## Choosing a Field of Education: Signaling, Mismatch, and Equilibrium Shifting

## Introduction

In the last 30 years, Polish universities have experienced a drastic change in the demand for tertiary education. The number of students in post-communist Poland increased dramatically. The percentage of students within the age group 19-24 has quadrupled in the 20 years from $9.8 \%$ in 1990/91 to a peak of $40.8 \%$ in 2010/2011 and then experienced a small decline to $35.6 \%$ in 2018/19 (see Table 1).

Table 1
Enrolment ratio in higher education

| Enroll. ratio | $1990 / 91$ | $1995 / 96$ | $2000 / 01$ | $2005 / 06$ | $2010 / 11$ | $2015 / 16$ | $2018 / 19$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Gross | 12.9 | 22.3 | 40.7 | 48.9 | 53.8 | 47.6 | 46.2 |
| Net | 9.8 | 17.2 | 30.6 | 38.0 | 40.8 | 37.3 | 35.6 |

Gross: number of people (regardless of age) in tertiary education, as a percentage of the total population aged 19-24.
Net: number of people aged 19-24 in tertiary education as a percentage of the total population aged 19-24.
Source: Statistics Poland (2018).
There is evidence that a growing number of alumni might exceed the capacity of the labor market, resulting in a gap between a high supply of educated workers and little demand for them. Some fields of studies (humanities and social sciences) are claimed to be of little value on the labor market and since the higher education in Poland is publicly funded, the question of the effectiveness of the investment in students may naturally arise.

I employ a well-known model of signaling through education to show how different fields of studies may convey different signals about the workers' productivitv. Then I present a simple numerical exercise on how increasing the cost of

[^0]education may decrease 'overeducation' in the labor market. There are two main contributions of the paper: first, I extend the classic setup to examine a model with a continuum of types and discrete (larger than 2) set of efforts, that is a non-injective mapping from types to signals. Second, I present a small numerical exercise on how the exogenous intervention might decrease inefficiencies in the labor market.

The structure of the paper is as follows. First, I examine a standard model of signaling with a continuous set of agents (of measure 1) and discrete effort choice, which I shall interpret as a choice of field of study - with a possibility of not going to university at all. I assume that there is a fixed cost of going to university and additional disutility from studying a certain field with a certain productivity level. I prove the existence and find the characterization of semi-separating equilibria, in which the agents segregate themselves into subgroups, depending on their productivity level. I also prove that for a given cost of education the equilibrium is uniquely defined through the subset of efforts, chosen by agents.

For a numerical exercise, I shall assume that the demand side of the labour market is mildly imperfect - in particular, there exist a maximal capacity of high-productivity job positions, or, more realistically, that the demand and supply sides on the labor market adapt with some lag. However, this friction is not recognized or internalized by agents in the moment of making educational decisions. As a result, some mismatch arises in the market, which can be decreased only by changing the agents' incentives through policies. Given the data on recent alumni's field of study and professional career, I calibrate a stylized disutility function that would rationalize the choices within a signaling model. I also examine what is the range of overeducation on the labor market suggested by the data, in terms of the ratio of employers whose jobs require less schooling than they actually obtained. The calibration is followed by a sensitivity analysis - varying the cost of education, I verify whether a better match on the labor market could be obtained. In particular, I check whether the issue of overeducation could be reduced through the change in the fixed cost of education. The paper is concluded with a short discussion on the assumptions and implications of the exercise.

## 1. Literature review

### 1.1. Signaling

The standard signaling model of education dates back to the seminal article by Spence (1973), where the agents' types are discrete (high or low). A fully continuous case is considered in Mailath (1987) and Escriche et al. (2002), where important results considering differentiability and existence of equilibria are set. I investigate a model with a discrete set of signals (with $|S|>2$ ) and a continuum of types, which is a simple, yet novel, ${ }^{1}$ extension. In such a case, the standard

[^1]conditions from Mailath (1987); cannot be properly defined, however, following an idea from Crawford and Sobel (1982), I prove the equilibrium existence under mild conditions. Naturally, since the type space is larger than the signal space, only semi-pooling (and pooling) equilibria exist.

Signaling through education has been also tested, to some extent, empirically. While it is difficult to disentangle the effect of signaling effect vs. human capital accumulation, there were some attempts to do it. Ehrmantraut et al. (2020) find the signal of a Bachelor diploma to increase wages by $30 \%$. Kaymak (2012) estimates the relative value of the signal vs. human capital to be $23 \%$, while the popular book Caplan (2018) claims it to be as high as $80 \%$. Tyler et al. (2000) provide an interesting identification strategy to show that the signal of GED diploma increases the earnings of US high-school dropouts by $10-19 \%$, while Clark and Martorell (2014) find, on the other hand, little evidence of high school diploma signaling effect. While it is quite clear, that education is both human capital formation and signaling, in my model, I abstract from the former to focus simply on the latter.

### 1.2. Overeducation

Overeducation is defined as an excess of a worker's attained or completed level of schooling, comparing to the level of schooling required for the job the worker holds (Leuven and Oosterbeek 2011; Sicherman 1991; Duncan and Hoffman 1982). This general definition could be made more precise, by specifying, how the educational requirements for the job are analyzed and measured. In particular, as is pointed out in Groot and Maasen van den Brink (2000), overeducation could be measured either by respondent's subjective opinion, expressed in a questionnaire or by objective measures, referring either to requirements expressed in job offers or to an average level of education within the occupation of the worker. The latter approach is represented by e.g. Sicherman (1991), where a large sample from PSID data is used to measure overeducation in a relation to the level of schooling of other workers employed in the same job. The subjective approach is used e.g. in Alba-Ramirez (1993), for an analysis of ECTV Spanish data, or Linsley (2005) for statistics arising from the Australian NLC survey and in Meroni and Vera-Toscano (2017) for cross-country comparison within a few European countries. Due to the character of the dataset, I follow their path, and examine potential overeducation in worker's own evaluation of the appropriateness of their schooling, as reported in the 'Higher Education as a Generator of Strategic Competences' (HEGESCO) project. The empirical studies (e.g. Alba-Ramirez 1993; Duncan and Hoffman 1982; Linsley 2005) typically use Mincer human capital equation.

An important assumption needs to be made when referring to overeducation in the signaling model. Notice that in the classic signaling model, overschooling understood as attaining more education than strictly required to perform the tasks of one's jobs (Linsley (2005)) arises by assumption, since the only role of education is to signal one's productivity. Therefore, I propose another interpretation
of overschooling: I assume that if the agents send a signal by completing a given level of schooling, they would perceive themselves as overeducated if and only if, their job position does not compensate them for their cost of signaling, i.e. they consider their choice of education as ex-post too high. Such a mismatch may arise if the demand side of the labour market is not perfectly elastic, i.e. only some of the workers are correctly matched to the job position they are suitable for.

## 2. Model

### 2.1. Setup

## Incentives

Consider a measure one of agents with secondary education that have to decide about their future career. Agents differ in productivity - i.e. their type - which is private information of each agent. Productivity could be signaled through tertiary education: getting a Bachelor's diploma entails a fixed cost of studying, and provides some disutility to the student, who has to exercise his mind. However, a diploma serves as a signal on the job market, and, therefore, studying can result in a higher wage in the future. I shall assume that the tertiary education decision is a choice among a finite set of effort levels, which I shall interpret as fields of study. The students signal their productivity not only through the mere fact of obtaining a diploma, but also through the type of diploma, that would differentiate between more and less productive students. The fields of study could be therefore divided into those that are 'more pleasant' or 'less demanding' (i.e. provide less disutility), but result in a lower future wage and fields that provide more disutility, but guarantee higher wage.

Going to the university entails a fixed cost $c$ that represents a constant part of the disutility function and may be interpreted as the tuition fees and other costs associated with getting higher education - the cost of moving to another city, renting a flat etc. This cost is borne by every student, regardless of the field of study. Notice, that if some part of $c$ is tuition fees, then $c$ is subject to policy changes.

I shall assume that different fields of study require different efforts $e \in \mathbb{R}$. I abstract from individual's talents and preferences, assuming that every high-school graduate could in principle study any field - from humanities to medicine and engineering - and the choice would entail a field-specific effort that would signal the player's 'productivity' ${ }^{2} \theta \in[b, B]$ and result in field-specific wage.

[^2]Specifically, I shall consider a signal space $E$ with effort levels $0=\bar{e}_{0}<\bar{e}_{1}<\ldots<\bar{e}_{S}$. The agent's problem is

$$
\max _{e \in E}\left\{w(e)-v(e, \theta)-c \cdot 1_{\left\{e \neq \bar{e}_{0}\right\}}\right\},
$$

where $w$ is the wage function, $v(e, \theta)$ is the agent's disutility from effort $e$ given productivity $\theta$ and $c$ is a fixed component of the cost.

Notice that the wages offered by the firms $w(e)$ take into account the expected productivity of the agent, and therefore, are the function of his signal. Agent anticipates this process and knows the distribution of types in the whole population.

The assumptions about the disutility function are standard:
Assumption 1. Let $v:[0, \infty) \times[b, B] \rightarrow \mathbb{R}$ and satisfy: ${ }^{3}$
(a) $v(0, \theta)=0$,
(b) $v_{e}(e, \theta)>0$ and $v_{\theta}(e, \theta)<0$,
(c) $v_{e}^{2}(e, \theta) \geq 0$ and $v_{e \theta}(e, \theta)<0$,
(d) $\lim _{\theta \rightarrow B^{-}} v(e, \theta)=0$ and $\lim _{\theta \rightarrow b^{+}} v(e, \theta)=\infty$.

I claim that in the equilibrium, lower types chose lower effort levels and receive lower wages. Therefore, I will make a correspondence between efforts and wages and conclude that the order of the wages represents the order of efforts.

## Firms and the regulator

As in standard signaling exercises, I assume there are at least two firms, to make the market competition explicit. The competition drives the firms' profit to zero, therefore forcing the firms to offer wage contracts that cover the expected productivity of hired workers:

$$
w\left(\bar{e}_{i}\right)=E\left(\theta \mid \theta \in \Theta_{\bar{e}_{i}}\right)
$$

where $\Theta_{\bar{e}_{i}}=\left\{\theta: \bar{e}_{i}=\operatorname{argmax}_{e} w(e)-v(\theta, e)\right\}$ is the anticipated set of types, who chose a signal $\bar{e}_{i}$.

In the theoretical part of this article, no further assumption is needed to show the existence of equilibrium with different effort levels. For the numerical exercise, I shall employ an additional assumption to rationalize the observed mismatch on the market. More specifically, I assume that the demand is partially inelastic. The firms cannot offer an unlimited number of contracts for any number of productive workers they face. It might be due to the costs associated with creating new job positions, or by the fact, that firms require a certain ratio of productive vs. unproductive workers. In particular, for any class of jobs that require at least a given level of productivity $\bar{\theta}$, there could potentially be created at most $a(\bar{\theta})$ job positions available for new workers. If the number of workers with productivity signaled to be at least $\bar{\theta}$ is less $a(\bar{\theta})$, then all the workers are employed in jobs appropriate to their signal. If, however, the number of productive workers

[^3]exceeds $a(\bar{\theta})$, only $a(\bar{\theta})$ randomly chosen workers are employed, while the rest must search for jobs in a "lower" class, i.e. job positions that would also accept workers with a lower signal.

However, the limited elasticity of the firms is not observed (or not internalized) by the future workers, deciding about their level of education. The students therefore do not have rational beliefs about $a(\bar{\theta})$. It might correspond to the assumption of a lagged adaptation of the market. ${ }^{4}$ It seems reasonable to assume that the difference between the supply and demand of productive workers is small, and therefore, might not be perceived as a real threat to the market participants.

I shall assume, on the other hand, that $a(\bar{\theta})$ is observed, or might be forecasted, by the better-informed regulator, who decides about the cost of study. Manipulations with the cost can generally lead to oversupply or undersupply of workers, in reference to the market satiation level $a(\bar{\theta})$.

## Distribution of types

Consider a distribution of types that is atomless with full-support and a resulting cumulative distribution function $F$ (that is strictly increasing). To simplify the analysis, I will 'normalize' $F$ to a uniform distribution, by using the following fact:

Fact 1. Let $F$ be a continuous cumulative distribution function and $U$ be a uniform random variable on $[\tilde{b}, \tilde{B}]$. Then $X=F^{-1}\left(\frac{1}{\tilde{B}-\tilde{b}} U\right)$ follows a distribution $F$.

Corollary 1. Let $\theta$ follow any continuous distribution function $F$ and $v(\theta, e)$ be a disutility function satisfying Assumption 1 . Then there exist $\theta \sim U n i f[\tilde{b}, \tilde{B}]$ and disutility function $\tilde{v}(\tilde{\theta}, e)$, such that $\tilde{v}$ also satisfies Assumption 1.

Let $\tilde{v}(\tilde{\theta}, e)=v\left(F^{-1}(\theta /(\tilde{B}-\tilde{b})), e\right)$. Since $F$ is increasing, also $F^{-1}$ is increasing and it is easy to verify that properties (a)-(d) from Assumption 1 hold.

Notice that in my exercise, the distribution $F$ is only needed to estimate $v$. By Corollary 1, I can assume without loss of generality that $F$ is uniform and estimate the disutility function corresponding to the uniform distribution. To avoid excessive notation, I shall denote from this point the desired disutility function as $v(\theta, e)$, remembering that now $\theta \sim U n i f[b, B]$.

### 2.2. Equilibrium

## Semi-pooling equilibria

I shall look at a specific class of equilibria in the signaling model, namely symmetric equilibria, in which all the agents choose ex-ante (i.e. before learning their type) the same strategy, that defines their effort, given their type.

[^4]Definition 1. A symmetric equilibrium for a signaling game is a function $\tilde{e}: ~ \Theta \rightarrow S$, such that for an agent of type $\theta$, function $\tilde{e}(\theta)$ is the best response to all the other players choosing according to $\tilde{e}(\theta)$. Specifically, $\tilde{e}(\theta)$ satisfies:

$$
\tilde{e}(\theta)=\max _{e \in S}\left\{w(e \mid \text { other players follow } \tilde{e}(\theta))-v(\theta, e)-c \cdot 1_{\left\{e \neq \bar{e}_{0}\right\}}\right\} .
$$

Notice that since a single agent is of measure zero, by changing unilaterally her decision, she does not affect other player's payoff function.

Therefore, the equilibrium is simply defined through an equation:

$$
\tilde{e}(\theta)=\max _{e(\theta)} w(e(\theta))-v(\theta, e(\theta))-c \cdot 1_{\left\{e(\theta) \neq \bar{e}_{0}\right\}} .
$$

It is clear that the function $\tilde{e}: \Theta \rightarrow S$ cannot be $1-1$, therefore no separating equilibria exist. Moreover, the equilibrium function $\tilde{e}(\theta)$ is locally constant and weakly increasing in $\theta .{ }^{5}$ By the fact that $\tilde{e}(\theta)$ is an increasing step function, all the equilibria must be partition equilibria, in a sense that the preimages of subsequent effort levels define (possibly degenerate) intervals that sum up to the type space $[b, B]$.

This intuition is formalized in the two definitions:
Definition 2. Let $\Theta \ni \theta$ be a space of types and $\delta=\left\{\bar{e}_{0}, \ldots, \bar{e}_{S}\right\}$ be the effort space. In this environment, a (generalized) semi-pooling equilibrium of size $M+1$ is defined by a subset of efforts $\mathcal{M}=\left\{\bar{e}_{i_{0}}, \ldots, \bar{e}_{i_{M}}\right\} \subset \delta$ and a sequence of thresholds $0=\bar{\theta}_{0}<\bar{\theta}_{1}<\ldots<\bar{\theta}_{M+1}=B$ such that all agents of type $\theta \in\left[\bar{\theta}_{m}, \bar{\theta}_{m+1}\right]$ choose $e=\bar{e}_{i_{m}}$ for $m=0, \ldots, M$. The sequence $\left\{\bar{\theta}_{i}\right\}_{i}$ would be called an equilibrium partition for a subset $\mathcal{M}$.

In particular, if $\tilde{e}(\theta)$ is an equilibrium in a sense of Definition 1, then $\mathcal{M}=\tilde{e}^{-1}(\delta)$ and $\left[\bar{\theta}_{m}, \bar{\theta}_{m+1}\right]=\tilde{e}^{-1}\left(\bar{e}_{i_{m}}\right)$.

The semi-pooling equilibrium of size 1 is, by definition, a pooling equilibrium. However, to avoid confusion in my further results, I will keep the generalized notion of semi-pooling equilibria throughout the paper, keeping in mind that those include also a pure pooling outcome.

Thus we can write:
Corollary 2. In a signaling game with a continuum of types and discrete signals, the only symmetric equilibria are (generalized) semi-pooling equilibria.

We will now examine some properties of the equilibrium. Let us recall that, following Corollary $1, \mathrm{I}$ assumed that $\theta \sim \operatorname{Unif}[b, B]$.
Lemma 1. For effort levels $\bar{e}_{0}, \ldots, \bar{e}_{M}$, and $\theta$ following a uniform distribution on $[b, B]$, the semi-pooling equilibrium wages satisfy:

$$
w\left(\bar{e}_{i}\right)=\frac{\bar{\theta}_{i+1}+\bar{\theta}_{i}}{2} \text { for } i=0,1, \ldots, S
$$

[^5]Proof (simple): By an assumption that firms make no profits, in a semi-pooling equilibrium with thresholds $0=\bar{\theta}_{0}<\bar{\theta}_{1}<\ldots<\bar{\theta}_{S+1}=B$ we have $w\left(\bar{e}_{i}\right)=$ $=\left.E \theta\right|_{\left(\bar{\theta}_{i}, \bar{\theta}_{i+1}\right)}=\left(\bar{\theta}_{i}+\bar{\theta}_{i+1}\right) / 2$.
Corollary 3. Given market wages $w_{1}, \ldots, w_{S}$, the semi-pooling equilibrium is a set of thresholds defined sequentially:

$$
\begin{aligned}
& \bar{\theta}_{0}=b \text { and } \bar{\theta}_{S}=B \\
& \bar{\theta}_{i}=2 w_{i-1}-\bar{\theta}_{i-1} \text { for } i=1, \ldots, S
\end{aligned}
$$

## Binary signal

Let us first consider a basic model, where there are only two possible signals: getting a higher education diploma (denoted by $\bar{e}_{1}=1$ ) or not (denoted by $\bar{e}_{0}=0$ ). I argue that in such a model there exists a threshold of type $\bar{\theta}$, such that agents with $\theta>\bar{\theta}$ choose to study and those with $\theta<\bar{\theta}$ prefer to go to the job market with a secondary education diploma.
Proposition1. For a distribution of types $\theta:[0,1] \sim U[b, B]$ and cost of education $c$, such that $c \leq \mathrm{B}-\mathbb{E} \theta$, there exists a semi-pooling equilibrium in which all agents of type $\theta>\bar{\theta}$ choose $e=1$ and all agents of type $\theta<\bar{\theta}$ choose $e=0$ for some threshold $\bar{\theta}$.

Observe that a set of symmetric strategies $\left\{\left(e_{s}=1 \mid \theta>\bar{\theta}\right),\left(e_{s}=0 \mid \theta<\bar{\theta}\right)\right\}_{s}$ is a Nash equilibrium, if it satisfies:
a) Individual rationality:

$$
\begin{gather*}
w(1)-v(1, \theta)-c \geq 0 \text { for } \theta>\bar{\theta}  \tag{1}\\
w(0)-v(0, \theta) \geq 0 \text { for } \theta<\bar{\theta} \tag{2}
\end{gather*}
$$

b) Incentive compatibility:

$$
\begin{align*}
& w(1)-v(1, \theta)-c \geq w(0)-v(0, \theta) \text { for } \theta>\bar{\theta}  \tag{3}\\
& w(0)-v(0, \theta) \geq w(1)-v(1, \theta)-c \text { for } \theta<\bar{\theta} \tag{4}
\end{align*}
$$

Notice first that condition (2) is trivially satisfied by the assumption that $v(0, \theta)=0$ and $\theta>0$, so only (1) needs to be verified.

The incentive compatibility conditions could be rewritten as:

$$
\begin{aligned}
& w(1)-w(0) \geq v(1, \theta)-v(0, \theta)+c \text { for } \theta>\bar{\theta} \\
& w(1)-w(0) \leq v(1, \theta)-v(0, \theta)+c \text { for } \theta<\bar{\theta}
\end{aligned}
$$

Function $v(1, \theta)-v(0, \theta)$ is continuous in $\theta$, so by taking $\theta \rightarrow \bar{\theta}$ we infer that it must be:

$$
\begin{equation*}
w(1)-w(0)=v(1, \bar{\theta})+c \tag{5}
\end{equation*}
$$

To finish the proof of existence, it remains to verify that for a given $c$ and $v$ there exists $\bar{\theta}$ such that the equation (5) holds. Observe that the LHS for a uniform distribution is:

$$
w(1)-w(0)=\frac{B+\bar{\theta}}{2}-\frac{\bar{\theta}+b}{2}=\frac{B-b}{2},
$$

and the RHS of (5) is decreasing in $\bar{\theta}$. By Inada conditions (Assumption 1.d), there exists a solution to (5). Moreover, since the formula is decreasing in $\theta$, the solution is unique.

## Finite space of signals

Now, we can move to the general case of a finite number of possible signals. Assume the signal space is a discrete set of size $S+1$ with effort levels $0=\bar{e}_{0}<\bar{e}_{1}<$ $<\ldots<\bar{e}_{S}$ for $S>1$. I shall interpret $\bar{e}_{0}$ as a decision not to study, and $\bar{e}_{1}, \ldots, \bar{e}_{S}$ as a choice to study in one of $S$ fields.
Theorem 1. For a given distribution of types $\theta:[b, B] \rightarrow \mathbb{R}$ and cost of education $c$ there exists some number $\bar{M}$ such that for any $1 \leq M \leq \bar{M}$ there exists a semi-pooling equilibrium of size $M$. In particular, an equilibrium partition for an equilibrium of size $M$ could be chosen for a subset of $M$ minimal efforts, i.e. $\left\{\bar{e}_{0}, \bar{e}_{1}, \ldots, \bar{e}_{M}\right\}$.

Proof. The argument follows a similar argument to the one in Crawford and Sobel (1982), adapted to my setup. For details, see Appendix.

From the proof, we observe that a semi-pooling equilibrium with efforts $\left\{\bar{e}_{i_{0}}<\ldots<\bar{e}_{i_{M}}\right\}$ is a sequence $b=\bar{\theta}_{0}<\bar{\theta}_{1}<\ldots<\bar{\theta}_{M+1}=B$ that solves a system of equations:

$$
\begin{equation*}
\frac{\bar{\theta}_{m+1}-\bar{\theta}_{m-1}}{2}=v\left(\bar{e}_{i_{m}}, \bar{\theta}_{m}\right)-v\left(\bar{e}_{i_{m-1}}, \bar{\theta}_{m}\right)-c \cdot 1_{\left\{i_{m-1}=0\right\}} \text { for } m=1, \ldots, M \tag{6}
\end{equation*}
$$

with initial conditions:

$$
\bar{\theta}_{0}=b, \bar{\theta}_{M+1}=B .
$$

From this observation, we conclude that given a set of efforts, the equilibrium is unique.
Proposition 2. For a given subset of efforts $\delta=\left\{\bar{e}_{i_{0}}, \ldots, \bar{e}_{i_{M}}\right\}$, there exists at most one equilibrium in which every effort in $\delta$ is chosen by a positive measure of agents.
Proof: by contradiction - see Appendix.
Proposition 2 enables us to determine an equilibrium partition, defined uniquely for a given (i.e. observed in the data) set of efforts chosen by agents.

## 3. Numerical exercise

### 3.1. Application of the theory

In this section I provide a simple numerical exercise to examine how the scope of overeducation could be reduced by increasing the cost of studying, thus linking the theory from Section 3 to the data.

First, I examine the data and find an equilibrium that corresponds to the observed choices of efforts. In the data we can observe 8 different fields of studies
that could be interpreted as effort levels and denoted from now on as $\left\{\bar{e}_{1}, \ldots, \bar{e}_{8}\right\}$. Additionally, we shall consider a choice of not going to studies at all, denoted by $\bar{e}_{0}$. By Proposition 2 for a given set of 9 "non-redundant" effort levels and some underlying $\operatorname{cost} c$, the equilibrium is unique.

Next, I choose a simple functional form of the disutility function and calibrate the model with the data provided. In particular, I calculate the cost of education that rationalizes the choices. Additionally, I examine the scope of overeducation that corresponds to the assumption of non-perfectly flexible labour market.

Finally, I examine, whether by varying the cost of education (e.g. through tuition fees) the scope of overeducation could be reduced and what cost would support the equilibrium that minimizes the mismatch.

### 3.2. Data

The data comes from a survey of higher education graduates within a "Higher education as a generator of strategic competencies" (HEGESCO) project, conducted by the Maastricht University's Research Center for Education and the Labor Market. The data was gathered for five East European countries; however, in the model I used only data for Poland.

The initial sample consisted of 1200 observations of university alumni (MA and BA, with a majority of the former), who were questioned about their career history and current position five years after their graduation in 2002/2003. To avoid possible bias I excluded graduates, who chose more than one study program. On the labor side, I also excluded self-employed entrepreneurs, as it is hard to verify whether they are a part of exogenous market demand for educated workers, or rather create the demand by themselves. I also excluded some observations with a significant number of missing values. The final sample included therefore 799 observations on the choice of study and the following careers on the job market.

The variables of greatest interest are summarized in Table 2.
Table 2
The variables used in the analysis

| Name | Variable | Scope |
| :--- | :--- | :---: |
| a1foe1 | Field of study (general) | Code of field in $\{1,2, \ldots, 8\}$ |
| e2empsgm | Months of employment since graduation | $0-60$ |
| f7inctex | Gross monthly earnings truncated: $a$ <br> total (euro) in a current job | $1-99999$ |
| f8rlevs | What is the appropriate level of education for <br> your job position? | Code of answer in $\{1,2,3,4\}$ |
| f9apfos | What is the appropriate field of study for your <br> job position? | Code of answer in $\{1,2,3,4\}$ |

${ }^{a}$ The top $2.5 \%$ and lowest $0.5 \%$ earnings were excluded.
Source: own elaboration.

### 3.3. Preliminary observations

First, I verify the anecdotal evidence of an excessive supply of university graduates in the job market. Figure 1 shows a percentage of time spent in paid employment within the 5 years from graduation. It could be observed, that engineering graduates, as well as alumni of science and mathematics, have worked, on average, the longest, while the agents who chose humanities or services are either more often unemployed or take a longer time to find a job. ${ }^{6}$ While the vast majority of the

Figure 1
Employment time (\%) within the $\mathbf{5}$ years from graduation


Source: HEGESCO data, own calculations.
Figure 2
Total monthly earnings (in EUR) $\mathbf{5}$ years after graduation


Source: HEGESCO data, own calculations.

[^6]alumni are employed (even if not in the desired job), the differences in non-employment rate between the fields are substantial. This picture partly supports the common belief that some diplomas serve poorly as an asset on the job market.

Figure 2 presents the total monthly earnings of university graduates, 5 years after the completion of studies. We can observe that the wages differ significantly between employers of different education fields, with the maximum attained by alumni of engineering and minimum for education graduates. Those differences shall be used as an indicator that different fields of studies entail different efforts.

### 3.4. Calibration procedure

The calibration procedure was divided into two parts: gathering data on equilibrium outcomes and calibrating the value function.

To find the thresholds, defining a semi-pooling equilibrium in the job market, I calculated the average total monthly earnings in every field of education and multiplied them by average employment in the last 5 years, to get a measure of "expected" income of an agent. In our simple exercise, I abstract from the fact that wages change over time, and take wages as they are 5 years of graduation, just for the sake of finding a linear order. Then, the thresholds are calculated, in accordance with Corollary 3 (see Table 3).

Table 3
Market wages (in EUR) and corresponding thresholds

| Variable | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $w\left(\bar{e}_{i}\right)$ | $512.6^{a}$ | 605.3 | 714.2 | 778.8 | 829.6 | 895.6 | 985.6 | 1014.7 | 1172.6 | - |
| $\bar{\theta}_{i}$ | 512.5 | 512.6 | 697.9 | 730.5 | 827.1 | 832.2 | 959.0 | 1012.2 | 1017.1 | 1328.1 |

${ }^{a} w(0)$ was not taken from data, but calibrated to guarantee existence of a separating equilibrium. Since the data covers only educated workers, it does not provide information about the wage of an employee with just a secondary education. I assume that all agents with secondary education are on a small interval around $b$ (here $b=512,64$ ).
Source: own calculations.
I assume that the disutility function satisfies Assumption 1. For simplicity, I shall only consider functions of the form $v(e, \theta)=f(e) g(\theta)$. Function $f(e)$ is convex, therefore I shall approximate it by a quadratic function of the form $a e^{2}+b e+d$. Since $f(e)=0$, then $d=0$ and I shall only estimate two parameters of function $f$, that I denote now by $f^{(a, b)}$. I shall assume that $g$ function is of the form $g(\theta)=\theta^{-q}$ for $q>0$. Such a function (denoted by $g^{(q)}$ ) satisfies Inada condition in 0 and gives us a third parameter to estimate. The fourth parameter is the cost of education $c$.

The difficulty with estimating $v$ arises from the fact, that we have few observations and properties of that function. However, in the equilibrium, function $v$ must satisfy (6), which could be rewritten as:

$$
\begin{equation*}
v\left(\bar{e}_{m}, \bar{\theta}_{m}\right)-v\left(\bar{e}_{m-1}, \bar{\theta}_{m}\right)-c \cdot 1_{\{m=0\}}=w\left(\bar{e}_{m}\right)-w\left(\bar{e}_{m-1}\right) \text { for } m=1, \ldots, M . \tag{7}
\end{equation*}
$$

Therefore, I propose the following procedure, based on GMM: define $\operatorname{error}(a, b, q, c)$ as the vector of differences between RHS and LHS of (7) for a function $v$ with parameters $(a, b, q)$ and cost $c$; use GMM with restrictions to choose a set of $(a, b, q, c)$ such that the sum of squared elements of error $(a, b, q, c)$ would be minimal and all the restrictions given by individual rationality and incentive compatibility (see the Appendix) are satisfied.

Thus, calibration is solving an optimization problem:

$$
\begin{gathered}
\min _{a, b, q, c}\left\{\sum_{m=1}^{M}\left[f^{(a, b)}\left(\bar{e}_{m}\right) \cdot g^{(q)}\left(\bar{\theta}_{m}\right)-f^{(a, b)}\left(\bar{e}_{m-1}\right) \cdot g^{(q)}\left(\bar{\theta}_{m}\right)-c \cdot 1_{\{m=1\}}-w\left(\bar{e}_{m}\right)+w\left(\bar{e}_{m-1}\right)\right]^{2}\right\}, \\
\text { subject to: }\left\{w\left(\bar{e}_{m}\right)-v\left(\bar{e}_{m}, \bar{\theta}_{m}\right)-c \cdot 1_{\{m \neq 0\}} \geq 0\right\} \text { for } m=1, \ldots, M
\end{gathered}
$$

Results of the calibration procedure are summarized in the Table 4.
Table 4
Coeffficients of from GMM with constraints

| $a$ | $b$ | $q$ | $c$ |
| :---: | :---: | :---: | :---: |
| 0.013 | 180.8 | -0.032 | 55.5 |

Source: own calculations.

### 3.5. Increasing the cost

I shall now examine, how increasing the cost of education may dampen the excess supply of educated workers. First, I shall verify what was the scope of the overeducation in 2003. I define the scope of overeducation as the ratio of respondents that perceive their education as too high for their job position. I will call the workers strongly mismatched, if they answered a question "What is the appropriate level of education for your job position?" with answers specifying an education level lower than their actual level. I will call the workers to be weakly mismatched if they are strongly mismatched or answered a question "What is the appropriate field of study for your job position?" with an answer "completely different field" or "no particular field".

In Figure 3 I plotted the scope of subjective overeducation. Bottom bars indicate workers, who perceive their education level as too high for their job position, i.e. strongly mismatched workers. Upper bars describe the ratio of agents, who think their high education level is adequate for their job, but the field of their studies is not. The sum of both bars describes the set of weakly mismatched workers.

To examine, how the cost of education could possibly reduce the mismatch, I follow a reverse algorithm to the one previously used. First, I calculate the length of the subsequent intervals $\left(\bar{\theta}_{i}, \bar{\theta}_{i+1}\right)$ for the mismatched market. Then, assuming the mismatch arises from a non-perfectly flexible supply of job positions,

Figure 3

## Mismatch on the labor market: \% of employees who perceive their education level as inadequate for their job



Source: HEGESCO data, own calculations.

I estimate the market capacity - given the data, I calculate, how much each interval should be truncated to fit to the market demand, e.g. if there are $10 \%$ of mismatched workers in a given interval, it means, that the 'optimal' interval (i.e. an interval that would bring supply to $a(\bar{\theta})$ ) should be $10 \%$ shorter. Then I find a cost that would result in a pooling equilibrium with interval sizes as close to 'optimal' as possible.

In the first version, I try to diminish only the ratio of strongly mismatched workers, assuming that weak mismatch could be viewed as a random disturbance. In the second version, I assume that a weak mismatch is always an indication of choosing effort 'too high' for a given job. It represents the assumption that a worker can get a job for which she is too educated, but not a one, for which she does not have high enough signal $e$.

The optimal cost that reduces the mismatch is $c$ that minimizes the sum of differences between the desired and achieved interval lengths:

$$
\min _{c}\left\{\sum_{m=1}^{M}\left[\left(\bar{\theta}_{m}(c)-\bar{\theta}_{m-1}(c)\right)-d_{m-1}^{m}\right]^{2}\right\},
$$

where $d_{m-1}^{m}$ is the length of the interval $\left[\bar{\theta}_{m-1}, \bar{\theta}_{m}\right]$ that corresponds to the values in Table 3, corrected for overeducation ratio and $\bar{\theta}_{i}(c)$ are solutions to the optimization problem for $(a, b, q)$ specified during the calibration. The optimal cost
that minimizes the sum of squared differences between the optimal and achieved partition is summarized in Table 5.

Table 5
Costs that support different equilibria

|  | Equilibrium computed <br> from the real data | Equilibrium minimizing <br> strong mismatch | Equilibrium minimizing <br> weak mismatch |
| :---: | :---: | :---: | :---: |
| $c$ | 55.5 | 62 | 62 |

Source: own calculations.
Indeed, increasing the cost of education reduces mismatch. It might be surprising that the optimal cost is the same for weak and strong mismatch. However, in this simple numerical illustration, this could be a statistical fluke.

### 3.6. Limitation

The numerical example serves as a simple illustration of shifting the equilibrium in a signaling game. Its intention is by no means to estimate the effects of 're-al-life' policy changes.

First of all, the exercise is performed under very specific assumptions. In particular, education is reduced to pure signal and the agents are not forward-looking in a standard economic sense.

Second, the data I gathered served as an example, rather than actual calculations of return of education. I take an average income level five years after graduation, not taking into account the career perspectives and the expected path of income development. In particular, some fields of study (e.g. medicine) may result in jobs that are characterized by low income for inexperienced workers, and a further sharp increase with a certain level of experience. Moreover, the sample might not be representative of the whole population of Poland.

Third, the assumption of the form of the disutility function, that helped us to reduce the identification problem to a few parameters, might be inadequate. In principle, finding a general function to fit the data, given only a set of (a few) equality and inequality constraints is an impossible task.

The numerical exercise is therefore a simple illustration of a social planner's potential to influence the decentralized signaling equilibrium arising in the labor market.

## Conclusion

In this paper I have reviewed a well-known model of job market signaling through education to provide a simple framework for choosing a field of study. In the second part, I used the model to illustrate a possible reduction in overeducation in the labour market. I argued that overeducation might arise if the cost of getting a diploma is too small to give the right incentives to the agents. As a result, too
many agents would choose to pursue a MA or BA diploma and the market demand for educated workers cannot satisfy an excess supply of graduates.

In the theoretical part, a signaling game with continuum types of agents and discrete space of efforts was analyzed. I argued that since the set of types is bigger than the set of signals, the standard results of Spence (1973) or Mailath (1987) cannot be easily applied. I provided proof of existence and characterization of a semi-pooling equilibrium in this setup. Moreover, I showed that the equilibrium is uniquely defined through the efforts chosen by agents.

Next, a simple numerical exercise was executed. Under the assumption that the disutility function belongs to a specific parametric family, I used GMM to adjust the model to the real-world data. I used such a calibrated model to examine whether a change in the cost of study could diminish the issue of overeducation in the job market. The results are consistent with the intuition arising from the model - indeed, increasing the cost of education reduces agents' incentives to continue their education on the tertiary level and can therefore reduce the excess supply of educated professionals.
Received: 4 February 2021
(revised version: 17 December 2021)

## References

Alba-Ramírez A. (1993), Mismatch in the Spanish Labor Market: Overeducation?, "The Journal of Human Resources", no. 28(2), pp. 259-278.
Caplan B. (2018), The Case Against Education: Why the Education System Is a Waste of Time and Money, Princeton University Press, Princeton.
Clark D., Martorell P. (2014), The Signaling Value of a High School Diploma, "Journal of Political Economy", no. 122(2), pp. 282-318.
Crawford V.P., Sobel J. (1982), Strategic Information Transmission, "Econometrica", no. 50(6), pp. 1431-1451.
Duncan G.J., Hoffman S.D. (1982), The Incidence and Wage Effects of Overeducation, "Economics of Education Review", no. 1, pp. 75-86.
Ehrmantraut L., Pinger P., Stans R. (2020), The Expected (Signaling) Value of Higher Education, IZA Discussion Paper No. 13729, https://ssrn.com/abstract=3699846.
Escriche L., Gonzalez O.G., Sánchez R. (2002), Education and Job Competition in a Signalling Game, Technical Report, University of Valencia, http://centros.uv.es/web/departamentos/D10/data/investigacion/PDF178.pdf
Groot W., Maassen van den Brink H. (2000), Overeducation in the Labor Market: a Metaanalysis, "Economics of Education Review", no. 19(2), pp. 149-158.
HEGESCO (2008), Higher Education as a Generator of Strategic Competences [data], Research Center for Education and the Labour Market, Maastricht University. Published in DANS database: https://easy.dans.knaw.nl
Kaymak B. (2012), Quantifying the Signaling Role of Education, „Society of Labor Economists", no. 40.
Leuven E., Oosterbeek H. (2011), Overeducation and Mismatch in the Labor Market, IZA Discussion Paper No. 5523, http://ssrn.com/abstract=1771245.

Linsley I. (2005), Causes of Overeducation in the Australian Labour Market, Department of Economics - Working Papers Series 940, University of Melbourne.
Ma C.A., Weiss A.M. (1993), A Signaling Theory of Unemployment, "European Economic Review", no. 37, pp. 135-157.
Mailath G.J. (1987), Incentive Compatibility in Signaling Games with a Continuum of Types, "Econometrica", no. 55(6), pp. 1349-1365.
Meroni E.C., Vera-Toscano E. (2017), The Persistence of Overeducation Among Recent Graduates, "Labour Economics" no. 48, pp. 120-143.
Sicherman N. (1991), "Overeducation" in the Labor Market, "Journal of Labor Economics", no. 9(2), pp. 101-122.
Spence M. (1973), Job Market Signaling, "The Quarterly Journal of Economics", no. 87(3), pp. 355-374.
Statistics Poland (2010, 2015, 2018), Higher Education Institutions and their Finances [annual report], CSO, Warsaw.
Tyler J.H., Murnane R.J., Willett J.B. (2000), Estimating the Labor Market Signaling Value of the GED, "The Quarterly Journal of Economics", vol. 115, pp .431-468.

## Appendix: Proofs

I shall prove the existence of an equilibrium partition (and therefore, a semi-pooling equilibrium) for a subset $\left\{\bar{e}_{0}, \ldots, \bar{e}_{M}\right\}$ for some $M>1$. Observe that a set of symmetric strategies $\left\{\left(e_{s}=\bar{e}_{m} \mid \theta \in\left[\bar{\theta}_{m}, \bar{\theta}_{m+1}\right]\right)\right\}_{s}$ is a Nash equilibrium, if it satisfies:
(a) individual rationality:

$$
w\left(\bar{e}_{m}\right)-v\left(\bar{e}_{m}, \theta\right)-c \cdot 1_{\{m \neq 0\}} \geq 0 \text { for } \theta \in\left[\bar{\theta}_{m}, \bar{\theta}_{m+1}\right]
$$

(b) incentive compatibility:

$$
w\left(\bar{e}_{m}\right)-v\left(\bar{e}_{m}, \theta\right)-c \cdot 1_{\{m \neq 0\}} \geq w\left(\bar{e}_{j}\right)-v\left(\bar{e}_{j}, \theta\right)-c \cdot 1_{\{j \neq 0\}} \forall_{j \neq m} \text { for } \theta \in\left[\bar{\theta}_{m}, \bar{\theta}_{m+1}\right] .
$$

The constraints are binding only for neighboring subintervals, i.e. a sufficient set of conditions is just:

$$
\begin{aligned}
& w\left(\bar{e}_{m}\right)-w\left(\bar{e}_{m-1}\right) \leq v\left(\bar{e}_{m}, \theta\right)-v\left(\bar{e}_{m-1}, \theta\right)-c \cdot 1_{\{m=1\}} \text { for } \theta \in\left(\bar{\theta}_{m-1}, \bar{\theta}_{m}\right), \\
& w\left(\bar{e}_{m}\right)-w\left(\bar{e}_{m-1}\right) \geq v\left(\bar{e}_{m}, \theta\right)-v\left(\bar{e}_{m-1}, \theta\right)-c \cdot 1_{\{m=1\}} \text { for } \theta \in\left(\bar{\theta}_{m}, \bar{\theta}_{m+1}\right) .
\end{aligned}
$$

Notice that the LHS is in fact a function of the partition:

$$
\begin{aligned}
& \frac{\bar{\theta}_{m+1}-\bar{\theta}_{m-1}}{2} \leq v\left(\bar{e}_{m}, \theta\right)-v\left(\bar{e}_{m-1}, \theta\right)-c \cdot 1_{\{m=1\}} \text { for } \theta \in\left(\bar{\theta}_{m-1}, \bar{\theta}_{m}\right) \\
& \frac{\bar{\theta}_{m+1}-\bar{\theta}_{m-1}}{2} \geq v\left(\bar{e}_{m}, \theta\right)-v\left(\bar{e}_{m-1}, \theta\right)-c \cdot 1_{\{m=1\}} \text { for } \theta \in\left(\bar{\theta}_{m}, \bar{\theta}_{m+1}\right)
\end{aligned}
$$

By continuity, we must have:

$$
\begin{equation*}
\frac{\bar{\theta}_{m+1}-\bar{\theta}_{m-1}}{2} v\left(\bar{e}_{m}, \bar{\theta}_{m}\right)-v\left(\bar{e}_{m-1}, \bar{\theta}_{m}\right)-c \cdot 1_{\{m=1\}} \text { for } m=1, \ldots, M \tag{8}
\end{equation*}
$$

with boundary conditions:

$$
\begin{equation*}
\bar{\theta}_{0}=b, \bar{\theta}_{M+1}=B \tag{9}
\end{equation*}
$$

Notice that (8) for given $\left\{\bar{e}_{0}, \ldots, \bar{e}_{M}\right\}$ is a set of $M$ non-linear equations with $M+2$ variables $\bar{\theta}_{0}, \ldots, \bar{\theta}_{M+1}$. We already know from Proposition 1 that it has a solution at least for $M+1$. We want to determine what is the maximal $M$ that allows us to construct the equilibrium.

We shall proceed as in Crawford and Sobel (1982). Let us use the following notation: let $\theta^{i}$ denote a sequence $\theta_{0}^{i}<\ldots<\theta_{i}^{i}$ that satisfies a set of conditons (8). Let:
$K(x)=\max \left\{i: \exists \mathrm{b}<x<\theta_{2}<\ldots<\theta_{i} \leq B, \quad\right.$ satisfying (8) with a subset $\left.\left\{\bar{e}_{0}, \ldots, \bar{e}_{i}\right\}\right\}$.
Since we have a total of $S$ possible effort levels, $K(x)$ is bounded from above by $S$. Let $\bar{M}=\sup _{x \in(0,1]} K(x)$. Then $\bar{M}$ is a maximal size of an equilibrium.

I proved the existence of $\bar{M}$. Let us take $\bar{\theta}_{1}=\operatorname{argmax} K(x)$, so that $K\left(\bar{\theta}_{1}\right)=\bar{M}$. What remains to be shown is that that for any $1 \leq M \leq \bar{M}$ we could construct an equilibrium of size $M$.

Let $\theta^{K(x)}$ be a sequence of length $K(x)$ that solves (8) with a boundary condition $\theta_{1}^{K(x)}=x$. We might notice that since all the functions in (8) are continuous, then the solutions of the system vary continuously with boundary conditions. Therefore, if only $\theta_{K(x)}^{K(x)}<1$, then the function $K(x)$ is continuous at $x$ and since it is an integer-valued function, it is locally constant around $x$. Moreover, the function can change by at most one in a disutility point. Given that two arguments, since $K\left(\bar{\theta}_{1}\right)=\bar{M}$ and $K(B)=1$, we conclude that $K(x)$ must take all integer values $1 \leq M \leq \bar{M}$ on an interval $\left[\bar{\theta}_{1}, B\right]$.

I shall also examine, what happens to the choice of efforts when $K$ varies with $x$. If $K(x)$ is discontinuous at $x$, then $\theta^{K(x)}$ satisfies both (8) and (9), so it is an equilibrium of a size $K(x)$. We might also notice, that the discontinuity may happen only if $\theta_{K(x)}^{K(x)}=1$, so that the biggest effort $\bar{e}_{K(x)}$ becomes redundant, in a sense, that it is not chosen anymore by a positive measure of agents. Therefore at discontinuity points of $K(x)=M$, a set of effort choices changes from $\left\{\bar{e}_{0}, \ldots, \bar{e}_{M}\right\}$ to $\left\{\bar{e}_{0}, \ldots, \bar{e}_{M-1}\right\}$.

Assume there exists an equilibrium $b=\bar{\theta}_{0}<\bar{\theta}_{1}<\ldots<\bar{\theta}_{M+1}=B$, such as described in the Proposition. We shall show that such equilibrium is unique.

Assume the contrary, i.e. there exists another equilibrium $b=t_{0}<t_{1}<\ldots<$ $<t_{\mathrm{M}+1}=B$ such that all the efforts from $S$ are chosen with a positive probability. Notice that the sequences $\left\{\bar{\theta}_{i}\right\}_{i=1}^{M}$ and $\left\{t_{i}\right\}_{i=1}^{M}$ are two different solutions to a set of equations (8) with boundary conditions (9). Take $j=\max _{i}\left\{\bar{\theta}_{i} \neq t_{i}\right\}$. Without loss of generality, we can assume that $\bar{\theta}_{j}>t_{j}$. Since $1 \leq j \leq M$ and for $i>j$ we have $\bar{\theta}_{i}=t_{i}$ so we can use the conditions (8) to write:

$$
\left[v\left(\bar{e}_{i_{j}}, \bar{\theta}_{j}\right)-v\left(\bar{e}_{i_{j-1}}, \bar{\theta}_{j}\right)\right]-\left[v\left(\bar{e}_{i_{j}}, t_{j}\right)-v\left(\bar{e}_{i_{j-1}}, t_{j}\right)\right]=\frac{t_{j-1}-\bar{\theta}_{j-1}}{2}
$$

and the LHS by Lagrange's mean value theorem (used twice) could be rewritten in terms of $v$ 's derivative:

$$
v_{e \theta}\left(e^{*}, \theta^{*}\right)\left(\bar{e}_{i_{j}}-\bar{e}_{i_{j-1}}\right)\left(\bar{\theta}_{j}-t_{j}\right)=\frac{t_{j-1}-\bar{\theta}_{j-1}}{2}
$$

where $e^{*} \in\left[\bar{e}_{i_{j-1}}, \bar{e}_{i_{j}}\right]$ and $\theta^{*} \in\left[t_{j}, \bar{\theta}_{j}\right]$. By Assumption $1, v_{e \theta}$ is strictly negative, and the terms in brackets are strictly positive, therefore the LHS is strictly negative. We, therefore, conclude that the RHS must be strictly negative as well, and we get $\bar{\theta}_{j}>t_{j} \Rightarrow \bar{\theta}_{j-1}>t_{j-1}$. By repeating the exercise for $m=j-1$, we get:

$$
\left[v\left(\bar{e}_{i_{j-1}}, \bar{\theta}_{j-1}\right)-v\left(\bar{e}_{i_{j-2}}, \bar{\theta}_{j-1}\right)\right]-\left[v\left(\bar{e}_{i_{j-1}}, t_{j-1}\right)-v\left(\bar{e}_{i_{j-2}}, t_{j-1}\right)\right]+\frac{t_{j}-\bar{\theta}_{j}}{2}=\frac{t_{j-2}-\bar{\theta}_{j-2}}{2},
$$

and, proceeding iteratively, we conclude that $\bar{\theta}_{j}>t_{j} \Rightarrow \bar{\theta}_{i}>t_{i} \forall i<j$. But since it must also hold for $i=0$, it contradicts the fact that the sequences $\left\{\bar{\theta}_{i}\right\}_{i=0}^{M+1}$ and $\left\{t_{i}\right\}_{i=0}^{M+1}$ satisfy (9). Therefore, the initial assumption that there exist two different equilibria must have been false.

## CHOOSING A FIELD OF EDUCATION: SIGNALING, MISMATCH, AND EQUILIBRIUM SHIFTING

## Summary

In this paper the author reviews a well-known model of job market signaling through education, extending it to a choice of a field of study. In the theoretical part, she extends the classic model, by analyzing a game of education choice with continuum types of agents and discrete space of efforts, which is here interpreted as a field of study at the university level. In the second part, the author provides a simple numerical exercise to show how policy changes may influence the equilibrium. This exercise is used in the context of observed overeducation in the Polish labor market. Given the data on recent alumni's field of study and professional career, the author calibrates a stylized disutility function that would rationalize the choices within a signaling model with inelastic demand and some unobserved frictions. Then, she provides a simple illustrative argument on how an intervention by a better-informed social planner may shift the equilibrium. The author argues that overeducation may arise if the cost of getting a diploma is too small; this can lead to an over-supply of university graduates as compared with the labor market demand.

Keywords: higher education, labor market, signaling
JEL: D82, I26

# WYBÓR KIERUNKU STUDIÓW: SYGNALIZACJA, NIEDOPASOWANIA I PRZESUNIĘCIA RÓWNOWAGI 

## Stresuczenie

W artykule autorka rozszerza klasyczny model sygnalizacji na rynku pracy przez edukację, dodając do niego wybór kierunku studiów. W części teoretycznej badany jest model z nieprzeliczalną przestrzenią typów jednostek i dyskretną przestrzenią wyborów, które
tu są interpretowane jako dziedziny studiów. W drugiej częsci autorka przedstawia proste ćwiczenie numeryczne ilustrujące, jak zmiany polityki mogą wpłynąć na równowagę rynkową. Ćwiczenie to jest przeprowadzone w kontekście obserwowanego przeedukowania na polskim rynku pracy. Wykorzystując dane dotyczące kierunku studiów i kariery zawodowej absolwentów, autorka kalibruje stylizowaną funkcję dysużyteczności, która racjonalizowałaby wybory w ramach modelu sygnalizacji z nieelastycznym popytem i niedoskonałą adaptacją rynkową. Następnie pokazano prosty przykład, w którym interwencja lepiej poinformowanego planisty może zmienić równowagę. Autorka wskazuje, że zjawisko przeedukowania może występować w sytuacji, w której koszt uzyskania dyplomu jest zbyt niski, co może doprowadzić do nadwyżkowej podaży absolwentów wyższych uczelni w stosunku do rynkowego popytu na prace.

Słowa kluczowe: studia wyższe, rynek pracy, sygnalizacja
JEL: D82, I26

# ВЫБОР НАПРАВЛЕНИЯ УЧЕБЫ: СИГНАЛИЗАЦИЯ, НЕСООТВЕТСТВИЕ И ПЕРЕНОС РАВНОВЕСИЯ 

## Резюме


#### Abstract

В настоящей статье автор расширяет классическую модель сигнализации на рынке труда через образование, добавляя в него выбор направления учебы. В теоретической части исследуется модель с неограниченным пространством типов единиц и прерывистым пространством выбора, под которыми понимаются здесь направления образования. Во второй части автор представляет простое числовое упражнение как иллюстрацию того, как политические изменения могут повлиять на рыночное равновесие. Это упражнение делается в контексте имеющегося «чрезмерного образования» на польском рынке труда. Используя данные о направлениях учебы и профессиональной карьере выпускников, автор делает калибровку стилизованной функции бесполезности, которая позволила бы рационализировать выборы в рамках модели сигнализации с неэластичным спросом и несовершенной рыночной адаптацией. Затем демонстрируется простой пример, в котором интервенция более информированного плановика может изменить равновесие. Автор указывает, что явление «чрезмерного образования» может появляться в условиях, когда стоимость получения диплома слишком низка, что может привести к чрезмерному предложению выпускников вузов по отношению к рыночному спросу на труд.


Ключевые слова: высшее образование, рынок труда, сигнализация
JEL: D82, I26


[^0]:    * Joanna Franaszek, Ph.D. - College of Economic Analysis, SGH Warsaw School of Economics; ORCID: 0000-0001-9111-9055; e-mail: jfrana@sgh.waw.pl

[^1]:    ${ }^{1}$ Best to my knowledge.

[^2]:    ${ }^{2}$ With this assumption, 'productivity' should not be interpreted too literally. The fact that some people chose humanities over engineering - even if engineers' wages are higher than historians' - would in my model mean that humanities students have a lower type than engineering students. However, this abstracts from talent, intelligence or skills, and measures only some personal characteristics valued (through wages) in the market. The assumption that those characteristics are one-dimensional is clearly a severe simplification, but it makes the model tractable.

[^3]:    ${ }^{3}$ The Inada conditions in (d) could be easily relaxed, but I assume them as non-controversial.

[^4]:    ${ }^{4}$ The field of study is chosen at least 3-5 years before entering the labor market, and most often is determined by the choice of courses in high school. This choice is taken 6-9 years before entering the labor market.

[^5]:    5 The first part of the sentence is obvious. To give the idea of the proof of the second part, notice that if both $\theta$ and $e$ were continuous variables, the statement is an immediate consequence of an Envelope Theorem. For $e$ discrete, some minor technical difficulties must be considered, however, the proof follows the same simple argument.

[^6]:    ${ }^{6}$ Some reluctance in accepting the job below one's education level might be explained by the signaling theory itself - see Ma and Weiss (1993).

