
The Determinants of Self-Employment - Evidence from Self-Employed Persons in Poland

Submitted 28/03/21, 1st revision 25/04/21, 2nd revision 26/05/21, accepted 30/06/21

Elżbieta Jędruczyk¹

Abstract:

Purpose: The primary purpose of the research is to build a model that makes it possible to define the categories of entities through which self-employment is used as a form of tax optimization. This would facilitate the creation of more efficient tax instruments that prevent false self-employment without discouraging real entrepreneurship.

Design/Methodology/Approach: The research method is a survey carried out from September 2020 on a representative sample of 400 Polish self-employed persons representing specific trades. Subsequently, machine learning classification algorithms were implemented to find factors characterizing self-employed persons with different tax-optimization attitudes.

Findings: The research results have shown that classification and regression tree (CART) and bootstrap aggregation (bagging) tree models can be helpful to determine whether a self-employed person is likely to use self-employment as a tax optimization method. The three main factors determining classification are the self-employed person's attitude towards risk, the industry in which they work, and the length of their business experience.

Practical Implications: The research performed made it possible to identify factors that affect self-employed people's attitude towards reducing their tax burden, in particular, to verify the scope of taxpayers becoming self-employed to reduce their tax burden (treating self-employment as a form of tax optimization), and to determine whether there is a way to shape tax rules to encourage entrepreneurship while minimizing the risk of abuse such as forcing employees to become self-employed or people registering as self-employed to avoid taxation.

Originality/Value: I have proposed the opposite than previous approach: this research is performed among self-employed persons to extend knowledge about their attitude toward becoming self-employed. This approach should be more efficient in understanding the process of individual decision-making between becoming an employee or self-employed.

Keywords: Self-employment, taxation, tax policy, classification trees.

JEL classification: G38, H21, H26, H30.

Paper Type: Research Paper.

Acknowledgments: The project is financed within the framework of the program of the Minister of Science and Higher Education under the name „Regional Excellence Initiative” in the years 2019-2022; project number 001/RID/2018/19; the amount of financing PLN 10,684,000.00.

¹ Department of Economics, Finance and Management, University of Szczecin, elzbieta.jedruczyk@usz.edu.pl;

1. Introduction

Self-employment is an ambiguous phenomenon as a method of reducing unemployment. On the other hand, self-employment as a method of reducing unemployment and promoting innovation and invention became particularly important when the industrial structure and technological change became inevitable (Startienė *et al.*, 2010).

Self-employed persons are a very heterogeneous group. A proportion of professionals and highly skilled workers are more likely to be independent contractors (Engblom, 2003). On the other hand, there is a significant proportion of dependent (fake, bogus) self-employment workers classified as self-employed but still economically dependent on a single contractor. In such cases, companies have often contracted out work to their former employees to downsize and thus gain flexibility and reduced labor costs (Thörnquist, 2013). Regarding low-skilled workers, especially in low-paid service jobs, some companies can consciously transfer costs, risks, and responsibilities to the workers (Thörnquist, 2013).

Self-employment is equated by some researchers with the concept of entrepreneurship, while the literature on the subject agrees as to the purposefulness of promoting entrepreneurial attitudes. On the other hand, self-employment is also perceived as a pathology both of the labor market (in the case of a so-called dependent, fake, false, or forced self-employment) and of the tax system, as (due to the taxation framework and as the self-employed are subject to social security contributions different from those in employment) the self-employed have broader possibilities of shaping the tax base. In a situation of significant differences in the tax burden of entrepreneurs and persons earning income from other sources, the decision on self-employment may be motivated by the desire to reduce the broadly understood tax contribution burden (Adamczyk *et al.*, 2018).

This results in legislators trying to prevent the aggressive use of self-employment to minimize the tax burden using, among others, complicated taxation restrictions. The problem is that the regulations implemented affect all self-employed persons, not only those whose goal is to reduce the tax burden.

2. Literature Review

Several researchers try to build decision models for becoming self-employed. Eisenhauer (1995) builds an economic model of the decision to be self-employed based on the expected utility gained, not simply from the prospective income streams, but also dependent on utility derived from the “working conditions” of employment versus self-employed alternatives (Eisenhauer, 1995). Douglas and Shepherd (2000) expand this approach they distinguish between entrepreneurial attitudes and entrepreneurial abilities and link an individual’s income potential to these abilities and attitudes. They investigate the “working conditions” regarding the individual’s

attitudes to specific working conditions such as effort required, risk exposure, and decision-making autonomy. They develop a theory of entrepreneurship that explains, in part, an individual's choice to be self-employed or to be an employee of an existing organization by utilizing a utility-maximization model (an individual will choose the career option that promises the most significant expected utility). They consider three central attitudes that differ between those intending to be self-employed and those intending to be employees attitudes toward hard work, financial risk, and decision-making autonomy together called "independence" (Douglas and Shepherd, 2000, 2002).

Out of many economic theories, the prospect theory developed by D. Kahneman and A. Tversky in the 1970s may help assess the process preceding the decision to become self-employed (especially tax). An essential feature of prospect theory is the statement that it is not the absolute level of wealth but changes in the level of wealth that are the carriers of the expected utility. The authors of the theory state that the individual is focused on assessing changes or differences, not absolute quantities, and perception depends on the point of reference. The following postulates of the described theory are the decreasing sensitivity to changes (the marginal subjective value of profits and losses decreases with their increase) and human aversion to incurring losses (the values of the aversion coefficient determined by Kahneman and Tversky for losses are more than twice as strong as for gains of the same amount) (Kahneman and Tversky, 1979).

The tax burden levied impacts the taxpayer's net earnings. In the context of the presented theory, it can be assumed that the individual determinants of self-employment will differ in significance depending on the individual situation of the subject, namely:

- in the case of entities with relatively low earnings, the marginal utility of additional funds obtained due to lower taxation will be higher, due to a bigger relative change in the level of assets;
- in the case of entities with a very high financial status, the marginal utility of additional funds obtained from the use of self-employment to reduce the tax burden will be lower (the perception depends on the reference point);
- the relative difference between the level of fiscal burden on individual forms of earning money, according to the perspective theory, maybe the determinant of decisions about self-employment; however, its significance in the case of entities with a relatively high level of assets may lose its importance.

Among other nonfiscal factors determining the decision to become self-employed, the following may be necessary: individual propensity to take risks; nature of business; access to professional knowledge (education); situation on the labor market.

The analyzed problem gains importance in times of economic crisis when pathologies are the result of the situation on the labor market "illegal" work, widespread use of

non-contract employment, and an increasing scale of self-employment defined as “dependent,” often forced by the former or potential employer.

There is some research on the determinants of false self-employment in Poland. Recently, D. Nikulin performed research using primary data obtained in a survey carried out between November and December 2018 on a representative sample of Polish small and medium-sized enterprises (SMEs). The research was conducted by a professional survey company using the CATI (computer-assisted telephone interview) method. As a result, 952 effective surveys were conducted covering characteristics related to the phenomenon of false self-employment (FSE) in Poland and opinions on entrepreneurship barriers (economic, institutional, legal). Respondents comprise owners and high-level managers, and the main goal of the study was to determine the characteristics of companies engaged in false self-employment in Poland (i.e., forcing employees to become self-employed) (Nikulin, 2020).

The downside of this approach is that when asking the employing companies, there is a risk of social desirability bias due to underreporting behaviors perceived in society as reprehensible (for instance, forcing employees to become self-employed against their own will). To avoid the risk, I have proposed the opposite approach – the research is performed among self-employed persons themselves. The study aims to find out if it is possible to use classification methods to predict their attitude towards using self-employment as a tax optimization method.

3. Research Methodology

A survey was carried out in September 2020 and October 2020 on a representative sample of $n=400$ Polish self-employed persons representing specific trades (only enterprises of natural persons sole traders registered in the CEIDG register). The research was conducted by a professional survey company, using both an online CAWI (computer-assisted web interview) panel, and computer-assisted telephone interview (CATI) methods. To ensure the representativeness of the research, stratified random sampling was performed, with the population of each trade as the strata. Then, within each group, random sampling was performed. Those chosen for the survey were the ones that provide an easy transition between certain forms of work supply (employment/self-employment). The distribution of the sample is presented in Table 1.

The research tool used the interview questionnaire consisted of eight sociodemographic questions and 36 substantive questions. The study used both closed and semi-open questions (with predefined answers), allowing the respondent to expand the proposed criteria with additional options. The questionnaire was intended to answer the following questions:

- What are the respondents' attitudes towards self-employment, and does this differ among different social groups?
- What are the characteristics of people more willing to become self-employed?

Is it influenced by family or social status? What is the impact of the industry in which the respondents operate? Is the decision on self-employment made voluntarily or under duress (forced by the labor market conditions or the dominant position of the employer/principal)? This is to identify "false self-employed" persons.

- How important (from the respondent's point of view) are certain factors for deciding to become self-employed (subjective assessment of the safety/certainty/stability of a given form of earning, tax burden carried, independence, labor market situation, etc.)?
- The classification tree method has been used to find the influence of the tax burden factor on becoming self-employed. Two different models were performed: a classification and regression tree (CART) model and a bootstrap aggregation (bagging) tree model.

Table 1. Overall sample description according to PKD²

	Frequency	Percentage	Percentage valid	Cumulative percentage
Computer programming, consultancy, and related activities, PKD 62	42	10.5	10.5	10.5
Activities auxiliary to insurance and pension funding, PKD 66	30	7.5	7.5	18
Real estate activities, PKD 68	18	4.5	4.5	22.5
Professional, scientific, and technical activities, Section M	151	37.8	37.8	60.3
Services to buildings, and landscape activities, PKD 81	18	4.5	4.5	64.8
Office administrative service activities, and other business support activities, PKD	12	3	3	67.8
Education, PKD 85	30	7.5	7.5	75.3
Arts, entertainment, and recreation activities, Section R	12	3	3	78.3
Other service activities, Section S	87	21.8	21.8	100
Total	400	100	100	

Source: Own study.

4. Research Results

The main findings of the exploratory analysis could be divided into two main topics: the scope of false self-employment in the sample and the importance of taxation factors importance. In terms of the scope of fake/false self-employment among the sample, 67% of the respondents answered that the company's/contractor's will was not an essential factor in becoming self-employed, 6.5% answered that the contractors

² PKD2007 the official Code List of Classification of Business Activities in Poland was used.

had a very high impact on the decision. Considering certain factors indirectly influencing fake (false, forced) self-employment:

- 90.3% of the respondents claimed that they use their assets to perform their economic activities, while only 9.8% use the company's assets;
- 76.8% of respondents decide about the time and place of supplying services on their own, while 14.3% work at a time and place that depends on the services delivered; 9% work at the time and place indicated by a customer.
- Regarding the importance of taxation as a factor, it was essential to determine whether the respondents have sufficient knowledge to use self-employment as a form of tax optimization. The survey showed that:
- 87.3% get help with fulfilling their tax duties, among them 75.3% decided on outsourcing, while 12% directly employ an in-house accountant;
- only 12.8% of the respondents answered that they are not interested in the topic of taxation at all, 68.3% received information from their accountant/tax adviser/legal adviser, and 44.5% from the Internet;
- 25% of respondents said they are not aware of the tax status of their companies, 18% do not know what form of income tax they pay, and 7.3% of respondents are not aware whether they are VAT payers or not;
- 50.5% of respondents claimed that they would not choose to become an employee, even if offered better earnings, 34% of the respondents claimed that self-employment increased their quality of life, and 58% claimed that their quality of life increased significantly.

In order to distinguish between different groups of respondents, as the next step of the research, classification tree models- analysis was performed. The primary purpose was to determine the variables affecting the importance of taxation as a factor among the respondents. In all models, the same variables were used, an approach that made it possible to compare the usefulness of each model type.

The survey question used as an indicator of the importance of tax in becoming self-employed was the question about the respondent's attitude towards the possibility of a hypothetical future transition into employment. There were four classes of dependent variable implemented: willing to change, even if the earnings offered would be lower; willing to change if the earnings offered were the same; only willing to change if the earnings offered were higher, and not willing to change. The independent variables considered in models were coded as follows:

Table 2. Independent variables

Factor	Code
how long has been self-employed	P18
industry	M0
self-assessment of attitude towards risk	P13_SQ001
scope of clients/contractors	P19
the financial situation of the family	P8
background (parents' employment status)	P5

the size of village/city/town where resident	M5
interest in taxation	P31_5

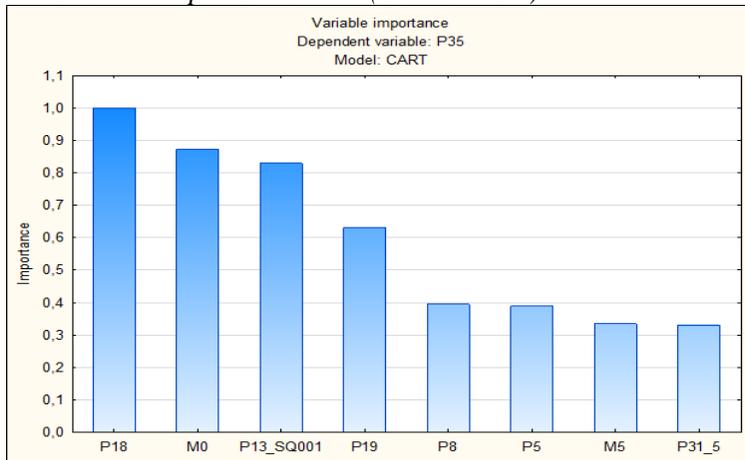
Source: Own study.

Of the two models, the performance of the CART and bagging tree models was the same, with an overall prediction error of 0.2275 (72.25% of cases were classified correctly by both models).

The first model implemented was a CART model. The equal costs of misclassifications and a priori probabilities calculated from the sample provided the best model performance. As the dependent variable was the ordinal variable, the Gini coefficient was used as a measure of inequality. The three most important variables for the model were: P18, the individual's experience; M0, the industry in which the entrepreneur operates; and P13_SQ001, the entrepreneur's attitude towards risk. The second model tested was a bagging tree classification model. The best performance was gained with a maximum size of seven leaves, and 389 calculations were performed.

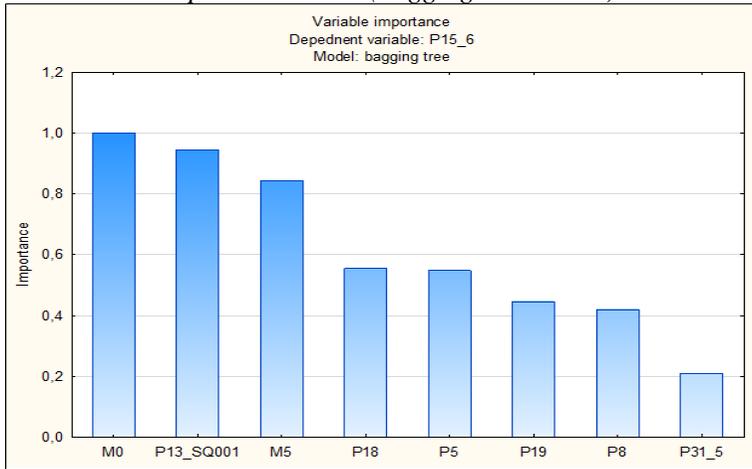
In this case, the most important variables were: P13_SQ001, the entrepreneur's attitude towards risk; M0, the industry in which the entrepreneur operates; and P5, the entrepreneur's origin.

Figure 1. The variable importance chart (CART model)



Source: Own study.

Figure 2. The variable importance chart (bagging tree model)



Source: Own study.

Table 3. Confusion Matrix (CART Model)

Items	Predicted: prefers employment, even if earns less	Predicted: prefers employment, only if earns the same	Predicted: would accept employment if earnings were higher	Predicted: would not change, even if earnings were higher	Total in row
Count	1				1
Column percentage		0.00%	0.00%	0.48%	
Row percentage	0.00%	0.00%	0.00%	100.00%	
Total percentage	0.00%	0.00%	0.00%	0.25%	0.25%
Count		8	7	9	24
Column percentage		47.06%	4.02%	4.31%	
Row percentage	0.00%	33.33%	29.17%	37.50%	
Total percentage	0.00%	2.00%	1.75%	2.25%	6.00%
Count		7	124	42	173
Column percentage		41.18%	71.26%	20.10%	
Row percentage	0.00%	4.05%	71.68%	24.28%	

*The Determinants of Self-Employment - Evidence from Self-Employed Persons
in Poland*

1022

Total percentage		0.00%	1.75%	31.00%	10.50%	43.25%
Count	would not change, even if earnings were higher		2	43	157	202
Column percentage			11.76%	24.71%	75.12%	
Row percentage		0.00%	0.99%	21.29%	77.72%	
Total percentage		0.00%	0.50%	10.75%	39.25%	50.50%
Total count	Total		17	174	209	400
Total percentage		0.00%	4.25%	43.50%	52.25%	

Source: Own study.

Table 4. Confusion Matrix (Bagging Tree Model)

	Items	Predicted: prefers employment, even if earns less	Predicted: prefers employment, only if earns the same	Predicted: would accept employment if earnings were higher	Predicted: would not change, even if earnings were higher	Total in row
Count	prefers employment, even if earns less	1				1
Column percentage		5.56%	0.00%	0.00%	0.00%	
Row percentage		100.00%	0.00%	0.00%	0.00%	
Total percentage		0.25%	0.00%	0.00%	0.00%	0.25%
Count	prefers employment, only if earns the same	2	17	3	2	24
Column percentage		11.11%	28.81%	2.03%	1.14%	
Row percentage		8.33%	70.83%	12.50%	8.33%	
Total percentage		0.50%	4.25%	0.75%	0.50%	6.00%
Count	would accept employment if earnings were higher	7	20	122	24	173
Column percentage		38.89%	33.90%	82.43%	13.71%	
Row percentage		4.05%	11.56%	70.52%	13.87%	
Total percentage		1.75%	5.00%	30.50%	6.00%	43.25%
Count	would not change, even if earnings were higher	8	22	23	149	202

Column percentage		44.44%	37.29%	15.54%	85.14%	
Row percentage		3.96%	10.89%	11.39%	73.76%	
Total percentage		2.00%	5.50%	5.75%	37.25%	50.50%
Total count	Total	18	59	148	175	400
Total percentage		4.50%	14.75%	37.00%	43.75%	

Source: Own study.

The comparison of the two models implemented shows (Table 5), that in both cases, the micro-F1 score was just the same, but in terms of accuracy the bagging tree model was better (overall accuracy 0.6969 vs. 0.4306).

Table 5. Confusion Matrix (Bagging Tree Model)

CART model				bagging tree model		
	Precision	Recall	F1	Precision	Recall	F1
Class1	0	-	0	1	0.055556	0.105263
Class2	0.333333	0.470588	0.390244	0.708333	0.288136	0.409639
Class3	0.716763	0.712644	0.714697	0.705202	0.824324	0.760125
Class4	0.777228	0.751196	0.76399	0.737624	0.851429	0.790451
MicroF1	0.7225	0.7225	0.7225	0.7225	0.7225	0.7225
Accuracy			0.430625			0.696994
Weighted F1			0.718336			0.752773

Source: Own study.

5. Conclusions

Self-employment is an ambiguous phenomenon considered an essential factor in economic growth. On the other hand, so-called false (fake) self-employment is perceived as tax evasion.

The use of the classification tree method made it possible to identify four categories of self-employed persons. The two methods used (CART and bagging tree models) gave the same overall performance, with a prediction error of 0.2275. However, the bagging tree model showed greater accuracy. Two of the three most important variables in both models were the self-assessment of attitude towards risk and the industry in which the taxpayer operates. In the CART model, the most crucial factor was the individual's level of experience. In the bagging tree model, the most critical factor was the attitude towards risk, and the third most important factor was the respondent's background (parents' employment status).

Identification of the categories of self-employed persons could contribute to shaping tax regulations in a way that would both encourage entrepreneurship and prevent abuse of self-employed status.

6. Research Limitations

The research focused on self-employed individuals providing a specific range of services. Consequently, the results do not include the manufacturing sector, which includes the construction sub-sector. As other research shows, false self-employment is a frequent phenomenon in construction. It would therefore be worthwhile to perform further research targeting self-employed persons in the manufacturing sector.

References:

- Adamczyk, A., Neneman, J., Morawski, L. 2018, Why do we need self-employed persons? Some economic reflections, mainly tax related (Po co nam samozatrudnieni? Refleksje ekonomiczne, głównie podatkowe. Social Science Research Network. <https://papers.ssrn.com/abstract=3393517>.
- Blanchflower, D.G. 2000. Self-employment in OECD countries. *Labour Economics*, 7(5), 471-505. [https://doi.org/10.1016/S0927-5371\(00\)00011-7](https://doi.org/10.1016/S0927-5371(00)00011-7).
- Douglas, E.J., Shepherd, D.A. 2000, Entrepreneurship as a utility maximizing response. *Journal of Business Venturing*, 15(3), 231-251. [https://doi.org/10.1016/S0883-9026\(98\)00008-1](https://doi.org/10.1016/S0883-9026(98)00008-1).
- Douglas, E.J., Shepherd, D.A. 2002. Self-employment as a career choice: Attitudes, entrepreneurial intentions, and utility maximization. *Entrepreneurship Theory and Practice*, 26(3), 81-90.
- Eisenhauer, J.G. 1995. The entrepreneurial decision: economic theory and empirical evidence. *Entrepreneurship Theory and Practice*. <https://doi.org/10.1177/104225879501900405>
- Engblom, S. 2003. Self-employment and the Personal Scope of Labour Law Comparative Lessons from France, Italy, Sweden, the United Kingdom and the United States. <https://data.europa.eu/doi/10.2870/68969>.
- Kahneman, D., Tversky, A. 1979, Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263. <https://doi.org/10.2307/1914185>.
- Nikulin, D. 2020. The determinants of false self-employment: A survey of Polish enterprises. *Eastern European Economics*, 1-12. <https://doi.org/10.1080/00128775.2020.1782754>.
- Startienė, G., Remeikienė, R., Dumčiuvienė, D. 2010. Concept of self-employment. *Economics & Management*, 262-274.
- Thörnquist, A. 2013. False (bogus) self-employment in EastWest labour migration: Recent trends in the Swedish construction and road haulage industries. *TheMES Themes on Migration and Ethnic Studies*, 41(13), 1-44.