amount of leisure time by the paid-employed		6.491	2.008
Average satisfaction of the way in which leisure time is spent by the self-employed		6.942	1.812
Average satisfaction of the way in which leisure time is spent by the paid-employed		6.861	1.643

As can be seen the average amount of children is 1.27. The average amount of hours worked is higher for the self-employed than it is for the paid-employed, which is in line with the research from Hyytinen and Ruuskanen (2007). The average satisfaction of the amount of leisure time and the way in which leisure time is spent are also both higher for the self-employed than for the paid-employed. This seems to be in contrast to hypothesis 3, but in line with hypothesis 4. However, these are just the means and no conclusions can be made from them regarding the hypotheses. This will be done in the next section.

Results

The first linear regression looked like this: $self\ employment = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * education + \beta_4 * netinc + \beta_5 * workhours + \beta_6 * leisure + \beta_7 * Wleisure + \varepsilon$

In this regression the variable *wchildren* describes whether a person has children or not, *gender* is a value of 1 when male and 2 when female, *netinc* is the net income of a person per month and *education* describes the type of education (the higher the number for this variable, the higher the education, as described in the methodology section). The dependent variable is self-employment and is equal to a value of 1 when a person is self-employed and equal to a value of 0 when not self-employed.

The second linear regression looked like this: $self\ employment = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * education + \beta_4 * netinc + \beta_5 * workhours + \beta_6 * leisure + \beta_7 * Wleisure + \varepsilon$

Here *amchildren* means the amount of children and the rest of the variables are the same as in the previous regression.

To more accurately test the hypotheses I will also take the individual fixed effects regressions into account. The first individual fixed effects regression looked like this: $self\ employment = \alpha_i + \beta_1 * wchildren_i + \beta_2 * netinc_i + \beta_3 * workhours_i + \beta_4 * leisure_i + \beta_5 * Wleisure_i + year_t + \varepsilon_{it}$

The second individual fixed effects regression looked like this: $self\ employment = \alpha_i + \beta_1 * amchildren_i + \beta_2 * netinc_i + \beta_3 * workhours_i + \beta_4 * leisure_i + \beta_5 * Wleisure_i + year_t + \varepsilon_{it}$

The results of all regressions are shown in the table below and each regression will be discussed in order.

Self-employment	Reg 1	Reg 2	IFE 1	IFE 2
wchildren	-.0083 (.0030)		-.0058 (.0042)	
amchildren		-.0005 (.0012)		.0016 (.0019)
gender	-.0236 (.0031)	-.0234 (.0031)		
netinc	4.63e-07(3.31e-07)	4.94e-07 (3.31e-07)	1.29e-07 (3.49e-07)	1.38e-07 (3.49e-07)
education	.0084(.0009)	.0084 (.0009)		
workhours	.0007 (.0001)	.0007 (.0001)	.0001 (.0000)	.0001 (.0000)
leisure	.0012 (.0008)	.0015 (.0008)	.0006 (.0005)	.0006 (.0005)
Wleisure	-.0014 (.0010)	-.0013 (.0010)	-.0006 (.0007)	-.0006 (.0007)
_cons	.0550 (.0106)	.0481 (.0105)	.0610 (.0064)	.0555 (.0063)
Year 2009			.0043 (.0025)	.0043 (.0025)
Year 2010			.0045 (.0026)	.0048 (.0026)
Year 2011			.0032 (.0028)	.0035 (.0028)
Year 2012			.0062 (.0029)	.0066 (.0029)
Year 2013			.0087 (.0029)	.0092 (.0029)
Year 2014			.0127 (.0030)	.0134 (.0030)
Year 2015			.0144 (.0031)	.0152 (.0031)

Note: the standard errors are in parentheses.

As it turns out, the independent variable wchildren in the first regression has a p-value of 0.006. This is lower than the p-value of 0.05, which I use to determine the significance of a variable. This means that the variable has a significant effect. Thus the results of this

regression imply that having children has a negative effect (because of the negative coefficient) on the chance to be self-employed. This is evidence for the first hypothesis to be rejected, however a standard linear regression can contain omitted variables and other errors, and measures things a little bit differently than an IFE regression, so I will not make any conclusions yet.

What is also an interesting result is that the variable 'education', which describes the type of education, is significant and positive in both the linear regressions. This means that the higher the education, the higher the chance to be self-employed. This is not in line with previous research done by Blanchflower (2000), who found that lower educated people have a higher chance to be self-employed. This might, however, be due to the way of measuring education. As mentioned, in my model the higher the number for education, the higher the education is that a person has completed. Blanchflower, on the other hand, uses a system where he defines people that left school at age 14 or lower as lowly educated and people that left school at age 22 or higher as highly educated. These are a bit similar, because mostly highly educated people are also older, but that does not always have to be so. Also, he does not specify what kind of education the people of ages in between 14 and 22 have, whereas my dataset does.

The second regression shows a negative and insignificant effect of amchildren (amount of children) on the chance to be self-employed. The p-value is 0.647 which is obviously higher than 0.05. This implies that the amount of children has no effect on the chance to be self-employed and is evidence for hypothesis 2 to also be rejected. Again, however, I will make no conclusions yet.

So, to sum it up, the first regression shows that when a person has children, they have a lower probability to be self-employed. The second regression shows that the amount of children has an insignificant effect on the probability to be self-employed, which is in contrast to the research done by Aronson (1991), since he found that more children implies a higher likelihood of being self-employed. This can, however, be due to the fact that I use a different dataset from a different country.

The first individual fixed effects regression shows an insignificant and negative effect of wchildren (that is whether a person has children or not) on the chance to be self-employed.

This is in contrast with the first linear regression which gave a significant effect. However, a linear regression can contain omitted variables which might not be present in an IFE regression. An IFE regression controls for any time-invariant factors, but this does not mean that there are no omitted variables in the IFE regression, it just means there is a smaller chance of it. Because the IFE regression compares individuals with themselves in different time periods, this regression should determine if the change in having children has an effect on the chance to be self-employed and should therefore be a bit more precise than the linear regression.

In the second individual fixed effects regression the variable *amchildren* has a p-value of more than 0.05. The p-value for this variable is 0.394, which is much too high for any significant effect, meaning that the amount of children has an insignificant effect on self-employment through this regression, which again compares individuals with themselves in different time periods.

So, the linear regression on whether a person has children or not gives a negative and significant effect, whereas the IFE regression gives an insignificant effect of *wchildren* on the chance to be self-employed. This might be a bit odd at first sight, as the linear regression should give a good insight into the effects of the used variables, but as was mentioned earlier, a linear regression can have omitted variables or other errors which are (almost) not present in an individual fixed effects regression. An IFE regression also compares individuals with themselves in different time periods and it controls for time-invariant factors.

Therefore, the IFE regression should be somewhat more reliable and precise. Based on this I can reject my first hypothesis which stated that the probability to be self-employed increases when people get children. In other words, my models suggest that if someone has children, they have the same chance to be self-employed which is not in line with my first hypothesis.

Because the first linear and the first IFE regressions differ, where the linear regression shows a significant effect and the IFE regression shows an insignificant effect, we can say that having children is not a determinant of self-employment, because the IFE regression weighs heavier. In other words, the probability to be self-employed does not depend on whether a person has children or not and thus the first hypothesis is rejected. This is in line with previous literature, namely Rees and Shah (1986), who also never found any significant

effect of having children on the probability to be self-employed, but did find other significant determinants, such as education and age. But it is also in contrast to Burke et al. (2002) who did find a significant effect of having children on the chance to be self-employed, albeit only for males and less qualified females.

The second linear and IFE regression both give an insignificant effect of the amount of children on the chance to be self-employed. Therefore I will also reject my second hypothesis which stated that the chance to be self-employed for a person increases when their number of children rises. The regressions showing an insignificant result means that this cannot be the case and thus it is rejected.

So both hypotheses are rejected. This means that the chance to be self-employed is apparently the same for people with children as it is for people without children. Regarding the second hypothesis; my results suggest that the chance to be self-employed does not rise with the number of children. My results are in line with the research done by Rees and Shah (1986) but not with those of Aronson (1991) and Burke et al.(2002). The reasons for why this might be, are discussed in the next section.

The first linear regression that was used to test the first possible mechanism, whether children have an effect on the satisfaction of someone's amount of leisure time, looked like this: $leisure = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

The other linear regression that I used to test whether the amount of children has an effect on the satisfaction of the amount of leisure time looked like this: $leisure = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

The first regression for the second possible mechanism, whether children have an effect on the satisfaction of the way in which leisure time is spent, looked like this: $Wleisure = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

The second regression for the second possible mechanism looked like this: $Wleisure = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

The results are shown in the next table, but the results for the third mechanism are in a separate table that will be discussed after this one.

(W)Leisure	Reg 1 (leisure)	Reg 2 (leisure)	Reg 3 (Wleisure)	Reg 4 (Wleisure)
Wchildren	-1.074 (.0192)		-.5719 (.0161)	
amchildren		-.4266 (.0083)		-.2307 (.0069)
gender	-.0554 (.0191)	-.0598 (.0192)	-.1456 (.0160)	-.1477 (.0161)
netinc	1.04e-06 (1.85e-06)	1.62e-06 (1.86e-06)	3.47e-06 (1.56e-06)	3.74e-06 (1.56e-06)
education	-.1145 (.0058)	-.1184 (.0058)	-.0601 (.0049)	-.0621 (.0049)
constant	8.175 (.0398)	8.101 (.0399)	7.802 (.0334)	7.765 (.0334)

Note: the standard errors are in parentheses.

In the first regression, whether someone has children or not has a significant and negative effect on the satisfaction of leisure time (p-value is lower than 0.05). This means that if a person has children, the lower their average satisfaction of their amount of leisure time will be.

The second regression also has a significant and negative effect of the amount of children on the satisfaction of the amount of leisure time. So, this implies that the more children a person has, the less satisfied they will be, on average, with their amount of leisure time.

So, having children and the amount of children affect the satisfaction of leisure time negatively. This means that the satisfaction of leisure time of a person goes down.

The results for the third regression imply a negative and significant effect of wchildren on Wleisure. This means that having children decreases the satisfaction of the way in which a person spends their leisure time.

The results for the fourth regression imply a significant and negative effect of the amount of children on the satisfaction of the way in which leisure time is spent. This means that the more children a person has, the less satisfied they will be with the way in which their leisure time is spent.

So, having children and the amount of children also affect the satisfaction of the way in which leisure time is spent negatively.

The third and last possible mechanism that was tested was the amount of hours worked per week. The first regression to test this mechanism looked like this: $workhours = \alpha + \beta_1 * wchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

The second regression looked like this: $workhours = \alpha + \beta_1 * amchildren + \beta_2 * gender + \beta_3 * netinc + \beta_4 * education + \varepsilon$

Where workhours is the amount of hours worked per week and the dependent variable in both regressions. Wchildren and amchildren are again the independent variables belonging to their respective regressions.

The results of these regressions are shown in the next table.

Workhours	Reg 1	Reg 2
Wchildren	-3.454 (.1640)	
Amchildren		-1.661 (.0698)
Gender	-10.035 (.1646)	-10.039 (.1642)
Netinc	.0002 (.0000)	.0002 (.0000)
Education	1.4685 (.0533)	1.459 (.0532)
Constant	40.349 (.3544)	40.319 (.3515)

Note: the standard errors are in parentheses.

Wchildren and amchildren both have a significant effect on the total hours of work per week (p-value is 0.000 for both). So, from these results we can conclude that workhours is negatively affected by having children and the amount of children. When someone has children they will work less on average. The same goes for the amount of children; the more children a person has, the less they will work on average.

That brings us to the next three and last regressions for the mechanisms. The first one was used to test whether the satisfaction of the amount of leisure time has an effect on the chance to be self-employed. It looked like this: $self\ employment = \alpha + \beta_1 * leisure + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education$

The regression for the second possible mechanism looked like this: $self\ employment = \alpha + \beta_1 * Wleisure + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education$ and was used to test the satisfaction of the way in which leisure time is spent.

The regression for the third possible mechanism, the effect of the amount of hours worked on self-employment, looked like this: $self\ employment = \alpha + \beta_1 * workhours + \beta_2 * wchildren + \beta_3 * amchildren + \beta_4 * gender + \beta_5 * netinc + \beta_6 * education$

The results of the above three regressions are shown in the table below.

Self-employment	Reg 1 (leisure)	Reg 2 (Wleisure)	Reg 3 (workhours)
Leisure	-.0035 (.0005)		
Wleisure		-.0024 (.0005)	
workhours			.0007 (.0001)
wchildren	-.0074 (.0041)	-.0052 (.0040)	-.0231 (.0054)
amchildren	.0058 (.0017)	.0059 (.0017)	.0077 (.0023)
Gender	-.0257 (.0020)	-.0260 (.0020)	-.0244 (.0030)
Netinc	5.13e-07 (1.97e-07)	5.23e-07 (1.98e-07)	4.61e-07 (3.27e-07)
Education	.0080 (.0006)	.0084 (.0006)	.0082 (.0009)
Constant	.0819 (.0059)	.0717 (.0062)	.0548 (.0074)

Note: the standard errors are in parentheses.

The results indicate that the variable leisure has a significant and negative effect on the chance to be self-employed. This means that if someone is very satisfied with their leisure time, their chance to be self-employed decreases. If someone is not satisfied with their leisure time, they have a higher chance to be self-employed.

From this table we can see that if someone is less satisfied about their amount of leisure time, their chance to be self-employed is higher than when someone is more satisfied about their amount of leisure time. Having children and the amount of children affect the satisfaction of the amount of leisure time negatively. This most likely means that a person has less leisure time available and is thus less satisfied with it. Because of this, and the fact that a lower satisfaction of the amount of leisure time means a higher chance to be self-employed, I can accept my third hypothesis which stated that the satisfaction of the amount of leisure time is a mechanism and if it goes down, the probability to be self-employed increases.

There is also a significant and negative effect of $W_{leisure}$ on self-employment. So, the satisfaction of the way in which leisure time is spent (flexibility) has a negative effect on the chance to be self-employed and is negatively influenced by having children and the amount of children. In other words, the more children a person has, the less satisfied they will be with the way in which they spend leisure time, and the more chance they have to be self-employed compared to others with higher satisfaction. This means that I will have to reject my fourth hypothesis which stated that the satisfaction of the way in which leisure time is spent (flexibility) is a mechanism and if it goes up, the probability to be self-employed increases. It is indeed a mechanism, but my results suggest that if the satisfaction of the way in which leisure time is spent goes up, you will have less chance to be self-employed. This seems to be a little odd, as flexibility is one of the main reasons why people become self-employed. However, this might be due to the way in which the variable is described. For lack of a better variable for flexibility I have used this one and it is closely tied to the satisfaction of the amount of leisure time. This may be why they both have the same effect. Also, when less leisure time is available, as was found to be true by Hyytinen and Ruuskanen (2007) for the self-employed and backed up by my results, people may actually be less flexible and thus less satisfied with the way in which their leisure time is spent.

The amount of hours worked per week also has a significant effect on the probability to be self-employed. It is positive and that means that the more a person works, the more chance they have to be self-employed. It is also a mechanism, because it is influenced by having children and the amount of children. This means I can accept my fifth and last hypothesis which stated that the amount of hours worked is a mechanism and if it goes up, the probability to be self-employed increases. But, compared to the other tested mechanisms, this one does have some odd results. As can be seen having children and the amount of children have a negative effect on the amount of work hours. But, workhours has a positive effect on self-employment. This means that a person will work less if they have children and even less if they have more than one child and the consequence of that is that their chance to be self-employed is smaller when they have children. This is not really contradicting anything, as Hyytinen and Ruuskanen (2007) indeed found that the self-employed tend to have less leisure time and should therefore work more, but the results from this mechanism suggest that having (more) children indirectly leads to a lower chance to be self-employed, whereas the mechanism of the satisfaction of the amount of leisure time suggests the opposite; having (more) children through that mechanism indirectly leads to a greater chance to be self-employed. This does not have to be wrong though, as different mechanisms can have different effects, but I thought it was noteworthy. Also, if we see the satisfaction of the amount of leisure time simply as the amount of leisure time then the results are actually not strange at all, since more leisure time means less working hours and vice versa. Thus, leisure having a negative effect and workhours having a positive effect on self-employment then actually makes good sense.

Conclusion

So, the first two hypotheses are rejected. This means that the chance to be self-employed is the same for people with children as it is for people without children, because of the insignificant effects. My results also suggest that the chance to be self-employed does not depend on the number of children that a person has. As was mentioned before, these results are in line with the research done by Rees and Shah (1986), but in contrast to that of Aronson (1991) and Burke et al. (2002). This could, however, be due to the fact that I used a different dataset. Aronson, for example, used data mainly from the United States, whereas I

used data from the Netherlands. It could be that people from the Netherlands feel different about their work and children than people in the United States do. Rees and Shah (1986) and Burke et al. (2002) used data from the United Kingdom and Rees and Shah (1986) also used a slightly different way of testing the determinants of self-employment, which might be why they contradict other literature, but my results do suggest the same as theirs. So, this does not mean that this is the only reason for the results to differ, but it is a likely one. Perhaps it is true that the effect of children has changed over time as I mentioned in the introduction, but we cannot know this for sure without further research.

It also turns out that the satisfaction of the amount of leisure time and the satisfaction of the way in which this leisure time is spent as well as the amount of hours worked per week, are mechanisms influenced by children and the amount of children. The first mechanism (leisure) and the second mechanism (Wleisure) then have a negative effect on the chance to be self-employed. Thus if someone has, say, two children, their satisfaction of the amount of leisure time and the way in which it is spent will be lower than for someone who has no children at all or only one child. This then makes their chance to be self-employed higher than for the person with no children (or one child). The mechanism for the amount of worked hours per week is negatively influenced by having children and the amount of children, but positively affects the chance to be self-employed as opposed to the other two mechanisms. Thus, if a person has (more) children, they will work less on average and their chance to be self-employed decreases.

To come back to my research question which was: What is the effect of children on the chance to be self-employed? The distinction is made between having children and the amount of children and my conclusion is that having children and the amount of children are not determinants for self-employment. If a person has children their chance to be self-employed is the same as for a person without children. The same goes for the amount of children. However, the mechanisms for the satisfaction of the amount of leisure time and the satisfaction of the way in which leisure time is spent both have a negative effect on the chance to be self-employed and are negatively influenced by having children and the amount of children. This means that, even though children do not have a direct effect on the chance to be self-employed, they do seem to have an indirect effect on the chance to be self-employed through leisure and Wleisure, according to my results. The same goes for the

amount of working hours, even though this affects the chance to be self-employed positively, children do indirectly affect the chance to be self-employed through this mechanism as well. So, to answer my research question somewhat simpler; there is no direct effect of children on the chance to be self-employed, but there does seem to be an indirect effect through the mechanisms.

Of course, there are some shortcomings to my research. As was mentioned, there might be omitted variables and other errors in the linear regressions. However, this should have been fixed in the IFE regressions, but it is possible that it was not completely fixed. Another shortcoming of my research was that I have only used data from people in the Netherlands. As I have said, this could be why my results are not in line with Aronson's (1991) research and the research by Burke et al. (2002), because it is probably not representative for groups of people outside the Netherlands. Then there is also the problem of self-selection. Some people that choose self-employment may already be intrinsically different than others. Unfortunately, there is no way for me to fix this with the dataset that I worked with, so some results could be inaccurate.

As for further research, it might be interesting to broaden or narrow down the subject a bit. For example, T. Bates (1995) has already done some research to self-employment entry across different industry groups and finds that different industries are characterized by different variables. The chance to be self-employed could then also be different for different industries and the effect of children might also be different per industry. For example, when a business man gets children he might be more inclined to become self-employed than a construction worker would. This could be because of various reasons and would be an interesting topic to do further research on. What could also be interesting is to get data from a broader region such as a continent, like Europe or North America, or even globally to test the effects of children on self-employment on a bigger scale. This is because my work was focused on data from the Netherlands, whereas Aronson's (1991) work was focused on data from the United States and the works from Rees and Shah (1986) and Burke et al. (2002) were focused on data from the United Kingdom and the results differ. So, I am interested in what the results would be on a more global scale.

References

- Aronson, R. L. (1991). *Self-employment: A labor market perspective* (No. 24). Ilr Pr.
- Bates, T. (1995). Self-employment entry across industry groups. *Journal of Business Venturing*, 10(2), 143-156.
- Blanchflower, D. G. (2000). Self-employment in OECD countries. *Labour economics*, 7(5), 471-505.
- Boden, R. J. (1996). Gender and self-employment selection: An empirical assessment. *The Journal of Socio-Economics*, 25(6), 671-682.
- Boden, R. J. (1999). Flexible working hours, family responsibilities, and female self-employment. *American Journal of Economics and Sociology*, 58(1), 71-83.
- Burke, A. E., Fitzroy, F. R., & Nolan, M. A. (2002). Self-employment wealth and job creation: The roles of gender, non-pecuniary motivation and entrepreneurial ability. *Small business economics*, 19(3), 255-270.
- Dippel, C., Gold, R., Heblich, S., & Pinto, R. (2017). *Instrumental Variables and Causal Mechanisms: Unpacking The Effect of Trade on Workers and Voters*(No. w23209). National Bureau of Economic Research.
- Goetz, S. J., Fleming, D. A., & Rupasingha, A. (2012). The economic impacts of self-employment. *Journal of Agricultural and Applied Economics*, 44(3), 315-321.
- Hamilton, B. H. (2000). Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political economy*, 108(3), 604-631.
- Hyytinen, A., & Ruuskanen, O. P. (2007). Time use of the self-employed. *Kyklos*, 60(1), 105-122.
- Rees, H., & Shah, A. (1986). An empirical analysis of self-employment in the UK. *Journal of applied econometrics*, 1(1), 95-108.

Scherpenzeel, A.C., & Das, M. (2010). "True" Longitudinal and Probability-Based Internet Panels: Evidence From the Netherlands. In Das, M., P. Ester, and L. Kaczmirek (Eds.), *Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. (pp. 77-104). Boca Raton: Taylor & Francis.

Self-employed (n.d.). In Investopedia. Retrieved from:

<http://www.investopedia.com/terms/s/self-employed.asp>