I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13564

# The Bottom 20%: Early Career Paths of Adolescents with Low GPA

Jørgen Hansen
Nicolai Kristensen
Henrik L. Andersen

AUGUST 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13564

# The Bottom 20%: Early Career Paths of Adolescents with Low GPA

**Jørgen Hansen**
*Concordia University and IZA*

**Nicolai Kristensen**
*VIVE - The Danish Centre of Applied Social Science and IZA*

**Henrik L. Andersen**
*VIVE - The Danish Centre of Applied Social Science*

AUGUST 2020



ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## The Bottom 20%: Early Career Paths of Adolescents with Low GPA*

Across nations, large proportions of younger birth cohorts obtain no professional qualications. Using a structural dynamic approach, we analyze policies targeted adolescents who leave grade nine with a GPA in the bottom 20%. We find that preparatory courses, offered to young people who are unable to commence a qualifying degree, have no positive impact on future labor market outcomes. Unobservable noncognitive qualities are more important for this group than are cognitive skills. Education is a good option for some, but not for all. Implications for mechanism design and wage support schemes are discussed.

**JEL Classification:** I2, I38
**Keywords:** education, adolescents, structural dynamic programming

**Corresponding author:**
Nicolai Kristensen
VIVE - The Danish Centre of Applied Social Science
Herluf Trolles Gade 11
1052 Copenhagen K
Denmark
E-mail: nikr@vive.dk

* We thank the TRYG Foundation for financial support. We thank Paul Bingley, Michael Svarer, Torben Tranæs and seminar participants at Lund University and The Rockwool Foundation Research Unit for helpful comments.

## 1 Introduction

Decades of increased internationalization and competition have reduced the number of jobs for low-skilled persons in most industrialized countries.[1] In the future, however, an even greater threat to low-skilled workers might come from automation. Advances in robotics and artificial intelligence challenge human performance of even cognitive intensive work activities, although they are primarily expected to transform jobs that require low cognitive skills.[2] These trends put pressure on all developed countries and their economies. Particular concern and attention has been given to low-skilled youths who, in many countries, already today face low demand for their labor and skills.

This is a worldwide phenomenon. For 2018, the US Census Bureau report that 10% of all 25-34-year-olds had educational attainment below upper secondary education[3], and in the EU, about 15% of 18-24 year-olds have, at most, a lower secondary education, EU (2010).[4] More and higher education has been the main policy response to the low labor market demand for unskilled adolescents, and reducing the incidence of early school leave to 10% by 2020 is one of the key benchmarks of the European Education Policy Cooperation.[5]

However, recent advances in studies of early childhood development question investments in education as the main policy response. According to applied human capital research, cognitive ability is the most important determinant of labor market outcomes (Herrnstein & Murray 1994, Hanushek & Woessmann 2008), but proponents of an alternative view argue that noncognitive skills, such as social skills, motivation, or perseverance, are the most important personal qualities, see e.g. Bowles & Gintis (1976) and Heckman et al. (2006). Lindqvist & Vestman (2011) hold the middle ground. Based

[1]See, e.g., https://www.oecd.org/general/jobsandskills.htm.

[2]OECD (2016) asssess that half of todays work activities can be automated by 2055 and in some scenarios even by 2035. McKinsey (2017) reaches similar conclusions.

[3]See Educational Attainment in the United States: 2018, Table 3. The number includes individuals with no diploma as well as individuals with Elementary or High School GED. Available at https://www.census.gov/data/tables/2018/demo/education-attainment/cps-detailed-tables.html.

[4]Early School Leavers (ESL) is defined as the percentage of the population aged 18-24 with at most lower secondary education and not in further education or training. In the EU, 16.9% of boys and 12.7% of girls are early school leavers, EU (2010).

[5]Raw numbers lend immidiate support to this view. For instance, the average OECD unemployment rate for 25-64-year-olds with an education below upper secondary level was 12.5% in 2015, while the average unemployment rate among workers with a post-secondary education was 4.8%, OECD (2017).

on military IQ scores they find that men with low earnings and high degrees of unemployment lack noncognitive skills rather than cognitive skills while higher in the wage distribution cognitive skills becomes a stronger predictor of labor market success. The tension between these two opposing views is fundamental, in that they point to very different policies.

Despite many advances in our knowledge of early life circumstances and long-term outcomes (e.g. Aizer et al., 2016), Almond et al. (2018) conclude their comprehensive review by noting that we still know relatively little about the interval between and the dynamics of choices made during late teenage life and the transition into adulthood. To what degree are labor market outcomes all set by the end of lower secondary school? Does GPA from lower secondary school predict earnings and employment status 13 years later, or do factors related to socio economic status (SES) and unobservable personality traits matter more? Answers to these questions are crucial for policy makers and are the focus of this study.[6]

More specifically, we investigate the impact of different educational options for 16-year-old adolescents with low cognitive ability. We focus on a group with GPA in the bottom 20% of the GPA distribution at the end of grade 9. This group is of particular concern since their risk of becoming marginalized, with weak connections to the the labor market and low earnings, is relatively high, and a natural reaction is to improve their cognitive skills.

We formulate and estimate a structural dynamic model of discrete choices of employment and schooling, following work pioneered by Wolpin (1992) and Keane & Wolpin (1997). This type of model allows past and current choices to affect future options and future expected returns, so that the dynamics incorporate how past choices affect current decisions through state dependence. It also allows for dynamic selection which, if ignored, may seriously bias the estimated effects of cognitive skills and family background on educational and labor market decisions.[7] Moreover, it is impossible to evaluate the dynamics of responses to particular policy changes without an explicit model of dynamic labor market behaviour in which unobserved heterogeneity plays a critical role. The estimated model allows us to simulate

[6]A large number of papers on educational attainment, using structural dynamic methods (e.g. Keane & Wolpin 1997), have pointed out that unobserved heterogeneity in preferences for schooling is more important than any other observed characteristic in the NLSY79.

[7]The notion of dynamic selection plays a key role in many areas of economics and its statistical implications within an optimal schooling model are discussed in Cameron & Heckman (1998)

outcomes over a 12-year period from three different policy interventions that take place at the end of secondary school, an exercise that is not possible without our modelling framework.

Access to administrative longitudinal records of the entire population over many years provides a sample size and panel length large enough for us to concentrate on this at-risk group of male adolescents with low cognitive abilities. Using six months as the period length, the structural parameters are estimated based on data for 25 periods. An important additional advantage of our approach is that we allow for the joint assessment of program substitutes. Failing to handle a joint assessment may lead to downward biased estimates, see Kline & Walters (2016).[8] In addition, Denmark is an interesting country institutionally as it offers tuition free universal programs, which makes it easier to ignore issues of credit constraints and other financial barriers in the interpretation of the results.

The problem of lifting the lower tail of the workforce into employment is not new, and our work has ties to studies of other programs targeted at disadvantaged youth. In the US, for example, a series of alternative US federal employment programs targeted at low-income youth started with the passing of the Area Development Act as early as 1961. Many of these programs have been evaluated and generally found to be ineffective (LaLonde 2003). In recent years, the largest US training program for disadvantaged youth has been the Job Corps program, designed to help youths between age 16 and 24 to become more employable and productive. The success of the Job Corps is partial and mainly found to benefit the oldest participants, Schochet et al. (2008).[9,10]

Despite the fact that we select male adolescents from the same birth cohort and in the lowest 20% of ninth grade GPA, the sample remains highly heterogenous. We find that noncognitive unobservables matter much more than cogntive skills, both for educational choices and for earnings. Improv-

[8]Kline & Walters (2016) and Heckman et al. (2000) argue that the presence of close program substitutes complicates the task of program evaluation if most participants would receive similar services in the absence of an intervention.

[9]In economics, a large literature is devoted to dropouts, but rarely do these studies look into alternative types of education strategies apart from high school graduation. Eckstein and Wolpin (1999) is an exception. They allow for different types of ability and model sequential choices of high school attendance and work decisions. Large strands of literatures in both Sociology and Psychology are also devoted to drop outs; see Freeman & Simonsen (2015) for a comprehensive review.

[10]Also worth mentioning is the literature on evaluations of adult literacy programs, like the GED. For the most disadvantaged and less able, the returns are generally found to be smaller than for the more advantaged (Carneiro & Heckman, 2003).

ing the cognitive skills for this group which will likely be difficult in itself at age 16 and onwards, will not make much of a difference in terms of labor market and educational outcomes. In contrast, we find that pre-market factors such as family background characteristics (notably fathers education and family ethnicity) are important. Based on estimates of finite-mixture unobservables we identify a group of individuals we refer to as "hard-to-help", of which half are unemployed at age 28 and very few complete a qualifying education. This group remains hard to lift and even tailor-made initiatives like preparatory courses do not succeeed as stepping-stones into a qualifying education or employment. Among individuals with favorable unobservables, vocational education results in stable employment and high wage returns.

The paper proceeds with an account of the Danish educational system relevant to adolescents. Following this section, we present, in turn, the data, model, and estimation approach. Next, we present estimates and subsequently use these to simulate a series of policies and counterfactual outcomes. Lastly, we discuss our findings and conclude.

## 2 Institutional Setup: Possibilities for Adolescents[11]

In Denmark, it is compulsory to complete grades 0-9, i.e. 10 years of schooling, and completion of ninth grade usually happens the year the adolescent turns 16. After ninth grade, choices commence.

One option is to continue in tenth grade, the curriculum of which resembles ninth grade in its focus on core skills such as literacy and numeracy.[12] The optional tenth grade is predominantly undertaken by pupils with a relatively low grade point average in ninth grade, and as we show later in our selected sample, about 60% choose to attend tenth grade.

After ninth or tenth grade, the educational system offers a choice between an academic or vocational track, as illustrated in Figure 1.

The academic track starts with general education (high school of some form), which is available to all pupils who were deemed fit for high school in the ninth (or tenth) grade. This will largely depend on their GPA, although no strict cut-point existed.

[11] The educational system changes over time. We present here the system that was in place for our sample period. The most important changes since then are included partly in the simulations and partly in the discussion section.

[12] The objects clauses state that teaching in tenth grade is an option for adolescents who after lower secondary school are in need of further academic qualification and clarification regarding choice of education to be able to complete a youth education program.

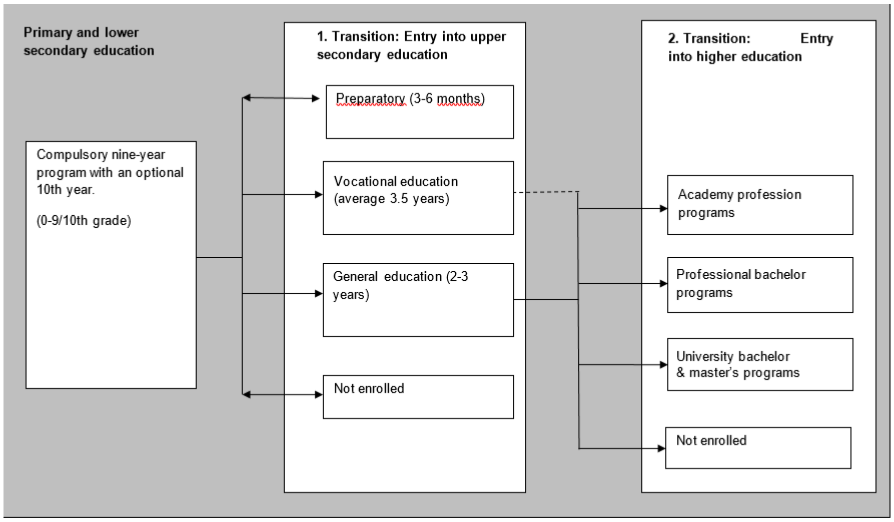



Figure 1: Illustration of the Danish education system

The vocational track, VET, is a large and multifaceted program in which education lasts 3.5 years on average. The vocational programs are based on a dual training system with a combination of school-based modules and in-company practical apprentice training. Upon completion, the students obtain a formal certificate in their profession and can undertake skilled employment in the labor market. Twelve programs exist in the Danish VET system.[13] They all start out with a basic training program, which typically lasts 20-30 weeks (one semester), and together they cover approximately 110 different vocational professions. The basic programs are school-based, whereas the remaining education is a mix of in-company training (50-70 percent of the time) and school periods. To continue from the basic to the main program, VET students normally have to obtain an apprenticeship contract with a public or private company. A major challenge over several decades

[13]The different programs (including examples of related professions) are: (1) Commercial trade: sales assistant, national mail service employee, bank clerk, office clerk. (2) Building and construction: bricklayer, plumber, glazier, woodworker. (3) Transport and logistics: truck driver, driver. (4) Mechanics: motor mechanic, bicycle mechanic. (5) Building and citizen services: building caretaker, security guard, and receptionist. (6) Media production: media graphic designer, photographer. (7) Food production and catering: baker, cook, butcher, miller, waiter. (8) Production and development: blacksmith, toolmaker, industrial operator. (9) Electronics and IT: electrician, IT supporter. (10) Styling: hairdresser, cosmetician, nail technician. (11) Animals and plants: farmer, gardener. (12) Health and care: social and health care assistant.

has been exceedingly high dropout rates from VET. The gross dropout rate is around 50%, but some 20 percentage-points of these eventually complete another education (or another VET program), leaving a net dropout rate of 30%.

Targeted at the most vulnerable adolescents, such as individuals with low cognitive skills who are not immidiately able enter high school or a VET program, a third educational option is the so-called "Preparatory Program". The purpose of a preparatory program resembles that of the Job Corps in that it serves the same age group, 16-24-year-olds, and it is supposed to make the participants more responsible and more productive. However, the main focus of the preparatory program is skills, maturity, and clarification to commence and sucessfully complete an education that ends with a qualifying degree. Failing this key objective, a secondary objective is employment. Access to preparatory programs is decided by the local city councils and thus varies across the 98 municipalities.[14,15]

# 3 Data

## 3.1 Registry Data and Selected Sample

Our study is based on very comprehensive, high-quality administrative records from Statistics Denmark based on the Central Personal Registry (CPR). Each individual has a unique CPR number, and this number is matched to the administrative registries, which makes it possible to follow individuals over time and across different types of registries. We choose a sample of the male 1986 birth cohort who completed their compulsory ninth grade schooling in the summer of 2002, and who are in the low end of the GPA distribution. Identification of at-risk students is found to be most accurate at this age group, and, in addition, low grades are found to be the most accurate single predictor of high school dropout for the US (Bowers et al. 2013).

The distribution of school grades is given in Figure 2, and our selected sample is found in the lowest 20%, indicated in grey.

In the model part of the paper we define a period as 6 months instead of the usual 12 months. This is particularly useful in the context of choices about working and education made by teenagers, who tend to move in and out of the the labor market and change field of study more often than older

[14]This variation provides variation previously used as an instrument, Rangvid et al. (2015).

[15]See Appendix 1 for more details on the content of preparatory courses.

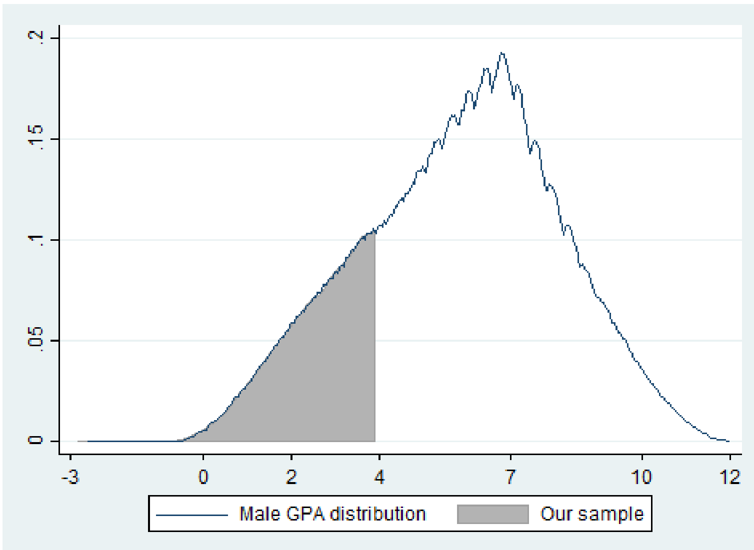



Figure 2: Distribution of male ninth grade GPA and selected sample

persons. From 2007 onwards, the registries include monthly data on income and employment. For the years 2002-2006, we constructed the biannual periods based on information about dates for begining and ending a job combined with information about hours in the job.[16]

## Measurement of Skills

We measure cognitive skills using ninth grade exam grades for selected courses, including math (written exam and mid-term test) and physics exams. However, even based on written exams, which purge teachers' assessment of individual students in the classroom, Borghans et al. (2016) establish that scores on achievement tests have independent power above and beyond "pure"cognitive (IQ) ability. In order to obtain a clean measure of cognitive skills we therefore follow Landersø & Heckman (2016) and regress the cognitive test scores on test scores on orderliness/organization/neatness grades from written Danish exam and mid-term test as well as as written math, and use the residuals from this regression as our measure of cognitive skills.

The orderliness/organization/neatness grades also enters our model, but

[16]The registries are called CON and RAS. We find complete smoothness in the average wage profile around the data source break.

we do not consider it a measure of noncognitive skills. As pointed out by Landersø & Heckman (2016), orderliness is more closely related to academic acheievement than a measure of socio-emotional skills, which probably capture noncognitive skills more accurately.[17]Socio-emotional skills are but one part of noncognitive skills, which among other features also include perseverance. Importantly however, note that with the residual-based measure of cognitive skills any remaining unobservables will, insofar as they represent skills, capture assorted noncognitive skills.

### 3.2 A First Look at the Data

In order to guide the formulation of a sequential choice model, it is instructive to analyze activity choices on a 6-month basis after completion of ninth grade. As Table 1 clearly shows, about 60% of our selected sample of male adolescents generally continue in the tenth grade for the first two semesters, after which the predominant activity is to enter the VET program (in period 3, more than 67% enter VET). In comparison, the preparatory program is relatively small, peaking in period 4 when 9.5% of the sample is enrolled in the program. After age 20, very few enter a preparatory program which reflects the intention of the Preparatory Program to serve as a stepping stone in the years immidiately following grade 9 (or 10) completion. While most of the individuals in our sample find some educational activity or employment, the share who do neither of these (and hence become a NEET, i.e. Not in Education, Employment or Training) stabilizes around one in four over the last 10-12 periods of our sample window. About 15% of our sample is in some kind of education (incl. VET) in period 25. This includes both some that are still enrolled and others who have re-enrolled.

[17]Almond et al. (2018: 1440) note that noncognitive skills typically are measured by a "hodge-podge of different measures". Our study is no exception although Landersø & Heckman (2016) report that this orderliness measure is three to five times more associated with factors such as driving under influence (DUI) and mental disorders (psychiatric admissions) than the measure of cognitive skills.

Table 1: Activity status by age and period (% of sample)

| | | | | | Choice of | Activity | | | |
|---|---|---|---|---|---|---|---|---|---|
| Period | Year | Semester | Age | School/Educ | Preparatory | VET | Employed | NEET | SUM |
| 1 | 2002 | Fall | 16.5 | 65.8 | 4.1 | 23.5 | 1.7 | 4.9 | 100.0 |
| 2 | 2003 | Spring | 17.0 | 59.3 | 5.8 | 24.8 | 3.9 | 6.2 | 100.0 |
| 3 | 2003 | Fall | 17.5 | 9.7 | 6.8 | 67.2 | 7.1 | 9.2 | 100.0 |
| 4 | 2004 | Spring | 18.0 | 8.7 | 9.5 | 61.9 | 7.8 | 12.1 | 100.0 |
| 5 | 2004 | Fall | 18.5 | 8.5 | 9.0 | 61.7 | 8.2 | 12.6 | 100.0 |
| 6 | 2005 | Spring | 19.0 | 8.1 | 8.1 | 59.0 | 9.7 | 15.1 | 100.0 |
| 7 | 2005 | Fall | 19.5 | 7.6 | 4.9 | 52.7 | 16.7 | 18.1 | 100.0 |
| 8 | 2006 | Spring | 20.0 | 7.2 | 3.8 | 52.1 | 17.9 | 18.9 | 100.0 |
| 9 | 2006 | Fall | 20.5 | 4.5 | 1.6 | 47.4 | 28.2 | 18.3 | 100.0 |
| 10 | 2007 | Spring | 21.0 | 4.3 | 1.2 | 42.4 | 31.6 | 20.5 | 100.0 |
| 11 | 2007 | Fall | 21.5 | 4.4 | 0.8 | 34.0 | 42.3 | 18.5 | 100.0 |
| 12 | 2008 | Spring | 22.0 | 4.6 | 0.8 | 23.9 | 48.3 | 22.4 | 100.0 |
| 13 | 2008 | Fall | 22.5 | 4.7 | 1.0 | 18.3 | 56.6 | 19.3 | 100.0 |
| 14 | 2009 | Spring | 23.0 | 5.0 | 1.0 | 15.8 | 52.4 | 25.9 | 100.0 |
| 15 | 2009 | Fall | 23.5 | 6.5 | 0.8 | 15.1 | 52.7 | 25.0 | 100.0 |
| 16 | 2010 | Spring | 24.0 | 7.0 | 0.8 | 14.1 | 50.6 | 27.6 | 100.0 |
| 17 | 2010 | Fall | 24.5 | 8.5 | 0.7 | 13.2 | 53.4 | 24.2 | 100.0 |
| 18 | 2011 | Spring | 25.0 | 8.7 | 0.9 | 12.1 | 52.8 | 25.5 | 100.0 |
| 19 | 2011 | Fall | 25.5 | 9.3 | 0.4 | 11.2 | 56.2 | 22.8 | 100.0 |
| 20 | 2012 | Spring | 26.0 | 9.0 | 0.4 | 10.0 | 54.2 | 26.3 | 100.0 |
| 21 | 2012 | Fall | 26.5 | 9.0 | 0.4 | 10.0 | 56.2 | 24.4 | 100.0 |
| 22 | 2013 | Spring | 27.0 | 8.6 | 0.4 | 9.6 | 55.5 | 25.9 | 100.0 |
| 23 | 2013 | Fall | 27.5 | 8.6 | 0.5 | 8.3 | 58.7 | 23.9 | 100.0 |
| 24 | 2014 | Spring | 28.0 | 8.8 | 0.4 | 7.9 | 58.1 | 24.8 | 100.0 |
| 25 | 2014 | Fall | 28.5 | 8.0 | 0.5 | 7.2 | 60.5 | 23.8 | 100.0 |

Note Number of individuals is 3,825.

# 4 The Structural Dynamic Model

Our sample consists of adolescents first observed the year they turn 16. The purpose of the model is to capture the choices of main activity and model the labor market implications of these choices over time. In turn this allows us to simulate the impact of various policies.

The available choices are:

- Enrol in general education/schooling (incl. tenth grade and high school)
- Enrol in a preparatory educational program
- Enrol in a *VET* program
- Employment
- Neither – which we will refer to as *NEET* (*Not in Education, Employment or Training*).

Every decision is made at the beginning of each 6-month period, and we follow individuals for a total of 25 periods until they are 28 years old. By modelling individual choices from as early as age 16, we minimize problems with endogenous initial conditions. Individuals are assumed to maximize the expected value of lifetime utility in a finite-time model with one choice per period (one control variable) and a number of exogenous and endogenous state variables.

## 4.1 Utility and Value Functions

A utility function adheres to each possible choice across the 25 periods. The utility associated with enrolling in formal **Schooling (education)** can be written as:

$$U_{it}^{s} = \alpha_i^s + \gamma_1^s d_{i,t-1}^s + \gamma_2^s S_{it} + \Upsilon^s X_{it} + \varepsilon_{it}^s$$

where $\alpha_i^s$ represents unobserved, permanent individual heterogeneity (ability) in the taste for education/schooling We parameterize these type-specific unobservables such that $\alpha_i^s = constant_i^s + \rho_1 Orderliness_i + \rho_2 Low\,HHincome_i + \rho_1 Immigrant_i$. The Orderliness is the orderliness/organization/neatness grades described in Section 3.1 and in Landersø & Heckman (2017). Low

HH income is an indicator that equals 1, if the family income is in the lowest quartile the year individual i completes ninth grade[18], and Immigrant is an indicator for families with a non-western background. Unobservales are further described in Section 5.2.

$S_{it}$ is a state variable that measures the accumulated number of periods of schooling (after grade 9) at the beginning of period $t$. Whether or not schooling/education was chosen in period $t-1$ is indicated by the dummy variable $d^s_{i,t-1}$, which will reflect the degree of state dependence. The term $X_{it}$ signifies a matrix of exogenous characteristics, some of which may vary over time. Among the observed covariates included is fathers education, GPA from ninth grade, indicators for tenth grade, age groups, and province (vs. greater Copenhagen area).[19] Finally, $\varepsilon^s_{it}$ is a stochastic utility shock. Similar utility functions adhere to each of the other possible choices. The utility associated with enrolling in a **Preparatory program** is:

$$U^p_{it} = \alpha^p_i + \gamma^p_1 reg_{it} + \gamma^p_2 Offer_{mt-1} + \Upsilon^p X_{it} + \varepsilon^p_{it}$$

where $\alpha^p_i$ represents unobserved, permanent individual heterogeneity (ability) in the taste for preparatory programs (parameterized in the same manner as $\alpha^s_i$), $Offer_{m,t-1}$ measures the proportion of youth in municipality $m$ (across 98 municipalities) that participated in a preparatory program lagged one period as a share of the local target group (adolescents with GPA below 4). This incorporates local variation and follows Rangvid et al. (2015), who use the municipal and time variation in this variable as an instrument. Again, $\varepsilon^p_{it}$ is a stochastic shock, and $X_{it}$ is a matrix of exogenous covariates.

Next follows the utility associated with choosing the **VET** program in period t. This utility can be written as:

$$U^{vet}_{it} = \alpha^{vet}_i + \Upsilon^{vet} X_{it} + \gamma^{vet}_1 region_{it} + \gamma^{vet}_2 PREP_{it} + \gamma^{vet}_3 dvet_{i,t-1} + \varepsilon^{vet}_{it}$$

where $\alpha^{vet}_i$ represents unobserved, permanent individual heterogeneity in the utility of enrolling in a $VET$ program. Here, exogenous variation comes from local differences in apprenticeship positions. PREP equals 1 if the individual was ever in a preparatory program; this parameter will indicate whether the program serves as a stepping stone or not, and $dvet_{i,t-1}$

[18]Computed using the OECD household equivalence scale.

[19]Means and standard deviations of these observed covariates are provided in the appendix, Table A1.

captures state dependence. The remaining parts of the utility function for VET follows the syntax of the other utility functions.

As a fourth option, the individuals may choose to work (**Employed**), and the utility of working is defined as:

$$U_{it}^{e} = \alpha_i^e + lnw_{it} + \gamma_1^e HS_{it} + \gamma_2^e VET_{it}^C + \gamma_3^e PREP_{it} + \gamma_4^e X_{it} + \gamma_5^e de_{i,t-1} + \varepsilon_{it}^e$$

where $\alpha_i^e$ represents unobserved, permanent individual heterogeneity in the taste for working, while $HS_{it}$, $VET_{it}^C$ and $PREP_{it}$ equal one, if the person has completed high school, completed $VET$ with a qualifying degree and (at least) one period of preparatory program, respectively (and zero otherwise). $de_{i,t-1}$ reflects state dependence in employment. A log-wage term enters in the utility for employment. This wage equation is expressed as follows:

$$\begin{aligned} lnw_{it} = \alpha_i^w + \gamma_1^w HS_{it} + \gamma_2^w VET_{it}^C + \gamma_3^w PREP_{it} + \gamma_4^w Immigrant_i + \\ \gamma_5^w Experience_{it} + \gamma_6^w Experience^2 + \gamma_6^w Cognitive_i + \varepsilon_{it}^w \end{aligned}$$

where $\alpha_i^w$ represents unobserved, permanent individual heterogeneity in the ability to earn a salary. The parameter $\gamma_1^w$ shows the return on completing high school or more while $\gamma_2^w$ and $\gamma_3^w$ show the economic returns to VET-completion and preparatory program participation, respectively. Ethnicity is part of the X-matrices in the various utility functions but is also included in the wage regression ($\gamma_4^w$). Finally, $\gamma_5^w$and $\gamma_6^w$represent the wage effect of experience and experience squared, and $\varepsilon_{it}^w$ is a random wage shock.[20]

The utility of the final activity, NEET, is normalized to 0.

**The value function.** The model is based on the underlying assumption that agents are forward-looking and consider the impact of current choices on future outcomes. Therefore, the choice-specific, instantaneous utility functions described above need to be adjusted to account for expectations over future utility and wage shocks. Specifically, the objective of individual $i$ in any period $t$, $t = 1, ..., T$ is to choose option $j$ in order to maximize the expected present value of his future utility stream:

[20]The model is set in a partial equilibrium framework, and therefore the employment option is available to all individuals at all times. Extending the model to include general equilibrium features with demand-side restrictions is beyond the scope of this paper. However, the permanent unobserved individual heterogeneity term ($\alpha_i^e$) may capture some of the potential lack of access to jobs in a given time period for individual i.

$$\max_j E\left\{\sum_{t=1}^{T}\beta^{t-1}U\left(\Omega_{i,t}|d_{i,t}^{j}=1\right)\right\} j=\{s,\, p,\, e,\, vet,\, neet\}$$

where $\beta$ is the discount factor and $E$ is the expectations operator. Using a Bellman equation, this complex dynamic optimization problem can be reduced to sequences of two period problems and solved recursively. Thus, for each possible choice there is a specific value function, $V_{it}^{j}\left(\Omega_{it}\right)$, equal to

$$V_{it}^{j}\left(\Omega_{it}\right)=U_{it}^{j}+\beta EV_{i,t+1}\left(\Omega_{i,t+1}|d_{it}^{j}=1\right)\; j=\{s,\, p,\, e,\, vet,\, neet\}$$

where

$$EV_{i,t+1}\left(\Omega_{i,t+1}|d_{it}^{j}=1\right)=Emax\left\{V_{i,t+1}^{s},V_{i,t+1}^{p},V_{i,t+1}^{e},V_{i,t+1}^{vet},V_{i,t+1}^{neet}\right\}$$

for $t=1,...,T-1$ and

$$V_{i,T}\left(\Omega_{i,T}\right)=max_{j=\{s,p,e,vet,neet\}}U_{i,T}^{j}$$

for period $T$. Starting from the last period ($T$), the model can be solved using backward recursions given the conditional expectation

$$EV_{i,t+1}\left(\Omega_{i,t+1}|d_{i,t}^{j}=1,\Omega_{i,t}\right).$$

In general, this expectation does not have a convenient analytical solution. However, by assuming that the utility shocks $\left\{\varepsilon_{i,t}^{s},\, \varepsilon_{i,t}^{p},\, \varepsilon_{i,t}^{e},\, \varepsilon_{i,t}^{vet},\, \varepsilon_{i,t}^{neet}\right\}$ follow an i.i.d. extreme value distribution, $EV_{i,t+1}\left(\Omega_{i,t+1}|d_{i,t}^{j}=1,\Omega_{i,t}\right)$, conditional on the wage shocks $\left(\varepsilon_{i,t}^{w}\right)$, can be expressed as

$$\begin{aligned}&\mathit{\Phi}_{i,t+1}\left(\Omega_{i,t+1}\right)=\tau\gamma+\\ &\quad\tau ln\left(\sum_{j}exp\left\{\frac{\left(U_{i,t+1}^{j}+\beta EV_{i,t+2}\left[\Omega_{i,t+2}|d_{i,t+1}^{j}=1,\Omega_{i,t+1},\varepsilon_{i,t+1}^{w}\right]\right)}{\tau}\right\}\right)\end{aligned}$$

where $\tau$ is a parameter of the extreme value distribution and $\gamma$ is Euler's constant.

### 4.2 The Optimization Problem and Solution Method

While the parametric assumptions made in this paper allow us to obtain an analytical expression of the $EV$ functions, they are still computationally demanding as they need to be evaluated at each possible value of the endogenous state variables $\left(S, P, E, VET^C, NEET\right)$ for each parameter iteration.[21] In addition, the wage shocks must be integrated out from the $EV$ function.

To reduce the computational complexity of the model and increase the speed of estimation, we follow Geweke and Keane (2000) and approximate $EV_{i,t+1}\left[\Omega_{i,t+1}|d^j_{i,t}=1,\Omega_{i,t},\varepsilon^w_{i,t}\right]$ by using a polynomial in the above mentioned state variables. Specifically,

$$\widetilde{EV_{i,t+1}}\left[\Omega_{i,t+1}|d^j_{i,t}=1,\Omega_{i,t},\varepsilon^w_{i,t}\right]=f\left(\Omega_{i,t+1}|d^j_{i,t}=1,\Omega_{i,t}\right)$$

where $f(\Omega)$ is a second-order polynomial in $S_{i,t}$, $P_{i,t}$, $E_{i,t}$, $VET^C_{i,t}$ and $NEET_{i,t}$:

$$\begin{aligned}
f\left(\Omega_{i,t+1}|d^j_{i,t}=1,\Omega_{i,t}\right) = \quad & \varrho_0+\varrho_1 S_{i,t}+\varrho_2 P_{i,t}+\varrho_3 E_{i,t}+\varrho_4 VET^C_{i,t}+ \\
& \varrho_5 S^2_{i,t}+\varrho_6 P^2_{i,t}+\varrho_7 E^2_{i,t}+\varrho_8\left(VET^C_{i,t}\right)^2+ \\
& \varrho_9 S_{i,t}P_{i,t}+\varrho_{10}S_{i,t}E_{i,t}+\varrho_{11}S_{i,t}VET^C_{i,t}+ \\
& \varrho_{12}P_{i,t}E_{i,t}+\varrho_{13}P_{i,t}VET^C_{i,t}+\varrho_{13}E_{i,t}VET^C_{i,t}
\end{aligned}$$

for $j \in (s, p, vet, e, neet)$.

This approach has recently gained additional support as it appears robust to misspecifications of the dynamic economic model being estimated (Jørgensen & Tô, 2018).

# 5 Estimation

## 5.1 Identification

To identify the conditional choice probabilities, we rely on the dynamic nature of choices and on exogenous variation in variables that partially determine access to certain options. Specifically, for the schooling, VET and

[21] In Appendix 2, the law of motion for each of these endogenous state variables is clarified.

employment options, identification is obtained in part by including indicators for choosing each option in the previous time period. In addition to the lagged choice indicators, we also utilize information on regional differences in apprenticeship (VET) positions over time. For the Preparatory program option, identification is aided by using local information on the proportion of youth in individual i's municipality (the share of the local target group) that participated in a preparatory program the previous period. Finally, we use information on past sequences of enrolment in VET to help identify the parameters in the probability of completing VET. Moreover, we exclude information on father's education from the wage equation (this information is included in the utility functions for schooling, preparatory program and VET). The assumption that family background characteristics only impact wages via educational choices is common in the literature and key in this paper in order to identify the wage equation parameters from the observed wages of the selected group of workers.

As is the case for all structural models set in an intertemporal framework, the estimation of our model requires some explicit functional forms.[22] Standard IV approaches, which are generally based on outcomes and choices observed at a single point in time, are not applicable in our framework as it is not possible to obtain valid instruments for the sequences of outcomes and choices over time.

### 5.2 The Likelihood Function

The utility error terms $\left(\varepsilon_{it}^{s}, \varepsilon_{it}^{p}, \varepsilon_{it}^{vet}, \varepsilon_{it}^{e}\right)$ are assumed to be i.i.d. and follow an extreme-value distribution (Rust, 1987). At the beginning of each period, the individual chooses an action based on the utilities associated with each option and an information set containing realizations of random shocks and the state variables described above is formed.

This means that we can write the probability for individual $i$ of choosing option $j$ ($j \in \{1, .., J\}$) in period $t$ as:

$$Pr\left(d_{it}^{j} = 1 | \Omega_{it}\right) = \frac{exp\left(V_{it}^{j}\left(\Omega_{it}\right)\right)}{\sum_{m=1}^{J} exp\left(V_{it}^{m}\left(\Omega_{it}\right)\right)}$$

In addition to information on choices, the likelihood function includes a density function for log earnings where we assume that $\varepsilon_{it}^{w} \sim N\left(0, \sigma_{w}^{2}\right)$.

[22] Although IV models are typically presented without explicit functional form restrictions, their interpretation also requires implicit restrictions, which may take the form of very specific functional form assumptions, cf. Keane (2010).

We have introduced persistent, unobserved heterogeneity in the form of finite mixtures (Heckman and Singer, 1984) and estimate them using an EM-approach (Train 2008, Arcidiacono & Miller 2011). In particular, we assume that there are $K$ $(= 4)$ types of individuals, each endowed with a value of $(\alpha_k^s, \alpha_k^p, \alpha_k^e, \alpha_k^{vet}, \alpha_k^w, \alpha_k^{neet}, \theta_k)$, $k = 1, ..., 4$. To relax the orthogonality conditions between unobserved heterogeneity and observed characteristics, we allow the type proportions to depend on a measure of noncognitive skills, an indicator for low family income and an immigrant indicator.

**VET completion**

As discussed above in Section 2, a common feature of the VET program is the high dropout rate after completion of the basic, one semester training program. In order to account for this feature of the data, we model the probability of completing a VET program as follows:

$$\begin{aligned} Prob\left(\triangle_{i,t} = 1 \mid dvet_{i,t} = 1\right) = \alpha_i^{vc} + vc_1 Venrol_{i,t}^4 + vc_2 Venrol_{i,t}^7 + vc_3 PREP_{i,t} + \\ vc_4 UE\, lowskilled_{a,t-1} + vc_5 Cognitive_i + \varepsilon_{it}^{vc} \end{aligned}$$

where $Venrol^l, l \,\epsilon\, (4, 7)$ are indicators that equal 1 if the number of time periods the individual has been enrolled in a VET program is at least as high as 4 (7), and $PREP$ is an indicator function that identifies those who enter VET after completing the preparatory program. The covariate $UE\, lowskilled_{a,t}$ measures the local unemployment rate in local area a in period $t-1$. This captures the outside option of dropping out and take up of unskilled employment. Finally, the completion probability is also a function of the cognitive skills of the individual.

Combining all elements, the likelihood contribution of individual $i$, conditional on the permanent, unobserved heterogeneity decsribed above, is:

$$\begin{aligned} l_{it}\left(|type\, k\right) = & \left\{Pr\left(d_{it}^j = 1 | \Omega_{it}, J\right)\right\}^{I(j \in \{1,..,J\})} * \\ & \left[Pr\left(\Delta_{it} = 1\right)^{I\left(VET_{it}^C = 1\right)} * \left(1 - Pr\left(\Delta_{it} = 1\right)^{\left(1 - I\left(VET_{it=1}^C\right)\right)}\right)\right]^{dvet_{i,t}=1} * \\ & f_{it}\left(w\right)^{I\left(d_{it}^e = 1\right)} \end{aligned}$$

Weighting the conditional likelihood contributions above by the type probabilities yields the unconditional-on-type log-likelihood contribution,

$$logL\left(|\right) = \sum_{i=1}^{N} ln\left(\sum_{k=1}^{K} \pi_k * \prod_{t=1}^{T} l_{it}\left(|type\, k\right)\right)$$

# 6 Model Estimates

The model comprises 59 parameters to be estimated, and we will therefore not discuss every single parameter estimate. Moreover, several of the covariates enter non-linearly in more than one equation, and simulating the model is therefore the best way to understand the net effect.

In order to develop these simulations we employ a bayesian approach to allocate each individual to the unobserved types.

## 6.1 Model Fit

It is instructive to first investigate the model's ability to fit the actual data. The model is able to fit the observed choices relatively precisely. In some periods the observed data show rather pronounced discrete jumps (this is the case after tenth grade – the shift between the second and third period), and this represents a challenge to any model, including ours, but the overall picture is that the model fits well, see Table 2.

From period 9 to approximately period 13, the model fit is weaker. This is due to large idiosyncratic variation in how long it takes to complete VET. In the data, we observe a large dicscrete jump from 17.9% working in period 8 to 28.2% working in period 9. The model follows the same pattern but with some discrepancy. However, in later periods it is quite precise.

Table 2: Predicted and actual choices, by period

| Choice | School | School | Prep | Prep | VET | VET | Work | Work | NEET | NEET |
|---|---|---|---|---|---|---|---|---|---|---|
| Period | Model | Data | Model | Data | Model | Data | Model | Data | Model | Data |
| 1 | 0.661 | 0.658 | 0.051 | 0.041 | 0.148 | 0.235 | 0.055 | 0.017 | 0.084 | 0.049 |
| 5 | 0.061 | 0.085 | 0.084 | 0.090 | 0.603 | 0.617 | 0.110 | 0.082 | 0.141 | 0.126 |
| 10 | 0.036 | 0.043 | 0.008 | 0.012 | 0.268 | 0.424 | 0.317 | 0.316 | 0.372 | 0.205 |
| 15 | 0.046 | 0.065 | 0.013 | 0.008 | 0.085 | 0.151 | 0.524 | 0.527 | 0.332 | 0.250 |
| 20 | 0.079 | 0.090 | 0.003 | 0.004 | 0.065 | 0.100 | 0.527 | 0.542 | 0.327 | 0.263 |
| 25 | 0.097 | 0.080 | 0.003 | 0.005 | 0.050 | 0.072 | 0.632 | 0.605 | 0.217 | 0.238 |

### 6.2 Parameter Estimates

All structural parameters are presented in the appendix. The parameters all appear reasonable and with expected signs. Here, we will mention a few general findings of importance for our study. The importance of many of the key parameters will be emphasized as part of the simulations below.

**Fathers education.** We find that father's education is very important for their sons' choices later in life; an intuitively plausible finding that links to both nature and nurture, see Holmlund et al. (2011).[23] In our selected sample, if the father has a college degree the son is much more likely to choose the academic track of further schooling/education. Similarly, if the father has a qualifying degree in VET the probability that the son will also enter VET increases substantially and, at the same time, the probability that the son will enter a preparatory program is reduced. The father's ability is likely passed on both genetically and through learning-by-doing in the upbringing. Also, fathers with a VET background can likely often assist their son in finding a VET apprenticeship position.

**VET and Preparatory Training.** The vocational education (VET) is known and recognized for having a high quality, and it's graduates are generally in high demand. This is consistent with our model estimates. Completion of VET enters with a high and positive parameter in the equation for Work (employment) and in the wage equation. In contrast, the parameter estimates for participating in preparatory training does not appear to serve as a stepping stone neither for VET participation nor completion. It also enters negatively in the work-equation, while it is insignificant in the wage equation. The net result and average impact are demonstrated in the simulations below.

**Immigrant.** Non-western immigrants enter the model estimates in the utility equations as well as the wage equation. In all places, their outcomes are unfavorable. An indicator for non-western immigrants is also included in the expression that captures unobservables.

**Unobservable Characteristics.** The unobserved components are of great interest. Their interpretation is mixed as they pick up several types of

[23] We recognize that nature and nurture are not separately identifiable; acquired skills and ability cannot be sharply distinguished (Heckman, 2007).

Table 3: Constant terms across Types and Equations

| Equation | Type 1 | Type 2 | Type 3 | Type 4 |
|---|---|---|---|---|
| Education | 0.7395 | 1.3559 | 0.0014 | 2.0863 |
| Preparatory | 0.1540 | 0.7436 | -0.4101 | 1.0606 |
| VET | 1.6937 | 3.1629 | 0.6706 | 3.8566 |
| VET Completion | -12.2394 | -1.9278 | -7.9888 | -6.3661 |
| Work | -7.5147 | -6.4946 | -8.0761 | -6.1164 |
| Wage | 11.5433 | 11.2073 | 11.1552 | 11.3174 |
| | | | | |
| Share | 28 pct. | 11 pct. | 14 pct. | 47 pct. |

unobservable time-constant characteristics such as noncognitive skills, social skills, motivation and health. In Table 3, we show all the type-specific time-invariant constants from the various equations that constitute the overall model.

Table 3 shows a remarkable consistency across the various equations. The 14% of the sample that are in Type 3 (and to a large degree also the 28% in Type 1) might be labelled "the hard to help". They have unfavourable unobservables in all activities we measure, and their chances of completing VET (should they ever enter) are much lower than for the remaining 58% of the sample. Type 1 is the type with the highest probability of being a non-western immigrant while non-western immigrants are absent from Type 3. Type 3 has a particularly high probability of being in NEET. We demonstrate this more in-depth in Section 7.2.

Type 2 and 4 perform much better. Type 4 is the best type cognitively and also the type that has the highest probability of participating in VET while Type 2 has the highest VET completion rate, conditional on having started.

Overall, this underscores how heterogeneity remains massive even though we have selected boys from the same birth cohort who all have low GPA in ninth grade. Despite this strong selection, our sample remains highly diversified with respect to both observables and notably unobservable noncognitive skills. We will investigate the extent and consequences of this heterogeneity further below.

# 7 Simulations of the Model

In this section, we simulate the consequences of three thought experiments.

Table 4: Simulated Percentage-point change in choices if cognitive skills could be increased by 1 standard deviation

| Period | School/Educ | Preparatory | VET | Work | NEET |
|---|---|---|---|---|---|
| 1 | 2.2 | -0.4 | -0.5 | -0.3 | -0.8 |
| 5 | 1.3 | -0.2 | 0.4 | -0.5 | -0.9 |
| 10 | 0.9 | -0.1 | 8.7 | -3.1 | -6.5 |
| 15 | 0.9 | 0.1 | 1.5 | -3.4 | 0.9 |
| 20 | 1.3 | 0.0 | 0.3 | -1.5 | -0.2 |
| 25 | 2.0 | 0.0 | 0.3 | -1.3 | -0.8 |

### 7.1 Improving the Cognitive Background

In our first simulation, we make the thought-experiment that an early intervention has already taken place and resulted in a rise in the cognitve skills of all pupils by one standard deviation, thus increasing the cognitive skill level but upholding the same unobserved, noncognitive individual features as well as keeping all other covariates constant.

A one standard deviation increase in cognitive skills – i.e. a substantial increase that would be very difficult to achieve through improvements in the educational system, even with massive early interventions – has a relatively modest impact on the choices made, see Table 4, although this simulated improvement in cognitive skills does increase the choice of schooling across (almost) all periods.[24] The extra participants come from a reduction in each of the four alternatives. However, the changes are generally small and after period 12 they are minimal. The impact on VET shows a more complex but rather intutive pattern. In the first three periods, VET participation is lower when the sample has better cognitive skills. This reflects that some of those who undertake VET initially would shift to an academic track. However, fewer would enter preparatory courses or stay NEET. Instead, they will choose tenth grade, and once this is completed they tend to start in the VET program. The net effect is an increase in VET participation after period 4. From period 16 onwards, the change in VET is miniscule.

In total, the average number of periods in the school option rises from 2.85 to 3.19, the share who complete VET by the end of period 25 decreases marginally from 55% to 54%, and average annual wages decrease by 299 USD (or 0.6%), see Table 5. This drop in wages is due to the fact that

[24]In the population, the standard deviation is 2.27 while in the sample it is 1.86.

Table 5: Period 25 outcomes, Baseline and simulation of cognitive skills increased by 1 standard deviation

| Model | Annual wage (USD) | Completion of VET | Avg. years in School |
|---|---|---|---|
| Baseline | 50,503 | 55 | 2.85 |
| Cognitive+1 | 50,204 | 54 | 3.19 |

higher cognitive skills increase investments in education and hence postpone entrance in the labor market.

### 7.2 The Importance of Unobservable Non-cognitive Skills

In order to better understand the role and importance of cognitive skills, it is instructive to compare the results of the above simulation of improved cognitive skills to the importance of non-cognitive skills, as captured by the unobservable, type-specific constant terms. We illustrate the value of unobserved type and cognitive skills by showing the predicted employment, NEET participation rates, and wages in period 25, as well as average years in school (after grade 9), by Type and cognitive ability. The cognitive ability is divided into quartiles (so quartile 1 is the lowest cognitive ability), see Table 6.

Table 6: Outcomes in period 25, by unobserved type and cognitive ability

| Type | Share (pct.) | Cognitive | Employeed | NEET | Annual Wage (USD) | Share completed VET | Avg. years in School |
|---|---|---|---|---|---|---|---|
| 1 | 28 | 1 | 42 | 28 | 48,143 | 2 | 3.3 |
| | | 2 | 50 | 23 | 48,287 | 2 | 3.3 |
| | | 3 | 43 | 24 | 52,106 | 3 | 3.8 |
| | | 4 | 43 | 15 | 51,845 | 3 | 4.4 |
| 2 | 11 | 1 | 72 | 22 | 47,927 | 91 | 1.7 |
| | | 2 | 70 | 27 | 48,361 | 85 | 2.1 |
| | | 3 | 71 | 24 | 47,966 | 91 | 2.1 |
| | | 4 | 68 | 19 | 47,270 | 85 | 2.7 |
| 3 | 14 | 1 | 28 | 48 | 32,344 | 3 | 2.1 |
| | | 2 | 22 | 55 | 30,365 | 3 | 2.1 |
| | | 3 | 20 | 56 | 31,689 | 7 | 3.0 |
| | | 4 | 25 | 48 | 30,965 | 7 | 3.4 |
| 4 | 47 | 1 | 77 | 14 | 55,052 | 85 | 2.3 |
| | | 2 | 80 | 11 | 53,779 | 83 | 2.8 |
| | | 3 | 78 | 14 | 54,368 | 90 | 2.4 |
| | | 4 | 80 | 10 | 54,772 | 87 | 2.6 |

Note: "Cognitive is divided into quartiles. DKK-USD exchange rate as of 2014.12.31 was 6.1531.

We generally see pervasive differences in period 25 outcomes between types but with very modest differences between quartiles of cognitive skills *within* types. The "hard to help" Type 3 group has a predicted average annual wage of approximately 31,500 USD in period 25. About half of these Type 3 individuals are NEET in period 25, and only one in four are employed in period 25 (when they are 28 years of age). In contrast, the type 4 group is predicted to have about 80% in employment in period 25, about 10-14% NEETs, and an average predicted annual wage in excess of 50,000 USD. These predictions do not depend much on cognitive skills, but the type-specific differences are substantial and much more important than the simulated increase in cognitive skills seen in the analysis above.

Cognitive skills do matter though. Within types, the highest quartile of cognitive skills has the lowest share in NEET in period 25 and for Type 1 this share is only half of the lowest quartile of cognitive skills (28 % and 15 %, respectively). Type 1 is the group of our selected sample that has the highest tendency to choose the school option and, conditional upon starting in VET, the lowest probability of completing VET. This can be explained by a low share of fathers with a VET background among Type 1 individuals, and a very high share of fathers with a VET background among Type 4 individuals.

So far, we have established that the 14% we find in unobserved Type 3 is a group of concern. However, looking at period 25 outcomes may mask important dynamics. We therefore plot the development in NEET and VET participation over all 25 periods, by Type, see Figures 3 and 4.

Figure 3: Probability of choosing VET, over 25 periods and by quartile of cognitive skills

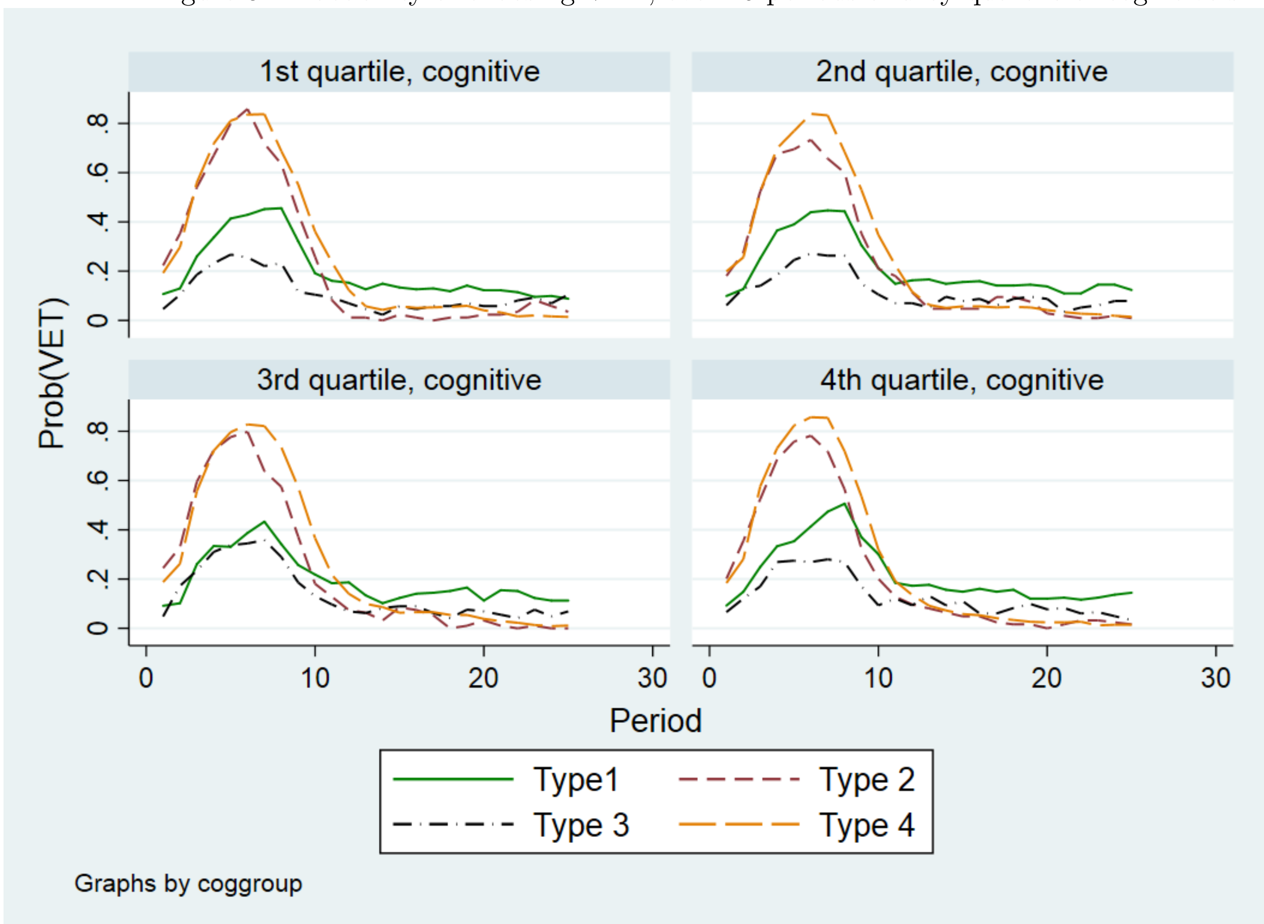
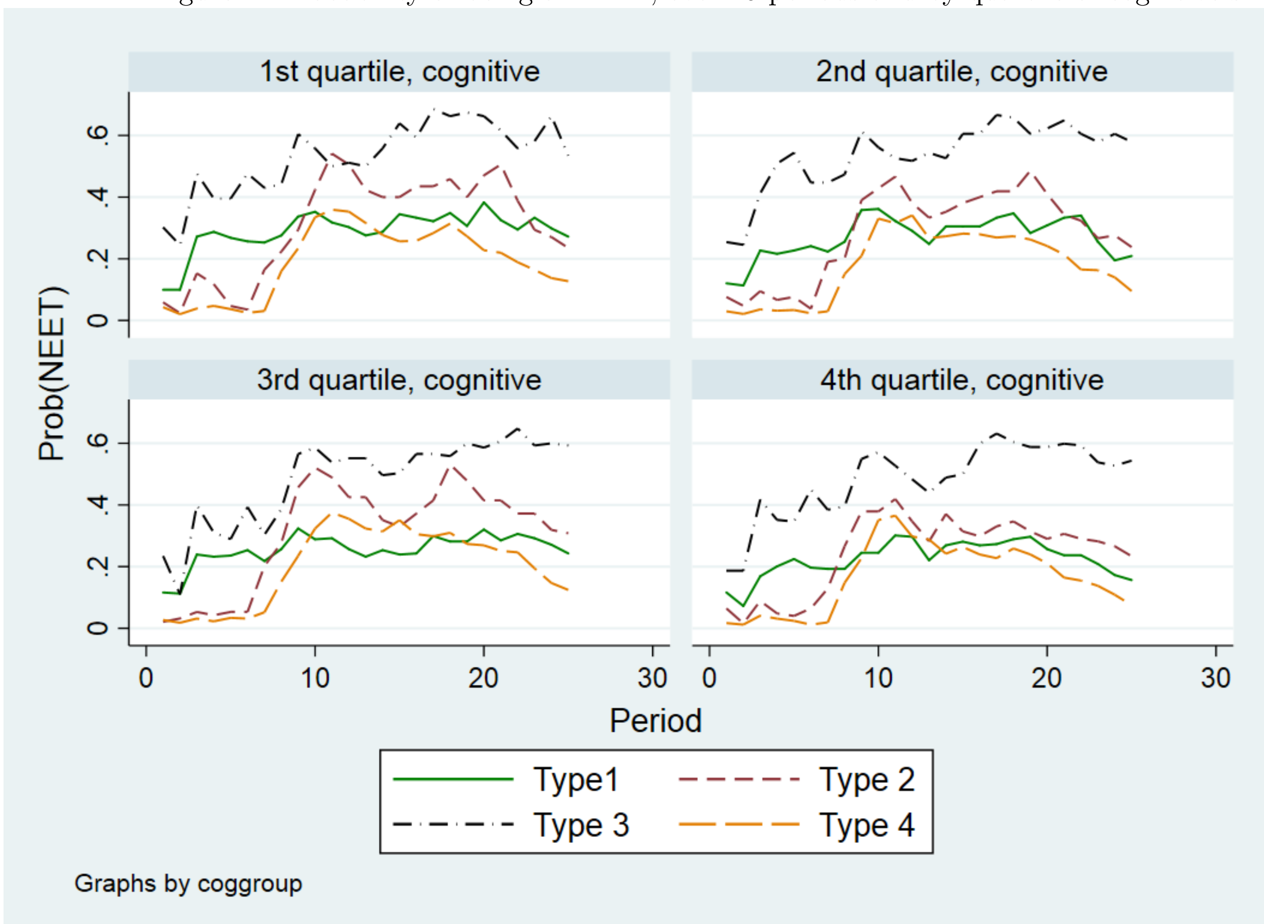


Figure 4: Probability of being a NEET, over 25 periods and by quartile of cognitive skills



Across all 25 periods, the probability of being NEET (Figure 4) is much higher for Type 3 than for the other three types, and cognitive skills matters almost nothing here, though the highest quartile cognitive group (i.e. quartile 4) has a lower level in period 25. These figures makes it clear how pervasive the differences are within this seemingly homogenous sample and that they relate to noncognitive unobservables rather than cognitive skills.

### 7.3 GPA Limit on VET Admission

As of January 2015, a GPA threshold requires a grade of 2 or above in Danish and Math to be eligible to enter a VET program. This limitation was introduced in order to mitigate the very high dropout rates from VET.[25]It is therefore obvious to simulate the outcome of such a policy on cohorts that came prior to the reform, in order to learn what outcomes the model predicts. The model's predicted impact on choices over time is shown in Table 7, and on period 25 outcomes in Table 8. When the VET eligibility constraint was imposed, we also constrained access to schooling (after grade 10) since anyone who failed the grade 2 threshold would also be unable to enter high school.

We find that VET activity decreases in all periods, especially in the first 10 periods, while more enter tenth grade, after which schooling also decreases as a result of the eligibility requirements. The vast majority of the constrained individuals obtain employment, while preparatory class participation is constant and NEET almost constant, except in period 10 where it decrease because former VET participants would otherwise be temporarily unemployed just around period 10. With the share working, in unskilled rather than skilled jobs, rise as a result of this policy, there is a temporary positive impact on average earnings, but after period 10 (where VET students would start to graduate and obtain skilled employment) the average earnings decrease, see Figure 5.

The new policy of an eligibility threshold of 2 in Danish and Math is estimated to reduce the share completing VET by 14 percentage points in our sample, Table 8. This compares to the actual drop seen in 2019, which is around 10 percentage points for the program overall. Since our selected sample is over-represented among participants in the VET program, these numbers appear reasonably in sync with one another.

[25]Prior to the passing of the law, politicians and professionals in the field debated whether cognitive or noncognitive skills was the cause of these dropouts.

Table 7: Simulated percentage-point change in choices if GPA > 2 requirement for VET eligibility

| Period | School/Educ | Preparatory | VET | Work | NEET |
|---|---|---|---|---|---|
| 1 | 2.1 | 0.1 | -2.6 | 0.4 | 0.3 |
| 5 | -1.3 | 2.7 | -11.6 | 7.1 | 3.2 |
| 10 | -0.6 | 0.0 | 1.4 | 6.2 | -7.1 |
| 15 | -0.8 | 0.1 | -0.5 | 2.7 | -1.6 |
| 20 | -1.3 | 0.1 | -1.1 | 3.5 | -1.2 |
| 25 | -1.6 | 0.2 | -0.9 | 3.0 | -0.7 |

Figure 5: Simulation of change in average annual wages (USD) with GPA threshold

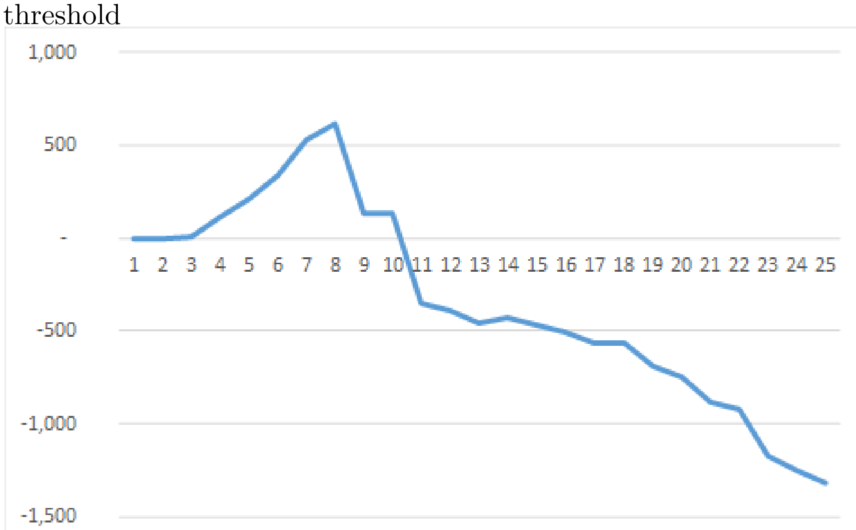



Table 8: Simulation of GPA threshold, impact in period 25

| Model | Annual wage (USD) | Completion of VET | Employment | NEET |
|---|---|---|---|---|
| Baseline | 50,503 | 55 | 63 | 22 |
| GPA over 2 | 49,188 | 41 | 66 | 21 |

Table 9: Period 25 outcomes, baseline and a simulation of forced preparatory

| Type | Change in annual wage (USD) | Change in share completing VET (pct-point) |
|---|---|---|
| 1 | -227 | -2 |
| 2 | -70 | -0 |
| 3 | -252 | -9 |
| 4 | -610 | -3 |

Table 10: Simulated percentage-point change in choices if forced education

| Period | School/Educ | Preparatory | VET | Work | NEET |
|---|---|---|---|---|---|
| 1 | 0.0 | 8.4 | 0 | 0 | -8.4 |
| 5 | 0.1 | 9.1 | -2.3 | -0.3 | -6.5 |
| 10 | 1.1 | 0.2 | 6.2 | -2.9 | -4.7 |
| 15 | 1.4 | 0.5 | 0.3 | -3.5 | 1.3 |
| 20 | 2.1 | 0.0 | -0.2 | -2.8 | 0.7 |
| 25 | 2.7 | 0.1 | -0.2 | -2.7 | 0.1 |

### 7.4 Forced Preparatory Participation

In this section, we simulate the outcome of "forced education". Such a policy is mostly relevant to individuals in the NEET group, and we choose to simulate the outcome where those who would otherwise be NEET in periods 1-8 are forced to participate in a preparatory class for up to 2 periods (one year). The literature on economics of human capital underscores the importance of self-selection and the principle of comparative advantage, see, e.g., the seminal paper by Willis & Rosen (1979), which generally means that forcing education beyond what an individual would choose by him/herself may not be optimal.

Given the one-sided parameter estimates, which all point to a negative outcome of partcipation in preparatory courses (no stepping stone to VET, nor to employment and with no positive effects in the wage equation or the VET completion equation) forced participation in preparatory courses turns out to be a bad policy avenue, see Tables 9 and 10.

Across all four types, especially type 4, period 25 wages decrease as a result of forced preparatory participation. The share completing VET decreases as does the share in employment across the various periods. The reason why forced preparatory course participation may be worse than being NEET is that NEETs do afterall spend time searching for a job and that forced course participation (with no positive impact) takes away some of the time for that search effort. Similar incapacitation effects are seen in evaluations of active labor market policies (see McCall et al. 2016 for a recent review). From a fiscal policy point of view, the bottom line is further detoriated when the average cost of some 16,000 USD/person (100,000 DKK) for preparatory courses is taken into account.

### 7.5 Transitions

The entries in Table 11 show predicted transitions across the five choice options. For each option in period t-1, the top figures show the conditional probability of entering a given state (row percentages), while the bottom numbers reveal the proportions in a given state who originated in one of the five possible states. For example, 69.6% of those who were enrolled in school in period t-1 remain enrolled in the subsequent time period while 87.5% of those in school in period t were in school in period t-1.

Overall, across types the entries also show that the most common destination after school is VET (12.9%) followed by NEET (9.5%). Further, the persistence in PREP is low in comparison with the other options (by construction), and after PREP the most common destination is NEET. The proportion entering NEET exceeds the sum of all other destination probabilities. VET students also leave for NEET but at a lower rate than PREP and the same is true for employment. Finally, the last set of numbers show that NEET is less persistent than both VET and employment and that 22.1% leave NEET for employment in any given time period (compared to 8.4% who enter VET). This illustrates the temporary nature of NEET for this group, where the risk of entering NEET is relatively high for those who are not in school or training but where about a third leave NEET for work or training.

In our discussions above, we have clearly demonstrated the importance of non-cognitive skills, and this is further illustrated in Tables 18-21 in the Appendix. These show transition matrices similar to the one in Table 11 but separately for each type. There are significant differences in some key transition probabilities across the types, especially between Type 3 and the rest. Specifically, the entry rates to NEET are substantially higher for Type 3 individuals than for the others, regardless of original state. For example, the transition probabilities from school, PREP, VET, and employment into NEET are 22%, 63.8%, 33.5%, and 25.4%, respectively. The corresponding rates for Type 4 individuals are substantially lower: 3.9%, 23%, 13.2%,

Table 11: Transition matrix: average across 25 periods

| | | | Choice (t) | | | |
|---|---|---|---|---|---|---|
| Choice (t-1) | School | Preparatory | VET | Work | NEET | Total |
| School | 69.6 | 3.5 | 12.9 | 4.5 | 9.5 | 100.0 |
| | 87.5 | 15.3 | 6.5 | 1.3 | 4.1 | 11.5 |
| Preparatory | 2.6 | 17.9 | 22.0 | 12.6 | 44.9 | 100.0 |
| | 0.8 | 19.2 | 2.7 | 0.9 | 4.7 | 2.8 |
| VET | 0.7 | 2.0 | 75.0 | 5.5 | 16.7 | 100.0 |
| | 1.9 | 18.1 | 76.4 | 3.4 | 14.6 | 23.4 |
| Work | 0.7 | 0.8 | 3.1 | 84.6 | 10.8 | 100.0 |
| | 2.9 | 10.4 | 4.9 | 79.3 | 14.5 | 36.1 |
| NEET | 2.4 | 3.7 | 8.4 | 22.1 | 63.4 | 100.0 |
| | 7.0 | 37.0 | 9.6 | 15.1 | 62.1 | 26.2 |
| Total | 9.1 | 2.6 | 23.0 | 38.5 | 26.8 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

and 7.9%, respectively. Further, Type 3 individuals have a much higher persistence rate in NEET than do any of the other types.

## 8 Discussion and Conclusion

Our discrete choice dynamic structural approach with choices made every 6 months is particularly useful when it comes to understand the early years of career paths of adolescents who relatively frequently choose new activities. It is clear from the results presented above that there is significantly more variation in labor market outcomes across unobserved types than across the distribution of cognitive skills. That is, even within a tightly selected group of high school students, there is lots of unobserved heterogeneity of large relevance for individuals' careers. Accounting for this heterogeneity in the model, and thereby purging the time-varying utility and wage shocks of time-invariant individual effects, is likely to reduce the bias in estimating the effects of cognitive skills and other variable on outcomes compared to a naïve model that ignores the presence of time-invariant shocks. Furthermore, our specification allows for correlated shocks to the utilities associated with each choice, wage shocks as well as shocks determining VET completion.

Allowing for a flexible covariance structure of the shocks is likely to lead to further bias reduction.

Adolescents with a low grade point average from compulsory school constitute a challenge to all industrialized countries. As a group, they are highly overrepresented among individuals who never complete a qualifying degree and they are more likely to have loose connections to the labor market later in life.

We find that their future success is closely related to their unobservable "Type" and pre-market, time-constant family background, notably father's education. Programs designed to help those who are unable to pass on directly to a qualifying education after ninth (or tenth) grade are unsuccessful. Participation in these preparatory courses does not lift them into a VET program, and even if they do start VET, having taken preparatory classes does not improve their chances of completing VET, nor does it improve employment chances or earnings. This underscores the inability of even the elaborate Danish welfare state to effectively tackle intergenerational distribution problems: In Denmark, access to education is generous. There are generally no tuition fees and students are even paid a Government subsidy to take an education. The welfare state can make these pecuniary reallocations, but we find that it is surprisingly inept when it comes to intergenerational educational mobility; a finding that corroborates Landersø & Heckman (2017).

This has important ties to the early intervention literature, notably a series a papers by Cunha, Heckman, and co-authors (Cunha et al. (2006, 2010) and Cunha & Heckman, 2009). They argue that it is too difficult and too expensive to lift the low-achieving adolescents compared to earlier interventions in kindergarden, primary, and lower secondary schooling. Complementarity between cognitive skills and investments becomes weaker as children age, and, as a consequence of this finding, in later stages of childhood (i.e. in adolescence) it becomes easier to remediate early disadvantage using investments in non-cognitive skills. Our results corroborate this result. Our sample is selected among those with low cognitive skills but the cure for this group is not further attempts to improve their cognitive skills. On the other hand, the preparatory courses do focus on maturity, vocational, and other non-cognitive skills (with the objective of further education), and yet they remain unsuccessful. The "hard to help" remain hard to help.

A classic result in economics, almost from time immemorial, is the principle of comparative advantage and self-selection, see Willis & Rosen (1979) for a seminal contribution. An important consequence of this is that one should be cautious about forcing education upon individuals rather than inducing education through (economic) incentives, since the educational outcome of those *already* in education obviously is irrelevant for those who choose otherwise.[26] When it comes to adolescents with the lowest cognitive skills, education may not be the best medicine. Especially not forced education. Maintaining a mechanism design that allows individuals to self-select (also based on characteristics unobserved to planners) appears wise.

According to the traditional human capital view, cognitive skills is the most important feature for future success in the labor market. For the population overall, this result may be true, but for the lower tail of the cognitive skill distribution, we find that unobservable noncognitive skills and time-constant family characteristics are much more important.

Either the preparatory program should be improved or some of the individuals from the "bottom 20%" are targeted too late. For the "hard to help", an alternative strategy could be to introduce wage-supported jobs where productivity is accepted to be below the normal minimum wage level. Future research should explore this promising avenue.

[26]Because the average treatment of the treated likely is higher than the marginal treatment effect. Increasing compulsory schooling is an ongoing debate in many countries, e.g., in the UK and the US following results by Clark & Royer (2013) and Oeropoulos (2006).

# 9 References

Aizer, A., S. Eli, J. Ferrie & A. Lleras-Muney (2016). The Long-Run Impact of Cash Trnasfers to Poor Families. *American Economic Review*, **106**(4): 935-971.

Almond, D., J., Currie & V. Duque (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature*, **56**(4): 1360-1446.

Arcidiacono, P., & R. A. Miller (2011): Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica*, **79**(6): 1823-1867.

Belfield, C. R. & H. M. Levin (2007). The Education Attainment Gap: Who's Affected, How Much, and Why It Matters. Ch. 1 in Levin, H. M. & C. R. Belfield (eds.) *The Price We Pay.* Brookings Institution Press, Washington, D.C.

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, & J. E. Humphries (2016), What Do Grades and Achievement Tests Measure?, *Proceedings of the National Academy of Sciences*, **113**(47): 13554-13559.

Bowers, A. J., R. Sprott & S. Taff (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity and specificity. *High School Journal*, **96**(2): 77–100.

Bowles, S. & H. Gintis (1976). Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books.

Cameron, S. V. & J.J. Heckman (1998). Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy,* **106**(2): 262-333.

Carneiro, P. & J. J. Heckman (2003). Human Capital Policy, in J. J. Heckman, A. B. Krueger and B. M. Friedman, eds., *Inequality in America: What Role for Human Capital Policies?*, MIT Press, Cambridge, MA, 77–239.

Clark, D.& H. Royer (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*, **103**(6): 2087-2120.

Cunha, F. & J.J. Heckman (2009). Human Capital Formation in Childhood and Adolescence. *CESifo DICE Report*, **07**(4): 22-28.

Cunha, F., J. J. Heckman & S. M. Schennach (2010). Estimating the Technology of Cognitive and Non-cognitive Skill Formation. *Econometrica*, **78**(3): 883-931.

Cunha, F., J. J. Heckman, L. J. Lochner & D. Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation, in E.Hanushek and F. Welch, eds., *The Handbook of Economics of Education*, 697–812, North Holland, Amsterdam.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, **132** (4): 1593–1640.

Eckstein, Z. & K. E. Wolpin (1999). Why Youth Drop Out of High Scool: The Impact of Preferences, Opportunities, and Abilities. *Econometrica*, **67**(6): 1295-1339.

EU (2010): Early School Leaving. Lessons from research for policy makers. European Commission, 2010.

Federation of Danish Industries (DI) (2016). Potentiale for flere private investeringer (In Danish). https://di.dk/.

Freeman, J. & B. Simonsen (2015). Examining the Impact of Policy and Practice Interventions on High School Dropout and School Completion Rates: A Systematic Review of the Literature. *Review of Educational Research*, **85**(2): 205-248.

Geweke, J. & M. Keane (2000). Bayesian inference for dynamic discrete choice models without the need for dynamic programming. In *Simulation-Based Inference in Econometrics: Methods and Applications,* R. S. Mariano, T. Schuermann, and M. Weeks, (eds.): 100–131, Cambridge University Press, Cambridge.

Hanushek, E. A. & L. Woessmann (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, **46**(3): 607–68.

Heckman, J.J. (2007). The economics, technology, and neuroscience of human capital formation. *PNAS*, **104**(33): 13250-13255.

Heckman, J.J. & B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, **52**: 271-320.

Heckman, J. J., N. Hohmann, J. Smith & M. Khoo (2000). 'Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment. *Quarterly Journal of Economics*, **115:** 651–694.

Heckman, J. J., J. Stixrud & S. Urzua (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, **24**(3): 411–82.

Herrnstein, R. J. & C. Murray (1994). The Bell Curve: Intelligence and Class Structure in American Life. New York: Free Press.

Holmlund, H., M. Lindahl & E. Plug (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, **49**(3): 615-651.

Jørgensen, T. & M. Tô (2018). Robust Estimation of Finitite Horizon Economic Models, *Computational Economics*, forthcoming.

Keane, M. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics*, **156**: 3–20.

Keane, M. & K. L. Wolpin (1997). The career decisions of young men. *Journal of Political Economy*, **105**(3): 473–522.

Kline, P. & C. R. Walters (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start. *Quarterly Journal of Economics*, **131** (November 2016): 1795-1848.

Landersø, R. & J. J. Heckman (2017). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US. *The Scandinavian Journal of Economics*, **119**(1): 178-230..

LaLonde, R. J. (2003). Employment and Training Programs. In *Means-*

*Tested Transfer Programs in the United States*, (ed.) Robert A. Moffit: 517-585. Chicago: University of Chicago Press.

Levin, H. M. & C. R. Belfield (2007). Educational Interventions to Raise High School Graduation Rates. Ch. 9 in Levin, H. M. & C. R. Belfield (eds.) *The Price We Pay.* Brookings Institution Press, Washington, D.C.

Lindqvist, E. & R. Vestman (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, **3**(January 2011):101–128.

McCall, B., J.A. Smith & C. Wunsch (2016). Government-Sponsored Vocational Education for Adults. Chapter 9, *Handbook of the Economics of Education*, Vol. **5**: 479-652.

McKinsey (2017). A Future That Works: Automation, Employment, and Productivity. McKinsey Global Institute, January 2017.

OECD (2016). The Risk of Automation for Jobs in OECD Countries. A Compartaive Analysis. OECD Social, Employment and Migration Working Paper No. 189. OECD Paris.

OECD (2017), Youth unemployment rate (indicator). doi: 10.1787/c3634df7-en

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. *American Economic Review*, **96**(1): 152-175.

Rangvid, B. S., V. M. Jensen & S. S. Nielsen (2015). Forberedende tilbud og overgang til ungdomsuddannelse. *SFI report* 15:4 (in Danish).

Rouse, Cecilia E. (2007). Consequences for the Labor Market. Ch. 5 in Levin, H. M. & C. R. Belfield (eds.) *The Price We Pay.* Brookings Institution Press, Washington, D.C.

Rust, J. (1994). Structural Estimation of Markov Decision Processes. In *Handbook of Econometrics*, Volume 4, edited by Robert F. Engle and Daniel L. McFadden, 3081–3143. Amsterdam and San Diego: Elsevier, North-Holland.

Rust, J. (2014). The Limits of Inference *with* Theory: A Review of Wolpin (2013). *Journal of Economic Literature*, **52**(3): 820–850.

Schochet, P. Z., J. Burghardt, & S. McConnell (2008). Does Job Corps Work? Impact Fnidings from the National Job Corps Study. *American Economic Review*, **98**(5): 1864-1886.

Train, Kenneth (2008). EM Algorithms for nonparametric estimation of mixing distributions, *Journal of Choice Modelling*, **1**(1): 40-69.

Willis, R. J. & S. Rosen (1979). Education and Self-Selection. *Journal of Political Economy*, **87**(5), pt.2: s7-s36.

Wolpin, K.I. (1992): The Determinants of Black-White Differences in Early Employment Careers: Search, Quits, Layoffs and Endogenous Wage Growth. *Journal of Political Economy*, **100**(3): 535-560.

## Appendix 1: Detailed Description of Data, Sample and Variables

Table 12: Control variables, means and standard deviations

| Variable | Mean | Std. Dev. |
|---|---|---|
| HH annual income (USD, OECD equivalents) | 39,050 | 4.8 |
| | | |
| Father's education | | |
| Compulsory only | 34.3 | 47.5 |
| High scool | 2.3 | 15.1 |
| VET completed/Skilled | 43.6 | 49.6 |
| College degree or more | 10.4 | 30.5 |
| Missing information | 9.4 | 29.2 |
| | | |
| Non-western Immigrant | 12.4 | 33.0 |
| GPA ninth grade | 2.7 | 0.9 |
| Living in a province (outside Cph.) | 90.3 | 29.6 |
| 1-year change in private investment level (pct) | -0.9 | 10.3 |
| Local uneployment of low skilled | 8.9 | 3.5 |
| | | |
| Number of observations | 95,625 | |
| Number of persons | 3,825 | |

Note: The change in private investments enters the model. Source: Federation of Danish Industries (2016).

## Appendix 2: Description of Preperatory Courses

Preperatory courses are targeted adolescents who are incapable of starting an education. The main offer given here is so-called Production Schools, which is an educational offer based on practical work and education in workshops. Usually, there will be 10-12 students in a team, and the branches mostly cover industries such as wood, metal, housing service, and kitchen/cooking. In some Prodcution Schools there is also access to pedagogy,, social, and health care service, while other schools have theater, music or design.

Teachers have usually completed VET with a qualifying degree as a craftsman and/or a pedagogical background. The teaching can also include so-called "preperatory adult education", which can include courses at rather basic levels (even below ninth grade if needed). Production School students are seen as individuals with individual needs, and continous, individual mentoring, guidance, and supervison is part of the package.

## Appendix 3: Law of Motion

The laws of motion for the relevant state variables are:

$$\begin{aligned}
S_{it} =& d_{it}^{s} + S_{i,t-1} \\
P =& d_{it}^{p} + P_{i,t-1} \\
VET_{it} =& d_{it}^{vet} + VET_{i,t-1} \\
E_{it} =& d_{it}^{e} + E_{i,t-1} \\
NEET_{it} =& d_{it}^{neet} + NEET_{i,t-1} \\
VET_{it}^{C} =& \Delta_{it} d_{it}^{vet} + VET_{i,t-1}^{C}
\end{aligned}$$

with initial conditions $S_{i,0} = PREP_{i,0} = VET_{i,0} = E_{i,0} = NEET_{i,0} = VET_{i,0}^{C} = 0$. $\Delta_{it}$ equals one if the individual completes $VET$ in period $t$. We model $Pr\left(\Delta_{it} = 1\right)$ as

$$\begin{aligned}
Pr\left(\Delta_{it} = 1\right) =& \frac{exp\left(\delta_i + \delta_1 DV46_{it} + \delta_2 DV7_{i,t} + \delta_3 Prep + \delta_4 UE\, unskilled + \delta_5 GPA\right)}{1 + exp\left(\delta_i + \delta_1 DV46_{it} + \delta_2 DV7_{i,t} + \delta_3 Prep + \delta_4 UE\, unskilled + \delta_5 GPA\right)}\ \forall t \leq t^{C} \\
=& 1\ \forall t > t^{C}
\end{aligned}$$

where $t^{C}$ is the time period in which the individual completes $VET$. $DV46_{it}$ is one if the person has completed 4-6 periods of $VET$, and $DV7_{it}$

is one if the person has completed 7 or more periods of $VET$. Further, $Prep$ is an indicator = 1 if the individual has completed at least one period of preparatory schooling, $UE\,unskilled$ is a measure of the unemployment rate of unskilled workers, which varies over time and by region (as a proxy for outside options if the individual does not complete). GPA is the individual's grade point average in ninth grade.

# Appendix 4: Structural Parameter Estimates

Table 13: Structural parameter estimates

| Equation | Parameter | Estimate | StdErr | t-value |
|---|---|---|---|---|
| Education | a1 In school last period | 3.959 | 0.040 | 99.558 |
| | a2 Periods in school | 0.236 | 0.008 | 29.869 |
| | a3 Father VET educated | 0.136 | 0.031 | 4.348 |
| | a4 Father college educated | 0.577 | 0.035 | 16.692 |
| | a5 Father education missing | 0.186 | 0.036 | 5.157 |
| | a6 Cognitive | 0.134 | 0.020 | 6.731 |
| | a7 Province | -0.096 | 0.037 | -2.604 |
| | a8 Immigrant | -0.241 | 0.026 | -9.119 |
| | a9 Tenth grade dummy | 0.808 | 0.017 | 47.044 |
| | a10 Period1*a1 | -1.383 | 0.021 | -66.049 |
| | a11 Tenth dummy*a1 | 1.973 | 0.017 | 113.314 |
| | a12 Age 16-19 | -0.289 | 0.088 | -3.295 |
| | a13 Age 24-28 | 0.343 | 0.048 | 7.221 |
| | alpha S type 1 | 0.740 | 0.063 | 11.671 |
| | alpha S type 2 | 1.356 | 0.069 | 19.694 |
| | alpha S type 3 | 0.001 | 0.043 | 0.033 |
| | alpha S type 4 | 2.086 | 0.055 | 37.779 |
| | | | | |
| VET | v1 Father VET educated | 0.102 | 0.029 | 3.489 |
| | v2 Father college educated | 0.041 | 0.028 | 1.448 |
| | v3 Father education missing | 0.007 | 0.039 | 0.185 |
| | v4 Cognitive | 0.069 | 0.016 | 4.268 |
| | v5 Province | 0.178 | 0.038 | 4.670 |
| | v6 Preperatory attended | -0.907 | 0.029 | -31.803 |
| | v7 Immigrant | -0.629 | 0.031 | -20.252 |
| | v8 In VET last period | 2.264 | 0.025 | 90.247 |
| | v9 Delta invest | -0.004 | 0.001 | -2.958 |
| | v10 Age 16-19 | 0.940 | 0.043 | 21.785 |
| | v11 Age 24-28 | -0.052 | 0.044 | -1.198 |
| | v12 VET spaces last year | 1.944 | 0.094 | 20.675 |
| | alpha VET type 1 | 1.694 | 0.054 | 31.339 |
| | alpha VET type 2 | 3.163 | 0.055 | 57.527 |
| | alpha VET type 3 | 0.671 | 0.068 | 9.914 |
| | alpha VET type 4 | 3.857 | 0.049 | 78.205 |

Table 14: Structural parameter estimates, continued

| Equation | Parameter | Estimate | StdErr | t-value |
|---|---|---|---|---|
| Preparatory | p1 Father VET educated | -0.276 | 0.041 | -6.793 |
| | p2 Father college educated | 0.222 | 0.029 | 7.728 |
| | p3 Father education missing | -0.012 | 0.020 | -0.590 |
| | p4 Cognitive | 0.018 | 0.029 | 0.641 |
| | p5 Province | 0.175 | 0.065 | 2.686 |
| | p6 Immigrant | -0.420 | 0.037 | -11.297 |
| | p7 Municip offers | 3.623 | 0.022 | 168.090 |
| | p8 Age 16-19 | 2.787 | 0.055 | 50.606 |
| | p9 Age 24-28 | -0.384 | 0.044 | -8.712 |
| | p10 Age 16 - 19*municip offers | -0.251 | 0.014 | -18.186 |
| | p11 Delta invest | -0.033 | 0.003 | -11.289 |
| | alpha Prep type 1 | 0.154 | 0.045 | 3.448 |
| | alpha Prep type 2 | 0.744 | 0.049 | 15.169 |
| | alpha Prep type 3 | -0.410 | 0.075 | -5.480 |
| | alpha Prep type 4 | 1.061 | 0.050 | 21.190 |
| Work | e1 High School completed | 0.480 | 0.026 | 18.152 |
| | e2 VET completed | 1.636 | 0.048 | 33.763 |
| | e3 Preparatoty attended | -0.751 | 0.030 | -25.194 |
| | e4 Immigrant | -0.628 | 0.031 | -20.207 |
| | e5 In employment last period | 2.456 | 0.022 | 109.298 |
| | e6 Delta invest | 0.005 | 0.001 | 4.944 |
| | e7 Unemployment of low skilled | -4.935 | 0.020 | -246.214 |
| | e8 Age 16-19 | -0.275 | 0.040 | -6.861 |
| | e9 Age 24-28 | -0.358 | 0.026 | -13.865 |
| | alpha emp type 1 | -7.515 | 0.059 | -128.503 |
| | alpha emp type 2 | -6.495 | 0.060 | -108.311 |
| | alpha emp type 3 | -8.076 | 0.069 | -117.477 |
| | alpha emp type 4 | -6.116 | 0.057 | -106.850 |

Table 15: Structural parameter estimates, continued

| Equation | Parameter | Estimate | StdErr | t-value |
|---|---|---|---|---|
| Wage | W1 High school/college | 0.155 | 0.010 | 14.881 |
| | W2 Completion of VET | 0.308 | 0.007 | 45.671 |
| | W3 Experience | 0.060 | 0.001 | 46.104 |
| | W4 Experience-squared | -0.002 | 0.000 | -21.947 |
| | W5 Preperatory attended | -0.004 | 0.006 | -0.770 |
| | W6 Immigrant | -0.051 | 0.007 | -7.364 |
| | W7 Cognitive | 0.011 | 0.002 | 5.000 |
| | W8 Sigma | -1.133 | 0.004 | -297.954 |
| | alpha Wage type 1 | 11.543 | 0.007 | 1700.819 |
| | alpha Wage type 2 | 11.207 | 0.008 | 1344.273 |
| | alpha Wage type 3 | 11.155 | 0.011 | 1043.576 |
| | alpha Wage type 4 | 11.317 | 0.007 | 1533.581 |
| | | | | |
| Polynomial | Constant | -3.994 | 0.068 | -59.102 |
| | S | -0.015 | 0.012 | -1.178 |
| | P | 0.174 | 0.015 | 11.399 |
| | E | -0.018 | 0.011 | -1.613 |
| | VET | -3.835 | 0.048 | -80.029 |
| | $S^2$ | -0.006 | 0.001 | -7.813 |
| | $E^2$ | 0.003 | 0.001 | 4.764 |
| | S*P | -0.018 | 0.004 | -3.992 |
| | S*E | 0.009 | 0.002 | 5.642 |
| | S*VET | 0.243 | 0.013 | 19.422 |
| | P*E | 0.000 | 0.003 | 0.018 |
| | P*VET | 0.246 | 0.028 | 8.866 |
| | E*VET | 0.105 | 0.007 | 15.314 |

Table 16: Structural parameter estimates, continued

| Equation | Parameter | Estimate | StdErr | t-value |
|---|---|---|---|---|
| Prob(VET completed) | vc1 4-6 vet periods | 1.715 | 0.026 | 65.864 |
| | vc2 7+ vet periods | 6.460 | 0.033 | 195.704 |
| | vc3 prep background | -0.048 | 0.018 | -2.589 |
| | vc4 Unemployment level low skilled (t-1) | 16.232 | 0.047 | 342.344 |
| | vc5 Cognitive | 0.046 | 0.015 | 2.992 |
| | alpha vc1 | -12.239 | 0.103 | -118.266 |
| | alpha vc2 | -1.928 | 0.024 | -81.176 |
| | alpha vc3 | -7.989 | 0.073 | -109.948 |
| | alpha vc4 | -6.366 | 0.033 | -192.802 |

Table 17: Structural parameter estimates, continued (types)

| Equation | Parameter | Estimate | StdErr | t-value |
|---|---|---|---|---|
| Types | Type 1 const | -1.086 | 0.022 | -49.131 |
| | Type 1 Orderliness | -0.152 | 0.023 | -6.683 |
| | Type 1 lowHHinc | 0.431 | 0.020 | 21.189 |
| | Type 1 Immigrant | 1.115 | 0.040 | 28.104 |
| | Type 2 const | -1.795 | 0.020 | -89.892 |
| | Type 2 Orderliness | -0.191 | 0.026 | -7.434 |
| | Type 2 lowHHinc | 0.339 | 0.018 | 18.773 |
| | Type 2 Immigrant | -0.214 | 0.007 | -28.898 |
| | Type 3 const | -2.009 | 0.093 | -21.574 |
| | Type 3 Orderliness | -0.513 | 0.049 | -10.453 |
| | Type 3 lowHHinc | 0.699 | 0.051 | 13.779 |
| | Type 3 Immigrant | -392.607 | 3.506 | -111.986 |
| | | | | |
| | Type 1 | 28 pct. | | |
| | Type 2 | 11 pct. | | |
| | Type 3 | 14 pct. | | |
| | Type 4 | 47 pct. | | |

The name ”Orderliness” refers to Orderliness/organization/neatness.

Table 18: Transition matrix, Type 1

| | Choice (t) | | | | | |
|---|---|---|---|---|---|---|
| Choice (t-1) | School | Preparatory | VET | Work | NEET | Total |
| School | 73.3 | 3.3 | 6.9 | 4.4 | 12.2 | 100.0 |
| | 85.5 | 12.0 | 4.8 | 1.8 | 6.7 | 14.7 |
| Preparatory | 3.1 | 20.0 | 15.3 | 13.4 | 48.1 | 100.0 |
| | 1.1 | 21.2 | 3.1 | 1.6 | 7.6 | 4.2 |
| VET | 1.2 | 3.0 | 68.3 | 8.3 | 19.2 | 100.0 |
| | 2.0 | 15.6 | 68.1 | 4.8 | 15.0 | 20.8 |
| Work | 1.0 | 1.4 | 3.8 | 81.7 | 12.0 | 100.0 |
| | 2.8 | 12.1 | 6.2 | 77.7 | 15.4 | 34.1 |
| NEET | 4.1 | 5.9 | 14.3 | 19.4 | 56.3 | 100.0 |
| | 8.6 | 39.0 | 17.9 | 14.2 | 55.3 | 26.2 |
| Total | 12.6 | 4.0 | 20.9 | 35.9 | 26.6 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

# Appendix 5: Additional Tables and Figures

Table 19: Transition matrix, Type 2

| | Choice (t) | | | | | |
|---|---|---|---|---|---|---|
| Choice (t-1) | School | Preparatory | VET | Work | NEET | Total |
| School | 66.3 | 4.0 | 18.3 | 4.5 | 6.9 | 100.0 |
| | 89.3 | 21.9 | 7.1 | 1.1 | 2.1 | 9.0 |
| Preparatory | 3.4 | 16.9 | 34.8 | 13.5 | 31.5 | 100.0 |
| | 0.9 | 18.8 | 2.7 | 0.6 | 1.9 | 1.8 |
| VET | 0.3 | 1.8 | 76.5 | 4.2 | 17.2 | 100.0 |
| | 1.2 | 26.3 | 79.2 | 2.6 | 14.2 | 24.1 |
| Work | 0.6 | 0.5 | 3.2 | 83.8 | 11.8 | 100.0 |
| | 3.2 | 11.9 | 5.0 | 78.5 | 14.7 | 36.5 |
| NEET | 1.3 | 1.2 | 4.8 | 23.7 | 69.0 | 100.0 |
| | 5.3 | 21.3 | 5.9 | 17.3 | 67.0 | 28.5 |
| Total | 6.7 | 1.6 | 23.3 | 39.1 | 29.3 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Table 20: Transition matrix, Type 3

| | Choice (t) | | | | | |
|---|---|---|---|---|---|---|
| Choice (t-1) | School | Preparatory | VET | Work | NEET | Total |
| School | 66.0 | 4.3 | 4.8 | 2.9 | 22.0 | 100.0 |
| | 80.4 | 9.4 | 4.1 | 1.5 | 4.7 | 11.1 |
| Preparatory | 1.6 | 18.6 | 8.2 | 7.9 | 63.8 | 100.0 |
| | 1.0 | 19.3 | 3.4 | 1.9 | 6.6 | 5.3 |
| VET | 1.1 | 4.0 | 54.5 | 6.9 | 33.5 | 100.0 |
| | 1.6 | 10.1 | 54.4 | 4.1 | 8.3 | 12.8 |
| Work | 1.3 | 1.5 | 3.6 | 68.2 | 25.4 | 100.0 |
| | 2.9 | 6.2 | 5.9 | 65.7 | 10.2 | 20.7 |
| NEET | 2.6 | 5.6 | 8.3 | 11.5 | 72.0 | 100.0 |
| | 14.3 | 55.0 | 32.3 | 26.8 | 70.1 | 50.1 |
| Total | 9.1 | 5.1 | 12.8 | 21.5 | 51.5 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Table 21: Transition matrix, Type 4

| | Choice (t) | | | | | |
|---|---|---|---|---|---|---|
| Choice (t-1) | School | Preparatory | VET | Work | NEET | Total |
| School | 68.2 | 3.3 | 19.6 | 5.0 | 3.9 | 100.0 |
| | 91.5 | 26.3 | 7.4 | 1.1 | 2.1 | 10.3 |
| Preparatory | 2.7 | 13.7 | 44.5 | 16.1 | 23.0 | 100.0 |
| | 0.5 | 15.6 | 2.4 | 0.5 | 1.8 | 1.5 |
| VET | 0.6 | 1.4 | 80.5 | 4.4 | 13.2 | 100.0 |
| | 2.0 | 29.7 | 82.7 | 2.7 | 19.3 | 27.9 |
| Work | 0.5 | 0.4 | 2.8 | 88.5 | 7.9 | 100.0 |
| | 2.9 | 11.6 | 4.2 | 82.2 | 17.1 | 41.6 |
| NEET | 1.3 | 1.1 | 4.8 | 32.0 | 60.8 | 100.0 |
| | 3.1 | 16.8 | 3.3 | 13.4 | 59.8 | 18.8 |
| Total | 7.6 | 1.3 | 27.1 | 44.8 | 19.1 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |