NBER WORKING PAPER SERIES

MISMATCH IN HUMAN CAPITAL ACCUMULATION

Russell Cooper
Huacong Liu

Working Paper 22010
http://www.nber.org/papers/w22010

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

MisMatch in Human Capital Accumulation
Russell Cooper and Huacong Liu
NBER Working Paper No. 22010
February 2016
JEL No. E24,I26,J24,O43

ABSTRACT

This paper studies the allocation of heterogeneous agents to levels of educational attainment. The goal is to understand the magnitudes and sources of mismatch in this assignment, both in theory and in the data. The paper presents evidence of substantial mismatch between ability and educational attainment across 21 OECD countries, with a focus on Germany, Italy, Japan and the US. In the model, mismatch originates from: (i) taste shocks, (ii) binding borrowing constraints and (iii) noisy measures of ability in test scores. The model is estimated using a simulated method of moments approach. The main finding is that measured mismatch arises largely from noise in test scores and does not reflect borrowing constraints. Differences in tastes for education across households play a minor role in explaining mismatch. Further, the estimation allows us to decompose the college wage premium, isolating cross-country differences in selection effects from the return to education.

Russell Cooper
Department of Economics
The Pennsylvania State University
611 Kern
State College, PA 16802
and NBER
russellcoop@gmail.com

Huacong Liu
College of Education
The Pennsylvania State University
State College, PA 16801
hxl39@psu.edu

MisMatch in Human Capital Accumulation*

Russell Cooper† and Huacong Liu‡

February 9, 2016

Abstract

This paper studies the allocation of heterogeneous agents to levels of educational attainment. The goal is to understand the magnitudes and sources of mismatch in this assignment, both in theory and in the data. The paper presents evidence of substantial mismatch between ability and educational attainment across 21 OECD countries, with a focus on Germany, Italy, Japan and the US. In the model, mismatch originates from: (i) taste shocks, (ii) binding borrowing constraints and (iii) noisy measures of ability in test scores. The model is estimated using a simulated method of moments approach. The main finding is that measured mismatch arises largely from noise in test scores and does not reflect borrowing constraints. Differences in tastes for education across households play a minor role in explaining mismatch. Further, the estimation allows us to decompose the college wage premium, isolating cross-country differences in selection effects from the return to education.

JEL classification: I26, J24

1 Introduction

This paper studies the allocation of heterogeneous agents to levels of educational attainment. Observed outcomes are often at odds with the stark predictions of assortative matching: i.e. mismatch occurs whereby high ability agents are not always the most educated and some low ability agents have high educational attainment. Our primary goal is to understand the magnitudes and sources of this mismatch, both in theory and in the data.

The paper presents and analyzes cross-country OECD data on mismatch. The measure of education attainment is dichotomous: (i) below college and (ii) college and above. PIACC scores, an OECD sponsored assessment of adult skills, are used in our analysis as noisy measures of ability for each individual by country.[1] The use of these data is key to facilitating a cross-country comparison of the relationship between ability and education attainment.[2]

*Comments and suggestions from Jonathan Eaton, Martin Hackmann, Eric Hanushek, Marc Henry, Immo Schott, Jon Willis and Guozhong Zhu as well as seminar participants at the Federal Reserve Bank of Kansas City, The University of Montreal, the Pennsylvania State University and the University of Alberta are greatly appreciated.

†Department of Economics, the Pennsylvania State University and NBER, russellcoop@gmail.com

‡College of Education, the Pennsylvania State University, hxl39@psu.edu

[1]See `http://www.oecd.org/site/piaac/` for a complete description of this “survey of adult skills”. The use of this test in our analysis as a proxy for ability is explained in detail below.

[2]Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) use the PIACC score as a measure of cognitive skills in Mincer wage regressions. It is clear from that analysis that the PIACC score is highly correlated with labor market outcomes, it is not simply noise. The PIACC score is significant in predicting wages even when schooling is included. We use these results as moments in our estimation. Section 5 of Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) discusses causal interpretations, particularly the reverse causality whereby individuals with particular high skilled jobs, say obtained as the outcome of a training program, consequently score higher on the PIACC test.

The empirical analysis starts with the relationship between PIACC scores and educational attainment across countries. Not surprisingly, the distributions of these scores conditional on educational attainment overlap: there are individuals with a low level of education (no college) but a higher PIACC score than those with high education (college). This is a simple, but informative, indicator of mismatch.

Our formal analysis of mismatch estimates the probability an individual will obtain higher education given an observed PIACC score.[3] Using these estimates, "under-matching" occurs if the predicted probability of a college education is relatively high but the agent does not have a college degree. In a similar manner, "over-matching" occurs for individuals with a college degree but a relatively low predicted probability of attending college. These estimates are obtained using country specific regressions.

The theoretical framework focuses on the assignment of individuals to education levels. It allows individuals to differ in a number of dimensions: (i) ability, (ii) tastes and (iii) wealth. If the only source of heterogeneity is ability, then the optimal allocation will assign higher ability agents to higher levels of education. There is no mismatch. Once differences in tastes are present, the optimal allocation assigns education attainment based on both ability and tastes so that some high ability agents will attain relatively low levels of education. The methodology described above would indicate mismatch, though the allocation may still be efficient. Differences in wealth are relevant to the assignment process in a decentralized setting with borrowing restrictions. In this case, relatively high ability agents may choose a low level of education simply because of a binding borrowing constraint.

In the model, there is another source of mismatch associated with ability being measured rather than observed. Individuals make education decisions based upon their true ability. Test scores, such as the PIACC assessment, are an imperfect indicator of ability. Hence, some agents may appear to be high ability based upon test outcomes though they choose low education based upon their true, relatively low, ability. This form of mismatch reflects noise in the measure of ability. As we shall see, this is an important source of measured mismatch in the data.

The analysis uses this theoretical framework to identify the sources of measured under- and over-matching. To do so, the country-specific parameters of the individual choice problems are estimated using a simulated method of moments approach.[4] The degrees of over- and under-matching, the mean education rate, the coefficients from the logistic regression used to predict education outcomes and coefficients relating wages to PIACC scores are computed for each of the countries.[5] These moments are used as a basis for the estimation of model parameters. The estimation allows us to determine the source of mismatch across countries.

In this project, the emphasis is on the choice of education based on ability rather than the matching of workers by skill to appropriate jobs. This complements the study of mismatch in labor markets. To the extent high ability individuals have low educational attainment and thus low skill jobs, these forms of mismatch are related.[6]

There are four main findings in this study. First, there is evidence of substantial mismatch in our sample, including both over-matching and under-matching. Countries with high education rates tend to have low under-match and high over-match rates.

[3] This follows Dillon and Smith (2013), Smith, Pender, and Howell (2013) and others.

[4] The parameters estimated include the borrowing constraint of the household, the distributions of ability, taste shocks and noise in the test score as well as the returns to education.

[5] Some of these moments are taken from Hanushek, Schwerdt, Wiederhold, and Woessmann (2015).

[6] An example is the famous taxi driver in Singapore with a PhD `https://en.wikipedia.org/wiki/Cai_Mingjie`. Is he under-matched in his job or over-matched in education? Section 8.3 returns to this theme and discusses the contribution of education mismatch to apparent job mismatch.

Second, by country, mismatch reflects noise in the test scores and is not due to imperfect capital markets nor to variations in tastes for education across agents. The estimation of the model finds no support for the presence of binding borrowing constraints. Further, taste shocks contribute essentially nothing to the fit of the model. Instead the noise in the test score is enough to generate the observed mismatch in a manner that is consistent with the estimated dependence of the education decision and compensation on the test score. Matching these latter moments in the estimation critically disciplines the explanatory power of the noise in the test score. From this over-identification, matching these moments from noise in the test score is non-trivial.

Third, mismatch is not a signal of inefficiency. Our simple model, relying solely on a noisy test score, does a remarkable job of capturing cross country variations in education rates, mismatch and wage premia. It does so by estimating differences in the distribution of ability, the noise in test scores and the return to higher education across countries. None of these sources of variation signal an inefficiency in the allocation of individuals to education attainment.

Finally, the estimated model facilitates the decomposition of the college wage premium into two sources: the returns to college (relative to no college) and the selection by ability into college. There are substantial cross country differences in the returns to education even though wage premia are similar. This reflects differences in selection into higher education.

2 Motivation

The paper is motivated by evidence of mismatch, i.e. the stark difference between the predictions of the sorting model of education attainment and the data. This section presents an initial model and an initial look at the data to make this inconsistency clear. The remainder of the paper uses a richer model to understand the sources and consequences of the mismatch through a simulated method of moments approach.

2.1 A Framework

This section presents a simple education choice.[7] It provides a benchmark for considering the evidence relating education attainment to measured ability. The framework is enriched as the analysis progresses to become the basis of the structural estimation.

Consider an economy with multiple agents, who differ in terms of their ability, denoted θ, with a cdf $G(\theta)$.[8] The lifetime utility of household θ is given by $u(c(\theta))$ where $c(\theta)$ is the consumption of a household with ability θ, $u(\cdot)$ is strictly increasing and strictly concave.

Each agent has a unit of time which is allocated to work and education. The resource constraint for the economy is given by:

$$\int_{\theta}[c(\theta)+pe(\theta)]dG(\theta)=\int_{\theta}[(1-e(\theta))+h(e(\theta))\theta]dG(\theta). \tag{1}$$

Here $e(\theta)$ is the time allocated to education. The left side is the use of output to finance consumption and education,

[7] Models of sorting such as this appear throughout the literature. See Spence (1973) and Weiss (1983) for early examples of sorting in equilibrium models of human capital accumulation and signaling. A key assumption is those models, retained here, is that education choice depends on actual ability.

[8] For this discussion, agents are indexed by ability.

with a resource costs of p per unit of time spent in school. The right side is total output, comprised of the output from unskilled work time, $1-e(\theta)$, and the type-specific return to education, $h(e(\theta))\theta$. The human capital accumulation function, $h(e)$, is assumed to be strictly increasing and strictly concave. The ability of the agent is complementary to time spent in school.

The planner chooses consumption allocations and education levels for all types $(c(\theta), e(\theta))$ to maximize social welfare of $\int_\theta [\Lambda(\theta)u(c(\theta))]dG(\theta)$ subject to the resource constraint, (1). In this expression of social welfare, $\Lambda(\theta)$ is a welfare weight.

The education decision, for each type θ, is characterized by:

$$p + 1 = \theta h'(e(\theta)). \tag{2}$$

The left side is the marginal cost of education in period 1 and the right side is the marginal return to education for ability θ. The optimal level of education is increasing in θ from the strict concavity of $h(\cdot)$. Note that the efficient allocation of time between work and education is independent of the welfare weight given to the type of an agent.

As for the consumption allocation, the necessary condition is:

$$\Lambda(\theta)u'(c(\theta)) = \lambda \tag{3}$$

for all θ, where λ is the multiplier on (1). This condition equates the weighted marginal utility of consumption across agents. It captures the optimal redistribution of output in the economy. In this economy, the assignment of households to education is independent of the allocation of total output.

If the education choice was discrete, say $e \in \{0, \bar{e}\}$, then the solution of the planner's problem is to set $e = 0$ for agents with $\theta < \theta^*$ and $e = \bar{e}$ for agents with $\theta > \theta^*$. The critical level of ability, θ^* solves

$$(p + 1)\bar{e} = \theta^* h(\bar{e}). \tag{4}$$

2.2 An Initial Look at the Data

The model makes a stark prediction about sorting: agents with higher ability obtain higher levels of education. This section analyzes that prediction across countries. It requires a measure of ability and education attainment.

2.2.1 Data

The primary data source for this study is PIAAC, also called the Survey of Adult Skills. PIAAC assesses the proficiency of adults aged 16-65 in literacy, numeracy, and problem solving in technology-rich environments. It is an ongoing data collection effort at the OECD, with 22 countries participating in the first round of data collection that took place between 2008 and 2012 in most participating countries. These countries include Austria, Flanders (Belgium), Canada, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, Russian Federation, Slovak Republic, Spain, Sweden, England and N. Ireland (United Kingdom), and the United States. This first round of PIAAC data include a total of 152,514 individuals between 16 and 65 years of age from these 22 countries, with a majority of countries having 4,000 to 9,000 participants.

For this analysis, we included participants aged 25-39 who have finished the initial cycle of formal schooling. Given that our main interest is the mismatch between ability and schooling choice, the PIAAC numeracy score serves as a signal of cognitive ability since it is highly correlated with literacy and problem-solving skills. In the analysis, we use standardized numeracy scores in each country.

There are a couple of concerns with the use of the PIACC score as a measure of ability. First, as with all tests, the results are signals not direct measures of ability. Second, and more importantly for the PIACC score, the exam is given during working years as a measure of adult skills. Thus the exam reflects not only innate ability but also acquired skills from work experience and training. These concerns are confronted in our estimation by adding noise to the PIACC score and controlling, as best as possible, for the effects of experience.

Our study uses two samples of countries. We present evidence on mismatch for 21 OECD countries.[9] For these countries, we characterize the magnitude of mismatch and estimate our structural model.

We then go into considerable detail on the evidence for four countries: Germany, Italy, Japan and the US. As explained below, these countries stand out from the sample and provide particular insights into mismatch. For these countries, we discuss aspects of institutional structures that underlie the observed assignment of agents to education attainment.

For education attainment, we specify a dichotomous variable indicating two levels: (i) no college degree and (ii) college degree and beyond. The PIACC data does not contain any indicators of college quality so that a finer breakdown of education attainment is not feasible. Accordingly our focus is on the choice between a college degree or not rather than the ordering of agents relative to the quality of their college education. We rely on the International Standard Classification of Education (ISCED) to identify individuals who have obtained college and/or beyond degrees (ISCED 5 and above) and those whose highest educational attainment is below college (ISCED 1 through 4).[10]

2.2.2 Attainment and Ability

Figure 1 shows the distribution of PIACC scores by education attainment for Germany, Italy, Japan and the US. Table 1 reports moments from these distributions. In viewing these results, it is important to keep in mind that the same PIACC exam was given in each country, though translated into the local language.[11] A couple of patterns are clear.

First, for each of the countries, the distribution of PIACC scores for those with college degrees appears to be a rightward shift of the scores for the low education group. This difference in means between "no college" and "college" is clear from Table 1. These differences are statistically significant.

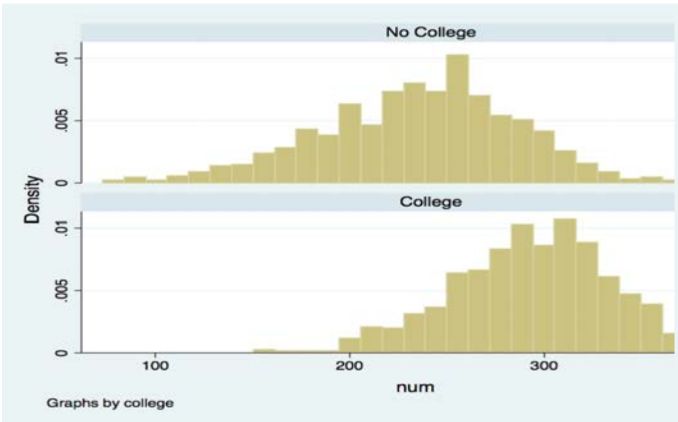
Second, for each of these countries, there is a wide dispersion of PIACC numeracy scores for each of the education attainment levels. Distributions of ability conditional on education attainment are far from degenerate. Except for Germany, the dispersion is lower for the college educated group. The dispersion is particularly low for those with college attainment in Japan.

[9] Russia was dropped due to measurement problems.

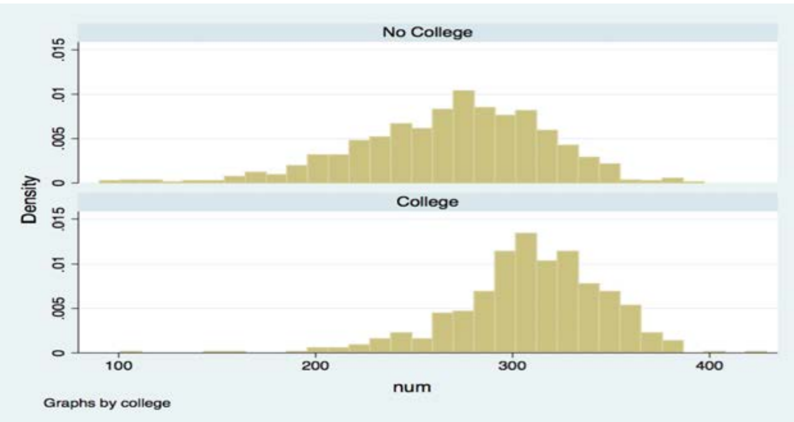
[10] Importantly, ISCED 5B is included as college for this analysis as seems customary.

[11] To be specific, the same set of questions was presented at each interview but the actual set of questions answered was determined through the interview process.

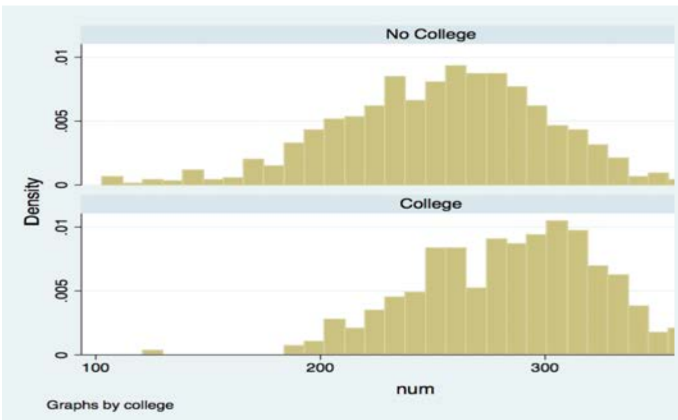
Figure 1: Distributions of PIACC Numeracy Scores



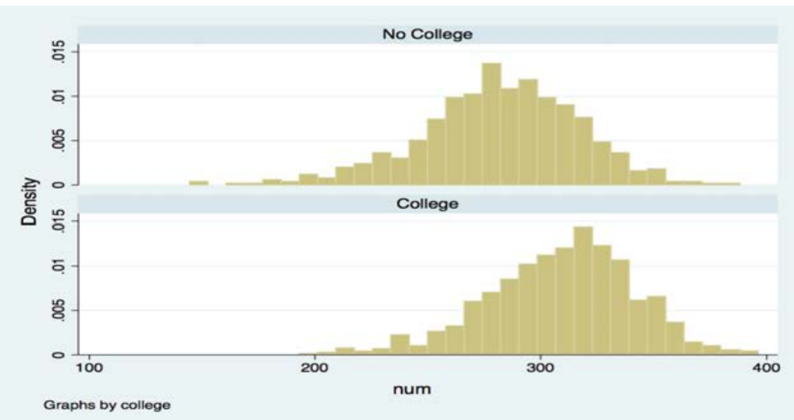
(a) US



(b) Germany



(c) Italy



(d) Japan

These figures show the distribution of PIACC numeracy scores by education (row) by country (columns). For each country, the first row is less than college and the second row is college and beyond.

Third, and most importantly for the purposes of our study, there is considerable overlap in these distributions. That is, many agents with low education attainment have scores that exceed those with high education attainment. This is true for each of these and the large sample of 21 countries.[12]

It is instructive to compare these figures against the model prediction. The model predicts the perfect sorting of agents by ability maps into education attainment. It is never the case that the education attainment of an agent with a given ability is strictly less than the attainment of an agent with a lower ability.

This property of perfect sorting by ability is clearly rejected by the data. This is the essence of mismatch: some lower ability agents obtain higher education and some high ability individuals do not. The next subsection goes on to better understand the determinants of this mismatch.

[12]Section 3.2 presents evidence for all 21 countries, measuring the levels of under- and over-matching.

Table 1: PIACC Numeracy Scores

	no college		college	
	mean	std.	mean	std.
Germany	246.80	50.39	309.48	37.58
Italy	253.60	45.61	284.31	40.12
Japan	281.99	35.52	309.11	32.21
US	235.55	50.71	293.71	40.40

This table reports the moments of the distribution of test scores by country and education attainment.

3 Evidence of MisMatch

This initial look at the data indicates a significant amount of mismatch. The next step of the empirical analysis is more formal. We study the empirical determinants of mismatch, separating over- and under-matching. To do so, we estimate state contingent probabilities of obtaining high education. Over-matching is defined by an agent obtaining a high level of education when the predicted probability of doing so is sufficiently low. Similarly, under-matching occurs if an agent does not obtain a high level of education despite a sufficiently high prediction probability of going to college.[13] A key part of the analysis is to determine, by country, the magnitude and sources of over-matching and under-matching.

Overall, this section is intended to answer two questions. How much mismatch is there in our sample? What are the magnitudes and empirical determinants of under- and over-matching? Using these results, we return to the theory models to infer the underlying sources of mismatch.

3.1 Under- and Over-Matching

This section goes beyond the unconditional distributions provided by Figure 1 to condition individual choices on individual attributes. In this manner, we generate empirical measures of mismatch, following the methodology of Dillon and Smith (2013) and Smith, Pender, and Howell (2013).

Specifically, consider a logistic model of education choice:

$$Pr(e_i = 1) = \frac{e^{\alpha_0 + \alpha_1 a_i}}{1 + e^{\alpha_0 + \alpha_1 a_i}} \tag{5}$$

where a_i is the PIACC score, treated as a proxy for ability, of individual i. The PIACC data reports ten plausible values for the numeracy score for each individual. The regression uses the mean of these plausible values as a proxy for ability. These regressions are run at the individual level by country generating country specific estimates of these parameters and ultimately rates of mismatch. Here $e_i = 0$ signifies that an individual has no college degree and $e_i = 1$ signifies college attainment and beyond.

This regression *per se* does not impose a direct interpretation of the coefficients (α_0, α_1). The structural estimation, based upon indirect inference, provides a framework for understanding (α_0, α_1) as they are used as moments.

The predicted values from these logistic regressions are used to obtain measures of under- and over-matching. In

[13]Throughout the paper, the terms under-match and over-match refer to the outcome of this empirical exercise and are not related to the efficiency of the assignment.

particular, a household is categorized as under-matched if: (i) the predicted probability that $e_i = 1$ exceeds the 80th percentile of all predicted values and (ii) the agent chooses $e_i = 0$. In a similar manner, a household is categorized as over-matched if: (i) the predicted probability that $e_i = 0$ is less than the 20th percentile of all predicted values and (ii) the agent chooses $e_i = 1$.

To be clear, at this point these cut-off values of 20th and 80th percentiles are arbitrary. The structural estimation provides an interpretation of this measure of under- and over-matching.[14]

Table 2: Moments

	college	under-match	over-match	α_0	α_1	N
Germany	0.373	0.104	0.062	−0.720 (0.07)	1.160 (0.09)	1440
Italy	0.230	0.146	0.069	−1.510 (0.08)	0.890 (0.09)	1381
Japan	0.597	0.078	0.108	0.230 (0.06)	0.860 (0.07)	1559
US	0.455	0.055	0.045	−0.360 (0.07)	1.510 (0.1)	1495

This table reports data moments including α_0 and α_1 from (5). Standard errors are provided for the logistic coefficient estimates. N is the sample size.

The results from these exercises are reported in Table 2. The first column reports college attainment rates by country. These are all advanced economies so that college attainment is relatively high, though it is noticeably lower in Germany and Italy than in Japan.

The second and third columns report the estimated under- and over-match rates. These are calculated as the ratio of the number of agents in an education group that is mismatched divided by the number of agents in that group.[15] The mismatch rates depend on the underlying regression, (5). The results for the logistic regression are shown in the fourth and fifth columns.[16]

The mismatch rate is highest in Germany and Italy and lowest in the US. There is some asymmetry in the mismatch rates: the under-match rate exceeds the over-match rate except for Japan. The asymmetry is particularly apparent in Italy. As we shall see, this difference will be important in assessing the role of capital market imperfections.

This discussion points to an important aspect of this methodology. Mismatch comes from large prediction errors which, in turn, depend on the specification of the model. One virtue of the structural estimation exercise, partly based upon indirect inference as discussed below, is that inference about the sources of mismatch are not coming directly from these regression results. So, for example, the argument that education and test scores are both influenced by unobserved ability is a valid criticism of any structural interpretation of α_1. In our model, this omitted variable bias is included in the structure since the agent's ability will impact both the education decision and test score. The estimates of the underlying parameters come indirectly from the reduced form coefficients in regressions such as (5).

[14]Section 6.4 explores the estimates with other cut-off values.

[15]For example, the under-match rate in the US is the ratio of the number of agents without a college degree ($e = 0$) and a predicted probability of $e_i = 1$ in excess of the 80th percentile divided by the number of agents without a college degree in the US.

[16]The model with additional controls, such as parent's education, is studied below.

3.2 Larger Sample

This analysis was performed on all 21 countries in our sample. The top panel of Table A2 contains the same moments reported in Table 2 for all of the countries.[17]

There is significant variation across countries. The college attainment rates range from a high of 64.8% in Korea to a low of 23% in Italy. For some countries, such as the U.S. the under-match rates is very low, only 5.5%. While for Italy it is 14.6%. There is also large variations in the over-match rate, from a low of 4% in the Czech Republic to a high of 10.9% in Korea. Across these countries, the estimate of α_1 is positive, indicating the correlation of the PIACC score with the education decision.

Countries with high education rates tend to have relatively low under-match rates and high over-match rates. These correlations across the 21 countries are -0.648 and 0.725 respectively. Evidently, higher education rates entail a reduction in under-matching and an increase in over-matching. We return to these patterns later through the lens of the estimated model.

4 Sources of Mismatch

This section studies models of the education decision structured to provide insight into the sources and consequences of mismatch. The initial framework from section 2.1 is reformulated as an intertemporal choice problem with heterogeneous agents and borrowing constraints. It is this model we ultimately take to the data.

The model has three stages of the lifecycle: (i) education, (ii) early employment and (iii) late employment. The framework is simplified by the assumption of stationary earnings within these stages. Nonetheless the structure allows us to study the potential impact of borrowing constraints during the education period as well as match with observations on the lifecycle pattern of labor earnings.

To study heterogeneity, we allow agents to differ along three dimensions. First, as in the baseline model, assume that the earnings of a college educated household are proportional to their human capital $h(\bar{e})\theta$. Here agents differ by their productivity, denoted θ, once they are educated. For agents with a higher value of θ, education is more productive.

Second, households have choice specific shocks which influence their education decision. These shocks are observed to households but not to the researcher.

Third, household can have different levels of wealth. This wealth can either come from an inheritance or stand for parental support. In the end, this third dimension of heterogeneity plays no role.

There are two sources of mismatch in the theory model: (i) shocks to the tastes of agents and (ii) borrowing restrictions. The next section uses these models to determine the sources of mismatch by country. The empirical model adds a third source of measured mismatch through a noisy test score. The test, as with the PIACC, is not linked to information about ability at the time of the education.

Importantly, the baseline model makes the informational assumption that is common to the literature linking ability and schooling: agents make their education decision knowing their true ability. This precludes mismatch arising from an uninformed education decision. College admissions based upon "luck" with a standardized exam are

[17] The coefficients (ν_1, ν_2) in Table A2 are explained below.

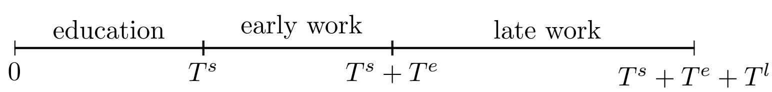
less likely to lead to a degree. Since our measure of attainment is college completion, the *ex ante* uncertainty about ability is reduced. But some educational systems, via an early tracking system, do force "early" college decisions. In Germany, for example, by age 10 individuals are sorted into multiple tracks, with only one leading to a college degree. We return to this variation in our framework in evaluating extensions of the baseline model in sub-section 7.1.

The model provides a structural interpretation of the estimates from (5): the test score is a proxy for ability which is an input into the education decision. The model imposes a restriction that education does not improve the test score.[18] Even with this restriction, the model fits the data quite well. We discuss the robustness of our results to this restriction in sub-section 7.2 where we extend the model to allow education to affect the test score.

4.1 LifeCycle

To study education choice and mismatch we consider a lifecycle model of the household. There are three phases of the lifecycle illustrated in Figure 2. In the first, termed education, phase, the household chooses an education level. This phase last T^s periods. The household works in the next two phases, termed early and late, of the lifecycle, lasting T^e and T^l periods respectively. The difference between these periods reflects experience: i.e. productivity increases with age, interacting with the education choice in the first stage. Let $T = T^s + T^e + T^l$ be the total lifetime of the household.

Figure 2: Phases of Lifecycle



This figure shows the phases over the lifecycle of the household.

4.2 Discrete Choice

As our empirical analysis focuses on two levels of education attainment, the model does as well. In the education phase the household chooses $e \in \{0, \bar{e}\}$. When $e = 0$, the household obtains only a high school education. When $e = \bar{e}$, the household attains a college degree and works $(1 - \bar{e}) > 0$ in youth.

The discounted present value of income over the education phase is given by:

$$Y^s(e) = \frac{\omega_1(1 - e) - pe}{\tilde{R}^{T^s}} \tag{6}$$

for $e \in \{0, \bar{e}\}$. Here p is tuition, ω_1 is the initial real wage and $\tilde{R}^{T^s}$ discounts this flow over the T^s periods back to the initial period.[19]

Agents differ by their productivity, denoted θ. For agents with a higher value of θ, education is more productive. The return to education is evident from labor earnings in the second and third phases of the lifecycle. Specifically,

[18] This is explicit in the specification of (18).

[19] Throughout this discussion, define $\tilde{R}^x = (1 + R + R^2 + ... + R^{x-1})/R^{x-1}$ where R is the real interest rate and x is the length of the period of the flow that is being discounted.

a household of ability θ that chooses education e obtains labor income of $\omega_j H(e,\theta)$ for $j=1$ (early work) and $j=2$ (late work). If $e=\bar{e}$, $H(\bar{e},\theta)=h(\bar{e})\theta$ where $h(\bar{e})$ represents the accumulation of human capital from college. In this specification, there is a complementarity between ability and the return to school. If instead the agent chooses no college, $e=0$, then labor income is ω_j in period $j=1,2$, where $H(0,\theta)=1$ for all θ.

The discounted present value of income over the early work phase of life is:

$$Y^e(e,\theta)=\frac{\omega_1 H(e,\theta)}{\tilde{R}^{T^e}} \tag{7}$$

for $e\in\{0,\bar{e}\}$ where $\tilde{R}^{T^e}$ discounts the flow of income during the middle phase back to the start of the early work period. Similarly, the discounted present value of income over the late work phase of life is:

$$Y^l(e,\theta)=\frac{\omega_2 H(e,\theta)}{\tilde{R}^{T^l}} \tag{8}$$

for $e\in\{0,\bar{e}\}$ where $\tilde{R}^{T^l}$ discounts the flow of income during the final phase back to the start of the late work period. Both of these flows depend on ability, θ, only if the agent attends college. Assume $\omega_2\geq\omega_1$ to allow for some experience effect on wages.

These three phases capture a couple of key dimensions of the education choice. First, in the initial phase the household bears the direct cost of education. In addition, if there are borrowing constraints (introduced below), the household bears an additional cost due to imperfect consumption smoothing. Second, in the second and third phases, the ability specific returns of education to the household accrue and interact with the phase of the lifecycle. Thus the gains to education have an individual component and an intertemporal component reflecting the shape of the lifecycle profile of wages. The distinction between the education and working phases is most important once borrowing constraints are present.

The lifetime discounted present value of income for the household is simply the sum of the discounted values from the three periods:

$$Y(e,\theta)=Y^s(e)+\frac{Y^e(e,\theta)}{R^{T^s}}+\frac{Y^l(e,\theta)}{R^{T^s+T^e}}. \tag{9}$$

Here the flows defined earlier are further discounted back to the initial period. Let $c(e,\theta)$ be the level of constant consumption such that over the T periods of life the discounted present value of consumption would equal $Y(e,\theta)$. That is

$$c(e,\theta)=\frac{Y(e,\theta)}{\tilde{R}^T}. \tag{10}$$

If the household has strictly concave utility over consumption in a period, $u(c_t)$, and discounts at a rate of $\beta=\frac{1}{R}$, then absent borrowing constraints, the household will choose to consume $c(e,\theta)$ in each period, given (e,θ). Let

$$V(e,\theta)=u(c(e,\theta))\tilde{\beta}^T \tag{11}$$

denote the lifetime flow of utility with $\tilde{\beta}^T=1+\beta+\beta^2+...\beta^{T-1}$.

If there are no capital market imperfections, then the household choice of education is simply a comparison of

$Y(0, \theta)$ and $Y(\bar{e}, \theta)$. This is equivalent to maximizing lifetime utility since consumption and utility flows are ordered by the discount present values of income associated with the two choices.

Since $Y(\bar{e}, \theta)$ is an increasing function of ability, there will exist a critical value of ability, denoted θ^* such that $Y(\bar{e}, \theta^*) = Y(0, \theta^*)$. For this ability and above, college is the optimal choice of the household, ie. $e^*(\theta) = \bar{e}$ iff $\theta \geq \theta^*$.

As in the static framework presented in section 2.1, this model predicts perfect sorting, i.e. the education choice depends only on a comparison of the individual's ability relative to a critical level of ability. There is no mismatch. We enrich the model to include sources of mismatch.

4.3 Tastes

So far the model assumes that agents have the same taste for education. In this discussion a choice specific shock, denoted ε, is added to the household problem to allow for taste differences.[20]

These taste differences can have multiple origins. For example, these taste could reflect differences in attitudes about education from parents and/or peer groups. In this case, the choice, for example, by high ability people not to go to college could simply indicate a taste for work relative to school. There is nothing inefficient about this choice. Alternatively, the taste variation may be interpreted as an added social cost of obtaining an education.

Assume ε affects the value associated with a college education, i.e. $e = \bar{e}$.[21] Then the value of attending college becomes

$$V(\bar{e}, \theta) = u(c(\bar{e}, \theta))\tilde{\beta}^T + \varepsilon. \tag{12}$$

The college choice again has a cut-off property. It entails a critical value of ability, denoted $\theta^*(\varepsilon)$, that depends on the taste shock. For $\theta > \theta^*(\varepsilon)$ the agent chooses higher education, $e = \bar{e}$; else the optimal decision is $e = 0$. For the optimal choice, $\theta^*(\varepsilon)$ will be decreasing in ε.

Figure 3 summarizes the solution. The function $\theta^*(\varepsilon)$ is shown as the downward sloping curve. To the right of this curve, the agent will choose $e = \bar{e}$, and to the left the optimal choice is $e = 0$. To be clear, this choice reflects both ability and the taste shock. Thus a high ability agent drawing a low taste shock, i.e. a large negative value of ε, may optimally choose $e = 0$. This is not mismatch.

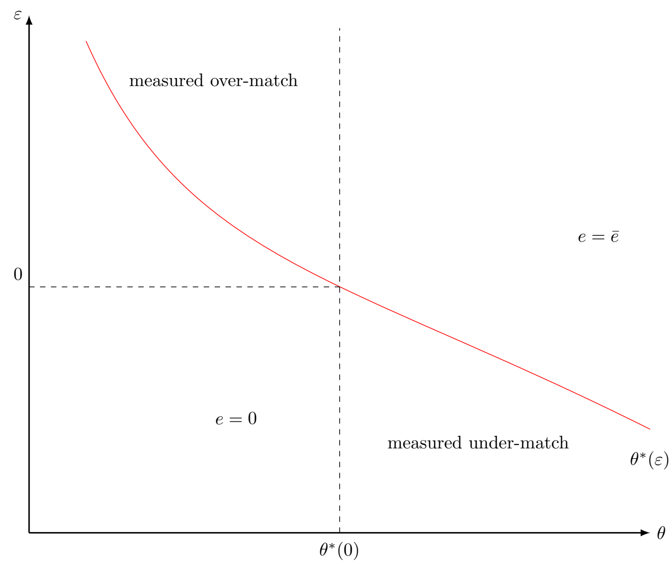
However this specification allows for **measured** mismatch if taste shocks are not directly observed. Consider a high ability agent with a large negative value of ε who chose $e = 0$. Conditioning on ability but not on tastes, this appears to be a mismatch.

4.4 Borrowing Constraints

The analysis focuses on the impact of borrowing constraints during the education phase. While the issue of outstanding student loans has been prominent in the US recently, student loans are important for higher education demand in other countries, such as Japan, where tuition is relatively high.[22]

[20] See Keane and Wolpin (2001) as a leading example of adding taste shocks to an education choice model.
[21] For now, assume that θ and ε are independently distributed.
[22] See `http://www2.jasso.go.jp/about_jasso/documents/e2015_01_28.pdf` for a presentation of facts about student loans in Japan.

Figure 3: Under- and Over-Matching with Taste Shocks



This figure shows the optimal choice of education given taste shocks. It also shows under- and over-matching based on merit alone.

To do so, we need to keep track of the debt of the household as well as initial assets. A borrowing constraint applies to the amount of debt accumulated during the school phase.[23] For the other phases of the lifecycle, the household is able to perfectly smooth consumption.

Let B represent the debt outstanding (assets held) at the start of the early working phase. Then the value of income over the early and late working years discounted back to the start of the early work period is given by

$$Y^{el}(e,\theta,B) = Y^e(e,\theta) + \frac{Y^l(e,\theta)}{\tilde{R}^{T^e}} - B \tag{13}$$

with $Y^e(e,\theta)$ and $Y^l(e,\theta)$ defined above.

Given the absence of borrowing constraints after the first phase, the household will smooth consumption over the last two phases of life generating a flow of utility captured by the value $V^{el}(e,\theta,B)$ given by

$$V^{el}(e,\theta,B) = u(c^{el}(e,\theta,B))\tilde{\beta}^{(T^e+T^l)} \tag{14}$$

where $c^{el}(e,\theta,B) = \frac{Y^{el}(e,\theta,B)}{\tilde{R}^{T^{el}}}$.[24]

During the school years, the household borrows b each period and consumes $c^s = \omega(1-\bar{e}) - p\bar{e} + b$. At the end of the school period, their debt outstanding is $B = b(1 + R + R^2 + ...R^{T^s-1})$.

Let $\bar{B}$ be a ceiling on debt outstanding at the end of the school period and $\bar{b}$ the borrowing limit imposed in each period of the school phase.[25] If the amount borrowed in each period of the school phase needed to finance the flow of consumption under the assumption of perfect capital markets, given by (10) with $e = \bar{e}$, is less than $\bar{b}$, then the

[23]The model does not include an endogenous borrowing constraint along the lines of Lochner and Monge-Naranjo (2011). But the specification is flexible in that the constraint is ultimately estimated.

[24]Here, following the notation developed earlier, $\tilde{R}^{T^{el}} = (1+R+R^2+...R^{T^e+T^l-1})/R^{T^e+T^l-1}$ and $\tilde{\beta}^{T^e+T^l} = 1+\beta+\beta^2+....+\beta^{T^e+T^l-1}$.

[25]These are linked by $\bar{B} = \bar{b}(1 + R + ...R^{T^s-1})$.

borrowing constraint is irrelevant.

Alternatively, if $\bar{b}$ is sufficiently low, then the borrowing constraint will bind. When the constraint binds, household consumption during each period of the school phase is given by

$$c^s(\bar{e},\theta,\bar{B})(1+\beta+\beta^2+....+\beta^{T^s-1})=\frac{\omega_1(1-\bar{e})}{\tilde{R}^{T^s}}-p\bar{e}+\bar{B} \tag{15}$$

where $\bar{B}$ is the amount of debt outstanding at the end of the school phase from borrowing $\bar{b}$ each period. The household will smooth consumption during the school phase but it is not able to smooth consumption between the school and working phases. The binding borrowing constraint creates an additional cost of college that distorts school choice.

During the school period the household choosing $e=\bar{e}$ has utility of

$$V^s(\bar{e},\theta,\bar{B})=u(c^s(\bar{e},\theta,\bar{B}))\tilde{\beta}^{T^s}. \tag{16}$$

Using (14) a household that chooses to go to college with a binding borrowing constraint has lifetime utility of

$$V(\bar{e},\theta,\bar{B})=V^s(\bar{e},\theta,\bar{B})+\tilde{\beta}^{T^s}V^{el}(\bar{e},\theta,\bar{B}). \tag{17}$$

In the presence of borrowing constraints, the household chooses $e=\bar{e}$ iff $V(\bar{e},\theta,\bar{B})\geq V(0,\theta)$. Clearly if the borrowing constraint binds the value of obtaining a college degree is lower, i.e. $V(\bar{e},\theta)>V(\bar{e},\theta,\bar{B})$.[26]

Thus the borrowing constraint distorts the education decision. It produces undermatching since some high ability household who would have chosen college do not do so given the additional cost of education created by the borrowing limit.[27] This will be more costly if $\bar{b}$ is close to zero and if the household is sufficiently risk averse so that the lack of consumption smoothing is costly. The magnitude of the under-matching due to a binding borrowing constraint will be a focus of the estimation.

Thus far we have assumed household wealth of zero. Of course, agents may have additional resources available to them during the school stage, say as transfers from parents. Denote this form of wealth, evaluated as a flow during each period of the school phase, as z, and the value of these transfers at the end of the school phase as Z. If, despite these transfers, the borrowing constraint binds, then the debt outstanding at the end of the school phase will remain $\bar{B}$. Of course, the transfer of z each period will increase utility during the school phase and reduce under matching. The magnitude of this affect is, again, an empirical issue studied below.

5 Estimation

The section puts the model and moments together. The point is to identify the factors that determine the levels of under and over-matching across the countries. While other studies have documented various measures of mismatch, the contribution here is to estimate the relative importance of the sources introduced by the theory model. In

[26] $V(0,\theta)$ and $V(\bar{e},\theta)$, the values of no education and education respectively, are given in (11).

[27] See Kim (2013) for a discussion of this in an equilibrium model and the consideration of the affects of education subsidies to relax these constraints.

addition, the estimation is informative about the returns to college, $h(\bar{e})$, by country and the economic costs of mismatch.

5.1 Functional Forms and Parameterization

In order to take the models to the data, a number of functional form assumptions are necessary. Some parameters are set and others, as described next, are estimated. After the baseline estimation, a lengthy robustness section explores alternative specifications of the model and these functional forms.

We assume a Pareto distribution for ability, with a shape parameter denoted ϕ. That is the CDF of ability, θ, is given by $1-\theta^{-\phi}$ with a mean of $\frac{\phi}{\phi-1}$.[28] Lower values of ϕ translate into larger tails of the Pareto distribution. We will estimate ϕ, allowing it to differ across countries.

The taste shocks are assumed to be uniformly distributed in the interval $[-\bar{\varepsilon}, \bar{\varepsilon}]$. The parameter $\bar{\varepsilon}$ controls the dispersion of taste shocks and is estimated as well. Note that taste shocks are symmetrically distributed around 0 by assumption.[29]

Agents make education decisions based on their true ability, θ. But, in the data analysis, the PIACC test score, which is assumed to be a noisy signal of ability, is observed and used to predict the education outcome. To mimic this, each household receives a test score, ts_i, equal to its true ability plus a normally distributed noise term with mean zero. The precision of the signal is parameterized by σ: i.e.

$$ts_i = \theta_i + \sigma\zeta_i. \tag{18}$$

Importantly, in the baseline model the test score does not factor directly into the individual's education choice. That decision is based on observed ability. Instead, as reported in Table 2, the test score is an input into the logistic regression predicting the education choice of an individual. It is thus a basis for the measured mismatch.

In our framework, only ability has a causal effect on the test score. Education and training are not included as covariates in (18).[30] Instead, both the college choice and the test score reflect differences in ability. As we shall see, the model does a very good job matching the data despite restricting the college choice not to have a direct influence on the test score. We return to the issue of other factors influencing test scores in two exercises reported in section 6.1. First, we add additional covarites to (18). Second, the model is re-estimated using moments for the US from the NLSY where the exam is taken prior to college. In addition, section 7.2 considers an alternative structural model where the test score depends on ability.

As noted earlier, the PIACC score is given to working adults, not to secondary or tertiary level students. The test captured in (18) could be given any time during the agent's life within the model. The education decision and return to experience are both, in our structural model, functions of ability and thus captured by (18).

As a normalization, set $\omega_1 = 1$ in each country. So all variables are relative to the compensation rate in period 1. For ω_2, we draw on Table 3 in Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) which reports wage profiles from PIACC. We match that wage profile with the two periods of work in our model, early and late.

[28] See Jones (2015) for a recent discussion of the origins and implications of the Pareto distribution for studying income and wealth inequality.

[29] The section on robustness considers another specification of taste shocks without bounded support.

[30] Arum and Roksa (2011) study the performance of over 23,000 students from 24 universities using the College Learning Assessment. They test students prior to the start of college and then after. They find little difference in performance after four years of college.

As a check on these profiles, a similar calculation was made using estimates of the lifecycle profile for someone without a college degree using data from the PSID. The calculated intertemporal return for the US was very close to that calculated from the regression results in Hanushek, Schwerdt, Wiederhold, and Woessmann (2015).[31]

Finally, the model includes a country specific price of education, p, and a time spent in school, $\bar{e}$, during the school phase. For the estimation, $\bar{e} = 0.75$ so that a student is in school for 9 or the 12 months of each college year. Our calibration of the price is from OECD, Education at a Glance and is determined for each of the 21 countries.[32] As a fraction of the US tuition, for our four countries, the price of education was near zero in Germany, 26% in Italy and 93% in Japan.

The distribution of outside wealth of agents, Z, is assumed to be Pareto with a shape parameter of aZ. We set aZ to be the same for all countries. We do not have information on household or parental wealth in the PIACC data. Thus we have not direct information allowing us to correlate tastes and household wealth induced by parental background. Accordingly the baseline model assumes that wealth is independent of other unobserved household attributes. In fact, for the baseline model we set outside wealth to zero, thus allowing the borrowing constraint to have a large affect. As we shall see, the borrowing constraint is not estimated to bind and thus has no influence on mismatch.

5.2 Simulated Method of Moments: Approach

The simulated methods of moments (SMM) approach finds the parameter vector, Θ, to minimize the weighted difference between simulated and actual data moments:

$$\pounds(\Theta) \equiv (M^d - M^s(\Theta))W(M^d - M^s(\Theta))'. \tag{19}$$

The logic is that the moments computed from the data summarize the relevant sources of variation. By choosing the parameters to match these data moments with the simulated data counterpart, the estimation brings the model and data together.[33]

The parameters estimated by SMM are $\Theta \equiv (\phi, \bar{\varepsilon}, \sigma, h(\bar{e}), \bar{b})$ where: ϕ is the shape parameter for the Pareto distribution of ability, $\bar{\varepsilon}$ parameterizes the taste shocks, σ parameterizes the noise in the test, $h(\bar{e})$ is the return to college and $\bar{b}$ is the borrowing constraint.

The data moments are those presented in Table A2 for the 21 countries. These include the moments presented in Table 2. These moments are supplemented by regression results reported in Hanushek, Schwerdt, Wiederhold, and Woessmann (2015), Table A5, that link the wages to test scores.[34] For their exercise, the sample was workers aged 35-54, similar to our late working phase. Specifically, letting i be an individual, there are two regressions that we use:

$$E[\omega_{2,i}|\cdot] = \nu_{01} + \nu_1 * test_i \tag{20}$$

[31] To obtain this, lifecycle income profiles were estimated by education group for the PSID and the return for the two age groups was calculated from those estimates. Thanks to Guozhong Zhu for this cross-check on the PIACC based results.

[32] See `http://www.oecd.org/education/eag.htm` indicator B5.

[33] See Adda and Cooper (2003) and references therein for a discussion of this approach and properties of the estimates.

[34] We rely on Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) for these moments as we do not have access to these data.

and

$$E[\omega_{2,i}|\cdot] = \nu_{02} + \nu_2 * test_i + \nu_3 * ed_i \tag{21}$$

The ν_1 coefficient relates the wage of individual i in the late work phase with that agent's test score. The ν_2 coefficient is similar though education is an added regressor.

In the results of Hanushek, Schwerdt, Wiederhold, and Woessmann (2015), the test results have significant explanatory power. Among other things, this implies that the PIACC score is not simply noise, uncorrelated with economic outcomes.

The simulated moments are constructed in exactly the same manner as those constructed from the data. So, for example, in the logistic regression of (5), the test score was normalized within each country to have a mean of zero and a standard deviation of unity. This was also done in the simulated data, using the test score generated by (18) as the raw input. Likewise, the (ν_1, ν_2) coefficients were created in a way parallel with the data-based regressions.

To be clear, the PIACC test results reported in Table 1 were not used directly as moments in the estimation. This is simply because of our inability to literally simulate the PIACC exam. Thus the raw test scores are normalized and used in the logistic regression, as in (5).

It is important to explain intuitively how these moments identify the key parameters. This discussion is continued in a more formal way after the presentation of the baseline estimates since the mapping is multidimensional. i.e. variations in some parameters are reflected in many moments.

The parameter controlling the distribution of ability, ϕ, has a direct affect on the college rate. If there were no taste shocks, as in the model of section 4.2, then ϕ could determine the college rate. There would be no mismatch.

The presence of the taste shocks, parameterized by $\bar{\varepsilon}$ creates mismatch and also breaks the tight link between the test score and the education outcome. The addition of a noisy test score also creates mismatch since true ability is not reflected in the PIACC score. This noise also weakens the link between the test score and education since the education choice is assumed to depend on true not measured ability.

Importantly, the taste shock directly influences the education choice while the noise in the test score has no affect on this decision. In theory, these are distinct channels.

The return to education, $h(\bar{e})$, directly influences the college rate. It also underlies the relationship between wages in the late work period and the level of education, i.e. the (ν_1, ν_2) moments.

A more formal way to understand local identification is to look at the affects of variations in the parameters on the moments near the parameter estimates. The point, in part, is to be sure that for each parameter there is a moment responding to it so there is local identification. Table 5 presents the elasticities of the moments with respect to variations in the parameters. These calculations are conducted at the baseline parameter estimates. From this table, variations in the ability shape parameter has large effects on all the moments, particularly the college rate, the mean of the logistic regression and the wage regressions. Variations in $\bar{\varepsilon}$ influence the mean of the logistic regression and the over-match rates. The test score noise σ has no influence on the college rate as that choice is based on true ability. It has a large and asymmetric affect on under- and over-match rates. The return on education matters for all moments, particularly the college rate and the constant in the logistic regression.

The estimation is undertaken with an identify matrix as a weighting matrix. The model has more moments than parameters so that the choice of the identity matrix matters. The estimates are consistent with the identify matrix but are inefficient, in large sample, compared to using the inverse of the variance-covariance matrix of the moments. For our study, we do not have access to the data underlying the estimates of (ν_1, ν_2) and thus cannot compute the variance-covariance matrix directly from the data.

5.3 Baseline Results

The estimation starts with the baseline model of no borrowing constraints. We then allow for borrowing constraints. We initially focus on the 4 key countries and then broaden the sample to include all 21 countries.

There are a couple of key findings. First, there is no evidence of binding borrowing constraints in any of the countries. Second, noise in the test score, parameterized by σ, plays the major role in matching the country specific moments. Third, the taste shocks, parameterized by $\bar{\varepsilon}$, plays a minor role.

5.3.1 No Borrowing Constraints

For this part of the analysis, capital markets are perfect. In this case, $\bar{b}$ is set at a large enough value that agents are able to borrow as much as needed to smooth consumption between the school and work phases.

The parameter estimates for all countries are presented in Table A1 and those for the select group of four countries are in Table 3. The data and simulated moments are presented in Tables A2 and A3 for all the countries and in Table 4 for the four countries. Table 5 is informative about the effects of the parameters on the moments and is used to interpret the estimation results.

Table 3: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
			Baseline		
Germany	2.545	1.354	1.186	0.803	na
Italy	2.835	0.893	1.586	0.728	na
Japan	4.243	0.464	0.511	1.227	na
US	3.137	0.498	0.583	1.056	na
			BC		
Germany	2.545	1.354	1.186	0.803	2.634
Italy	2.836	0.893	1.586	0.729	1.638
Japan	4.243	0.464	0.511	1.227	2.622
US	3.137	0.498	0.583	1.056	2.622
	Estimated No Taste Shocks: $\bar{\varepsilon} = 0$				
Ger.	2.542	na	1.198	0.804	na
It.	2.837	na	1.584	0.729	na
Jap.	4.247	na	0.510	1.227	na
US	3.143	na	0.582	1.056	na

This table reports parameter estimates for the baseline, the endogenous borrowing constraint and no taste shock models for the four leading countries.

These estimates are best understood relative to the moments presented in Table 4 for the countries. The education rate is lowest in Italy and Germany and highest in Japan. As noted earlier, tuition is relatively low in Italy and almost zero in Germany. A challenge is matching the college rates with the costs of education.

The estimated return to education is key. From Table 5, the education rate is very sensitive to $h(\bar{e})$. The estimated return to education is relatively low in Italy and Germany and much higher in Japan. In this way, the college rate is low in these countries despite the low tuition rates.

From Table 4, mismatch is relatively high in Italy and lowest in the US. Only Japan has more over-matching than under-matching. From Table 5, a high return to education leads to a high over-match and low under-match rate, as in Japan. For Italy, the high value of σ leads to a high level of mismatch, with the asymmetry in the direction of under-matching.

The estimated coefficient on the test score is relatively low in Italy and Japan and highest in the US. From Table 5, the value of α_1 is influenced (inversely) by the ability distribution, ϕ, and the return to education, since these impact the education decision, as well as the noise in the test score, σ. The low value of α_1 in Italy largely reflects the high level noise in the test score. For Japan there is relatively little noise in the test but a very high values of both the ability parameter and the returns to education. Evidently this implies that the education decision is less correlated with the test score. As for the US, the test score is not very noisy, i.e. α_1 is relatively large, and this is not offset by the ability estimate.

The coefficients from the wage regressions largely reflect the estimated returns to education and the ability distribution. For Italy the low return to education translates into low values of (ν_1, ν_2). In Japan, the high return to education is partially offset by a low estimated mean ability to match the (ν_1, ν_2) moments.

Looking across the parameter estimates, a couple of features stand out. First, the return to education is relatively low for Germany. This arises from two features: (i) the relatively low college education rate in Germany and (ii) the low tuition rates in Germany. To offset the low cost, the estimated return to college is low. Note that this return is relative to what is obtained without going to college. Germany has a well structured program of apprenticeships which increases the productivity and wages of those not going to college.

Second, the estimated mean ability is relatively low in Japan compared to, say, Germany and Italy. As noted above, this estimate comes from matching the college rate moment which, for Japan, would be very high with the estimated high return to education. To be clear, the raw PIACC scores were not included as moments for the estimation since we have no way to simulate those test results directly. These scores, normalized within each country, are used in the logistic regression as noisy signals of ability.

The fit is the unweighted sum of squared differences between the data and simulated moments, as in (19). From Table 4, the estimated model with only 4 parameters does a good job of matching the 7 moments.

To better understand the identification of the various parameters, Table 6 presents the simulated moments for parameter perturbations based on the estimated values. In particular, the block labeled $\bar{\varepsilon}$ sets the dispersion in taste shocks to 0, the block labeled $\sigma = 0$ eliminates the noise in the test score and the $h(\bar{e}) = 1$ case eliminates the return to college. For each of these treatments, the table presents the simulated moments when all other parameters are kept at their baseline values. **From this analysis, it seems clear that the measurement error in the test score is key element to matching the moments.**

As is clear the fit column, the noise in the test score plays a very prominent role in the analysis. Not surprisingly, the noise has a large influence on the mismatch rates and on the logistic regressions of education on the test score. Without the noise in the score, the model does not come close to matching those regression coefficients.

The deterioration in the fit from eliminating the taste shocks is minimal. This is consistent with the small

Table 4: Moments: Data and Simulated

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Data					
Germany	0.373	0.104	0.062	-0.720	1.160	0.235	0.144	na
Italy	0.230	0.146	0.069	-1.510	0.890	0.132	0.071	na
Japan	0.597	0.078	0.108	0.230	0.860	0.184	0.111	na
US	0.455	0.055	0.045	-0.360	1.510	0.279	0.149	na
			Baseline					
Germany	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Japan	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
			Estimated No Taste Shocks: $\bar{\varepsilon} = 0$					
Ger.	0.345	0.104	0.086	-0.714	1.162	0.225	0.142	0.002
It.	0.209	0.143	0.079	-1.507	0.891	0.135	0.064	0.001
Jap.	0.548	0.082	0.122	0.241	0.860	0.189	0.103	0.003
US	0.414	0.068	0.075	-0.352	1.512	0.261	0.152	0.003

This table reports data and simulated moments for the estimated models.

elasticities with respect to $\bar{\varepsilon}$ reported in Table 5. This does not mean that in general the taste shocks do not matter. At other points in parameter space, variations in $\bar{\varepsilon}$ can have large effects on moments. In particular, the taste shock can produce mismatch. But this is not the case at the estimated values.

Table 5: Elasticities of Moments

parameter	college	under-match	over-match	α_0	α_1	ν_1	ν_2
ϕ	-0.591	2.555	-0.107	-5.325	-1.587	-1.245	-1.819
$\bar{\varepsilon}$	-0.000	0.003	-0.001	-0.005	-0.003	-0.000	0.001
σ	0.000	1.576	0.565	-0.920	-1.145	-0.353	-0.457
$h(\bar{e})$	3.685	-1.498	4.710	-70.420	-1.843	0.269	1.603
$\bar{b}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000

This table reports elasticities of moments with respect to parameters for the Baseline model, US estimates.

Given these results on the apparent irrelevance of the taste shock, the model was re-estimated with the restriction of $\bar{\varepsilon} = 0$. The results appear in Tables 3 and 4 in the "Estimated No Taste Shocks: $\bar{\varepsilon} = 0$" block. Clearly, these estimates and moments are quantitatively very close to the baseline results indicating that taste shocks are adding almost nothing to the baseline model.

Shutting down the return to education has the expected affect of reducing the college rates in all countries except for Germany and Italy, where $h(\bar{e})$ was estimated to be less than unity. Interestingly, this reduces the under-match rate a little in Italy and increases the over-match rate significantly.

Returning to the sources of mismatch. From the results in Table 6, it is clear that the estimated model points to the measurement error in the test score as the source of mismatch. To be clear, the role of the test is solely through its use in predicting the likelihood of college, both in the actual and simulated data sets. The score is not used by agents in making the education decision. We return to this below.

Table 6: Moments from Perturbations

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Baseline					
Germany	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Japan	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
			Simulated No Taste Shocks: $\bar{\varepsilon} = 0$					
Ger.	0.343	0.103	0.086	-0.724	1.175	0.226	0.142	0.002
It.	0.209	0.143	0.079	-1.509	0.892	0.135	0.064	0.001
Jap.	0.547	0.082	0.122	0.237	0.859	0.189	0.103	0.003
US	0.413	0.068	0.075	-0.354	1.516	0.262	0.153	0.003
			No Noise: $\sigma = 0$					
Ger.	0.345	0.000	0.000	4.466	46.215	0.302	0.217	2056.900
Italy	0.209	0.002	0.000	-9.074	46.242	0.255	0.174	2114.073
Jap.	0.548	0.000	0.000	24.699	66.936	0.293	0.195	4964.765
US	0.414	0.000	0.000	18.079	96.689	0.326	0.217	9398.981
			No Return: $h(\bar{e}) = 1$					
Ger.	0.602	0.089	0.143	0.478	0.816	0.250	0.189	1.616
Italy	0.512	0.124	0.145	0.055	0.532	0.173	0.111	2.667
Jap.	0.231	0.111	0.046	-1.520	1.363	0.158	0.059	3.458
US	0.349	0.074	0.058	-0.723	1.667	0.253	0.140	0.169

This table reports data and simulated moments for the estimated models for large variations in parameters.

5.3.2 Borrowing Constraints

One of the key features of the model is the possibility that mismatch reflects capital market imperfections. This section builds upon the baseline and considers the role of these frictions in matching the moments. It adds the parameter $\bar{b}$ to the estimation. For this exercise, assume $u(c) = ln(c)$.[35]

For this analysis, outside wealth of the household was set to zero. If anything, this restriction will lead to an overstatement of the affects of capital market imperfections.

The panels of Tables 3 and 7 labeled "BC" present results for the estimation in which borrowing constraints were allowed. These are results for the four countries. Comparable results for all countries are reported in Tables A1 and A3.[36]

The results are striking: the estimates and moments with endogenous borrowing constraints are identical to those without them. That is, the estimation selects a level of the borrowing constraint such that it is not binding. In contrast to the *a priori* reasoning, capital market frictions are not needed to produce under-matching.

To be clear, this is not to say that borrowing is irrelevant. The panels labeled "No Borrowing" report simulation results when $\bar{b} = 0$ is imposed. From Table 4, the fit is considerably worse than the baseline. The borrowing restriction reduces the college rate and produces much more mismatch, particularly through under-matching. Clearly borrowing is not irrelevant.

One concern with this finding is that perhaps the estimation procedure would not find a binding borrowing

[35] Allowing more risk aversion would increase the costs of a binding borrowing constraint. But, as seen below, the borrowing constraint is estimated to be non-binding and so the choice of risk aversion is not pertinent.

[36] These estimates were obtained from many different starting values, including setting $\bar{b} = 0$.

Table 7: Moments with Borrowing Constraints: Data and Simulated

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Baseline					
Germany	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Japan	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
			BC					
Germany	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
It.	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Japan	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
			No Borrowing					
Ger.	0.139	0.133	0.025	-2.451	1.852	0.217	0.089	3.534
Italy	0.030	0.180	0.005	-5.101	1.998	0.094	0.018	14.176
Jap.	0.370	0.094	0.082	-0.612	1.082	0.200	0.081	0.812
US	0.328	0.077	0.053	-0.856	1.717	0.264	0.136	0.306

This table reports data and simulated moments for the estimated models. The "No Borrowing" treatment uses the baseline parameter estimates but sets the borrowing limit to zero.

constraint even if it was present: i.e. can we identify $\bar{b}$ when borrowing constraints are present?[37] To study this, we simulated data with a binding borrowing constraint. The same set of moments was calculated from the simulated data and used in an estimation exercise. Trying multiple initial guesses of the parameters, the estimation exercise did uncover the binding borrowing constraint when it was present.

5.4 All Countries

This discussion focuses largely on the four key countries. But, to be clear, Tables A1 , A3 and A4 show that our principal results hold for all of the 21 countries. That is: (i) there is no evidence of a binding borrowing constraint, (ii) the ability of the model to fit the data moments depends crucially on the presence of noisy test scores and (iii) taste shocks add very little.

As noted earlier, there are interesting correlations across countries between education rates and mismatch. As seem in Table 8, countries with higher education rates tend to also have higher over-match and lower under-match rates. These same patterns are seen in the simulated data from the estimated model.

Looking across countries, one of the most important factors influencing education rates is the return to college, $h(\bar{e})$. As indicated by the bottom panel in Table 8, from the simulated data, cross-country variations in the estimated value of $h(\bar{e})$ are positively correlated with the education rate and the over-match rate and negatively correlated with the under-match rate.

[37] Thanks to Immo Schott for suggesting this exercise.

Table 8: Cross-Country Correlations

correlation	actual data	simulated data
(ed., under)	-0.65	-0.62
(ed., over)	0.73	0.77
$(h(\bar{e}), \text{ed})$	na	0.90
$(h(\bar{e}), \text{under})$	na	-0.75
$(h(\bar{e}), \text{over})$	na	0.53

This table reports cross-country correlations from actual and simulated data. Here "ed." is the education rate, "under" is the under-match rate and "over" is the over-match rate.

6 Robustness

This section considers various alternative specifications to inspect the robustness of our findings. This includes looking at other moments of the data. Throughout, the presentation focuses on the role of borrowing constraints and the contributions of the taste shocks and the noise in test score. The tables for these exercises are in the text for the four main countries and in the Appendix for all countries. Though the parameter estimates respond to changes in the moments, the conclusions regarding the insignificance of borrowing constraints and the taste shocks remain intact.[38]

6.1 Isolating Ability

The baseline results rely on a framework in which the PIACC score proxies for ability which in turn influences the college choice. In the model, the test score itself is independent of education attainment and work experience.

Here we discuss three ways to isolate ability in the PIACC score. The first limits the data to a younger cohort. The second adds regressors and produces a conditional PIACC score used in the logistic regression. The third looks at a sample from the US in which the ability measure precedes college attainment. For all cases, we find that our results hold: mismatch reflects noise in the measurement of ability and not borrowing constraints nor taste shocks.

6.1.1 Young Cohort

The baseline results are for individuals aged 25-39. Pooling across cohorts generates a benefit of providing a large sample. But, particularly for older individuals, the PIACC test results may be far removed from their measured ability at the time of their college choice.

Here, we restrict attention to a sub-sample with individuals aged 25-29. The estimates and moments are reported in Tables 9 and 10 respectively. For this exercise, the ν_1 and ν_2 coefficients are the same as in the baseline.[39]

Many features of the baseline results remain. First, the fit is about the same despite the difference in the cohort. In fact, the fit is slightly better for Italy and a bit worse for the US. Second, estimating an endogenous borrowing constraint does not improve the fit: the estimated constraint does not bind. Third, the noise in the test score matters a lot for the fit of the model, the taste shocks matter little.

[38]With respect to the taste shocks, they do play a larger role when they are linked to parent's education.

[39]We do not have access to the restricted PIACC data necessary to estimate these parameters for this young cohort.

Table 9: Young Cohort: Parameter Estimates

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
	Young Baseline				
Ger.	2.085	0.019	2.536	0.703	na
It.	2.844	1.195	1.646	0.761	na
Jap.	4.296	0.678	0.484	1.244	na
US	3.157	0.556	0.603	1.068	na
	Young, BC				
Ger.	2.085	0.020	2.536	0.703	6.788
It.	2.850	1.284	1.638	0.762	2.978
Jap.	4.301	0.563	0.484	1.244	3.074
US	3.139	0.252	0.611	1.066	6.196

This table reports estimates from a sub-sample of 25-29 year olds.

Table 10: Young Cohort: Data Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
	Data							
Ger.	0.332	0.146	0.072	-0.840	0.880	0.235	0.144	na
It.	0.241	0.147	0.096	-1.310	0.800	0.132	0.071	na
Jap.	0.621	0.104	0.115	0.370	0.860	0.184	0.111	na
US	0.474	0.070	0.044	-0.300	1.400	0.279	0.149	na
	Baseline							
Ger.	0.316	0.131	0.108	-0.836	0.884	0.213	0.137	0.002
It.	0.236	0.143	0.091	-1.309	0.800	0.136	0.067	0.000
Jap.	0.578	0.077	0.127	0.379	0.857	0.190	0.108	0.003
US	0.427	0.070	0.082	-0.289	1.403	0.258	0.151	0.004
	No Noise; $\sigma = 0$							
Ger.	0.316	0.000	0.000	421.804	5947.227	0.295	0.206	35537671.469
It.	0.236	0.000	0.000	-4.642	37.289	0.266	0.182	1342.616
Jap.	0.578	0.000	0.000	18.108	44.933	0.290	0.196	2257.142
US	0.427	0.000	0.000	17.674	86.497	0.327	0.219	7564.661

This table reports data and simulated moments for the estimated models for the young cohort.

Given that the college decision of the young cohort is closer to the time of their PIACC exam, one might expect there to be less noise in the score, i.e. σ to be lower. Comparing the estimates with the young cohort against the baseline, the opposite seems to be the case. This can be traced to the logistic regression results where the coefficient on the test score for the young cohort is smaller for all countries (only slightly so for Italy) compared to the baseline. This is indicative of more noise in the score.

Further, compared to the baseline, the (unconditional) returns to education is lower for all countries but Italy. Still the order remains the same.

6.1.2 Inferred Ability

Beyond ability, there are a number of factors such as labor market experience, parent's education, gender, etc. that impact an individual's PIACC score. These effects are not in the model. To offset them, the test score in the data was regressed on a set of covariates.

Specially, this regression was estimated at the individual level in each country:

$$num_i = \Gamma_0 + \Gamma_1 * age_i + \Gamma_2 * gender_i + \Gamma_3 * Parent'sEd_i + \tilde{\theta}_i \tag{22}$$

The residual, $\tilde{\theta}_i$, is a proxy for ability and is used in (5) to produce a new set of logit regression coefficients, (α_0, α_1) as well as revised measures of under- and over-matching from the data.

Using these moments, Tables 11 and 12 present the parameter estimates and simulated moments. The estimates for this case differ from the baseline estimates. The ordering of countries by ability, ϕ, noise in the test score σ and the returns to college, $h(\bar{\varepsilon})$, remain the same.

Table 11: Inferred Ability: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ	$h(\bar{e})$	b
	Inferred Ability Baseline				
Ger.	2.252	1.693	1.949	0.752	na
It.	2.592	2.795	2.144	0.683	na
Jap.	3.812	0.534	0.685	1.206	na
US	2.839	0.609	0.815	1.026	na
	Estimated: No Taste Shocks				
Ger.	2.229	0.0	2.019	0.749	na
It.	2.588	0.0	2.172	0.687	na
Jap.	3.811	0.0	0.683	1.206	na
US	2.793	0.0	0.845	1.021	na

This table reports parameter estimates for the baseline for the four leading countries using the residual of (22) as a measure of ability in the logistic regression instead of the PIACC numeracy score.

Though not shown in the tables, allowing a borrowing constraint did not alter the simulated moments. Further, eliminating the noise in the test score led, once again, to a large deterioration in the fit. Finally, the model was reestimating with no taste shocks, i.e. $\bar{\varepsilon} = 0$. As indicated in the blocks "Estimated: No Taste Shocks", the simulated moments and thus the fit of the model is almost identical to the baseline with inferred ability.

6.1.3 US Data: NLSY 1997

In this section of the paper we report estimates of the model using data from the National Longitudinal Study of Youth 1997 (NLSY97). The value added of this exercise comes from the inclusion in the data of an ASVAB score, which we use as a proxy for ability, that individuals took prior to their college years.[40] Thus concerns about the PIACC score reflecting the accumulation of human capital during college and after are mitigated. Our findings about the noisy test score being the main source of mismatch remains.

The individuals in the sample were between 12 and 16 at the end of 1996.[41] The sample has 8,894 observations. Individuals were between 28 and 32 years old in 2012 when the moments for the estimation were collected.[42]

[40]The ASVAB measures the respondents knowledge and skills in 12 components (detailed in NLSY97 Userss Guidebook p.82), including math and reading comprehension as measured in SAT or ACT. Individuals in the sample were between 13 and 17 years old when they took the ASVAB.

[41]For a full discussion of the data see `http://www.bls.gov/nls/nlsy97.htm`.

[42]The regressions for the two wage regressions are from the PIACC sample not from the NLSY97.

Table 12: Inferred Ability: Data and Simulated Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Data: Inferred Ability					
Ger.	0.373	0.123	0.078	-0.749	0.908	0.235	0.144	na
It.	0.230	0.154	0.072	-1.50	0.780	0.132	0.071	na
Jap.	0.597	0.094	0.125	0.207	0.747	0.184	0.111	na
US	0.455	0.085	0.061	-0.359	1.253	0.279	0.149	na
			Simulated Moments					
Ger.	0.336	0.123	0.106	-0.739	0.912	0.216	0.139	0.003
It.	0.205	0.150	0.090	-1.495	0.782	0.132	0.066	0.001
Jap.	0.546	0.094	0.130	0.218	0.746	0.188	0.106	0.003
US	0.415	0.082	0.090	-0.352	1.255	0.261	0.155	0.003
			No Noise					
Ger.	0.241	0.000	0.000	-3.391	42.304	0.270	0.188	1720.663
It.	0.379	0.000	0.000	0.552	8.501	0.195	0.128	63.881
Jap.	0.507	0.000	0.000	39.405	125.454	0.307	0.205	17088.314
US	0.164	0.043	0.000	-30.301	51.685	0.189	0.099	3440.030
			Estimated: No Taste Shocks					
Ger.	0.336	0.124	0.106	-0.741	0.912	0.217	0.138	0.003
It.	0.205	0.150	0.090	-1.497	0.783	0.132	0.065	0.001
Jap.	0.546	0.094	0.130	0.218	0.746	0.188	0.106	0.003
US	0.415	0.083	0.090	-0.352	1.254	0.263	0.157	0.003

This table reports data and simulated moments for the estimated models using the residual of (22) as a measure of ability in the logistic regression instead of the PIACC numeracy score.

Table 13: NLSY97: Moments: Data and Simulated

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			1997					
Data	0.306	0.097	0.036	-0.884	1.139	0.279	0.149	na
Baseline	0.313	0.114	0.087	-0.884	1.145	0.244	0.141	0.0042
BC	0.312	0.114	0.087	-0.884	1.145	0.244	0.141	0.0042

This table reports data from NLSY97 and simulated moments for the estimated models.

Table 13 presents the moments for the sample. Compared to the moments for the US sample from the PIACC data, reported in Table 4, the NLSY97 sample has a lower education rate, a higher under-match rate and a lower over-match rate.[43] The response of the education decision to the test score is lower in the NLSY97 sample.

Table 14 reports the parameter estimates. Compared to the estimates from the PIACC sample, the estimate of ϕ is smaller, there is considerably more noise in the test score (i.e. the estimate of σ is higher) and the return to education is lower. These parameter differences are consistent with the lower education rate in the NLSY97 sample as well as the lower response of the education choice to the test score. Though larger, the estimated variability of the taste shock remains irrelevant for the moments: i.e. setting $\bar{\varepsilon} = 0$ does not change the fit of the model.

Table 14: NLSY: Parameter Estimates

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	bbar
			1997		
Baseline	2.3446	1.6495	1.5158	0.8500	na
BC	2.3446	1.6495	1.5158	0.8500	19.997

This table reports parameter estimates using moments from the NLSY97.

Table 15: Adding Parental Education: Parameter Estimates

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
			Baseline		
Ger.	2.509	5.588	1.309	0.799	na
It.	4.163	6.442	0.635	0.942	na
Jap.	4.091	3.146	0.572	1.225	na
US	3.114	3.439	0.619	1.055	na
			BC		
Ger.	2.504	5.617	1.316	0.798	2.946
It.	4.163	6.442	0.635	0.942	3.062
Jap.	4.091	3.146	0.572	1.225	2.622
US	3.114	3.439	0.619	1.055	4.520

This table reports parameter estimated for the models with Parental Education and the taste shock equated.

6.2 Controlling for Parents' Education

The data includes very little additional information except parent's education. There is no direct counterpart in the model. But it is natural to associate the taste for education with parental education.[44]

The implications of doing so are explored by making two adjustments to the baseline specification. First, the logistic regression includes parental education with a coefficient α_2. Second, when the logistic regression is run on the simulated data, the realized taste shock is used as a proxy for parental education. Thus an additional moment, the regression coefficient on parent's education from the logistic regression, is added.

The estimation and moments are reported in Tables 15 and 16. The parameter estimates are quite different from the baseline. Matching the regression coefficient α_2 evidently requires much more variability in the taste shock than estimated in the baseline model.

The fit of the model, with the exception of Italy, remains quite good. This is particularly true for the regression coefficients, including the new moments. For Italy, this specification is unable to match the low college rate. This was not the case for the baseline. Also, there is more mismatch for Germany, Italy and the US in the simulated model relative to the baseline.

Still, the main results of the baseline are intact. There is no evidence of a borrowing constraint and the test noise is very important for matching the moments. However, in this specification the taste shocks are assumed to be a proxy for the effect of parent's education. So, the taste shocks play a larger role. As seen in Table 16, the fit worsens

[43]These differences in moments may reflect differences in the age of the sample and in the sampling structure.

[44]As suggested to us by Eric Hanushek, another exercise would allow parental education to also impact the return to education and not just tastes alone. Here we focus on the effect of parent's education on tastes in an attempt to generate a more prominent role for tastes and hence mismatch in our estimation.

Table 16: Parental Education Moments: Data and Simulated

	college	under-match	over-match	α_0	α_1	α_2	ν_1	ν_2	fit
				Data					
Ger.	0.373	0.101	0.042	-1.200	1.030	1.120	0.235	0.144	na
It.	0.230	0.131	0.053	-1.730	0.860	1.970	0.132	0.071	na
Jap.	0.597	0.064	0.086	-0.270	0.800	1.310	0.184	0.111	na
US	0.455	0.049	0.026	-0.760	1.390	0.890	0.279	0.149	na
				Baseline					
Ger.	0.366	0.101	0.065	-1.198	1.033	1.121	0.219	0.150	0.001
It.	0.358	0.097	0.033	-1.753	0.856	1.955	0.129	0.081	0.019
Jap.	0.577	0.072	0.096	-0.266	0.800	1.313	0.184	0.110	0.001
US	0.426	0.069	0.069	-0.754	1.394	0.893	0.257	0.155	0.004
				No Taste; $\bar{\varepsilon} = 0$					
Ger.	0.344	0.109	0.091	-0.717	1.084	0.000	0.221	0.139	1.494
It.	0.291	0.115	0.076	-1.038	1.008	0.000	0.135	0.060	4.386
Jap.	0.556	0.086	0.128	0.269	0.798	0.000	0.187	0.104	2.011
US	0.416	0.071	0.079	-0.346	1.428	0.000	0.258	0.150	0.970
				No Noise; $\sigma = 0$					
Ger.	0.366	0.000	0.000	-1.140	24.397	6.585	0.302	0.230	575.896
It.	0.358	0.001	0.000	-3.930	7.774	6.404	0.226	0.179	72.360
Jap.	0.577	0.000	0.000	4.481	20.753	6.325	0.293	0.206	445.884
US	0.426	0.000	0.000	2.174	29.026	6.356	0.326	0.223	802.258

This table reports data and simulated moments for the estimated models with Parental Education and the taste shock equated.

if the taste shocks are removed since $\alpha_2 = 0$ must hold in the simulated data.

According to Brunello and Checchi (2007) and Dustmann (2004), education attainment is more dependent on parental education in high early tracking countries. Evidently, parental intervention on teacher's recommendation can influence the choice of tracking and this effect seems more potent for more highly educated parents. In our estimates, the education decision does seem particularly sensitive to parental education in Italy, but not excessively in Germany.

6.3 Distribution of Taste Shocks

The baseline model restricts taste shocks to be uniformly distributed on a bounded interval while the test noise is not bounded. It might be that this restriction is the source of the prominent role of noisy test scores for our results. Here we explore our findings by allowing tastes to be normally distributed with a mean of zero.[45]

The results of the new estimation are reported in Tables 17 and 18. The estimates, except for $\bar{\varepsilon}$, are similar across countries to the baseline estimates. The fit of this model is about the same as the baseline. Further, there is no evidence of capital market imperfections, as indicated by the "BC" estimates and moments being so close to the baseline. The noisy signal remains critical to our results.[46]

As another gauge of the relative importance of taste relative to test scores as a source of variation, the model was re-estimated using the normally distribution taste shocks assuming no noise in test scores. The results are reported

[45] Martin Hackmann and Jonathan Eaton drew our attention to this point.
[46] The other cases of no taste shocks and no return to education are not shown in the interest of brevity.

Table 17: Normal Taste Shocks: Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
				Baseline				
Ger.	0.345	0.104	0.087	-0.714	1.161	0.227	0.145	0.002
It.	0.209	0.143	0.080	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.082	0.122	0.240	0.861	0.189	0.104	0.003
US	0.413	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
				Simulated No Noise: $\sigma = 0$				
Ger.	0.345	0.000	0.000	2.825	30.150	0.302	0.218	853.014
It.	0.209	0.005	0.000	-6.006	30.145	0.255	0.174	876.097
Jap.	0.548	0.000	0.000	16.100	43.727	0.292	0.195	2089.503
US	0.413	0.000	0.000	15.908	85.188	0.326	0.217	7266.602
				Estimated No Noise: $\sigma = 0$				
Ger.	0.328	0.108	0.104	-0.713	1.162	0.179	0.180	0.008
It.	0.200	0.138	0.088	-1.505	0.893	0.086	0.093	0.004
Jap.	0.541	0.097	0.148	0.242	0.852	0.206	0.158	0.008
US	0.393	0.078	0.094	-0.349	1.514	0.219	0.174	0.011

This table reports data and simulated moments for the estimated models allowing normal taste shocks.

in Tables 17 and 18 in the block labeled "Estimated No Noise: $\sigma = 0$". The fit is worse for all of the models though clearly this specification captures much of the variation in the data captured by these moments. We return to this specification in our discussion of output losses due to mismatch.

Table 18: Normal Taste Shocks: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
			Normal Taste Shocks		
Ger.	2.524	1.196	1.203	0.799	na
It.	2.842	0.806	1.577	0.729	na
Jap.	4.255	0.399	0.506	1.227	na
US	3.139	0.323	0.581	1.056	na
			BC		
Ger.	2.524	1.196	1.203	0.799	3.713
It.	2.842	0.806	1.577	0.729	3.696
Jap.	4.255	0.399	0.506	1.227	4.216
US	3.139	0.323	0.581	1.056	3.053
			Estimated No Noise: $\sigma == 0$		
Ger.	3.663	11.212	0	0.739	na
It.	5.447	8.815	0	0.751	na
Jap.	5.016	11.995	0	1.235	na
US	4.292	8.757	0	1.019	na

This table reports parameter estimates for models allowing normal taste shocks.

6.4 Alternative Measures of MisMatch

The results presented thus far rest on a particular view of the tails of the distributions of ability by education. Undermatch rates were calculated as the fraction of those obtaining a predicted value of college attendance over 80% who

Table 19: Additional Moments: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
	Baseline				
Ger.	2.690	1.096	1.054	0.823	na
It.	3.077	6.801	1.230	0.730	na
Jap.	4.267	0.370	0.507	1.228	na
US	3.309	0.394	0.530	1.072	na

This table reports parameter estimates for a specification with additional moments.

did not go to college. Over-match rates were calculating using a cut-off of below 20% probability of attaining college. These same cut-offs are used in the structural estimation.

It is useful to study the robustness of our findings to alternative cut-off values. To do so, we re-estimated the models for the four countries using under-matching (over-matching) cut-off values of 90% (10%) and 95% (5%). The estimates and moments are presented in Tables 19 and 20. The main results about the irrelevance of the borrowing constraint and the prominent role of noise in the test score remain.

Table 20: Additional Moments: Data and Simulated

	col.	u(80)	u(90)	u(75)	o(20)	o(10)	o(25)	α_0	α_1	ν_1	ν_2	fit
	Data											
Ger.	0.373	0.104	0.043	0.138	0.062	0.022	0.082	-0.720	1.160	0.235	0.144	na
It.	0.230	0.146	0.069	0.192	0.069	0.016	0.100	-1.510	0.890	0.132	0.071	na
Jap.	0.597	0.078	0.034	0.106	0.108	0.052	0.149	0.230	0.860	0.184	0.111	na
US	0.455	0.055	0.022	0.076	0.045	0.007	0.065	-0.360	1.510	0.279	0.149	na
	Baseline											
Ger.	0.346	0.101	0.030	0.144	0.084	0.034	0.113	-0.714	1.166	0.218	0.136	0.003
It.	0.209	0.142	0.057	0.188	0.082	0.034	0.112	-1.510	0.894	0.126	0.070	0.001
Jap.	0.548	0.082	0.022	0.119	0.122	0.054	0.160	0.240	0.861	0.188	0.103	0.003
US	0.414	0.066	0.013	0.101	0.074	0.030	0.100	-0.352	1.515	0.254	0.145	0.006
	No Noise $\sigma = 0$											
Ger.	0.346	0.000	0.000	0.000	0.000	0.000	0.000	4.789	49.804	0.298	0.215	2396.643
It.	0.209	0.049	0.002	0.088	0.000	0.000	0.000	-1.870	5.748	0.230	0.174	23.791
Jap.	0.548	0.000	0.000	0.000	0.000	0.000	0.000	30.666	82.957	0.292	0.195	7666.281
US	0.414	0.000	0.000	0.000	0.000	0.000	0.000	20.952	111.364	0.320	0.211	12522.175

This table reports data and simulated moments for the estimated models using different measures of under- and over-match.

7 Alternative Structures

The analysis and estimation is based upon a particular structural model. The inferences are conditional on that structure. In particular, the model assumes: (i) agents decide on education knowing their ability and (ii) educational attainment has no influence on the test score. This section relaxes those parts of the structure and re-estimates the model. Our findings about the irrelevance of borrowing constraints and the role of taste shocks remain.

7.1 Decisions under Imperfect Information

Table 21: Imperfect Information: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ_t	σ_d	$h(\bar{e})$	b
	Baseline: Imperfect Information					
Ger.	2.532	1.370	1.194	0.188	0.798	na
It.	2.832	1.171	1.583	0.167	0.728	na
Jap.	4.189	0.544	0.508	0.175	1.196	na
US	3.095	0.503	0.590	0.230	1.037	na
	No Noise in Test: $\sigma_t == 0$					
Ger.	3.906	6.225	0	0.628	0.836	na
It.	6.195	2.075	0	0.553	1.007	na
Jap.	6.452	2.230	0	0.494	1.204	na
US	4.353	0.558	0	1.264	1.069	na

This table reports parameter estimates for models allowing a noisy signal of ability to influence the education choice.

As an alternative to the maintained assumption that agents make education decisions knowing their true ability, suppose the choice of education itself reflects imperfect information on ability. This might, for example, be more likely in education systems like Germany and Italy with early tracking, a point we return to below.

A risk neutral household makes an education choice by comparing expected lifetime income with and without higher education. At the time of the education decision, the household receives a signal about true ability. Thus there are two sources of noise in the model: the signal at the time of the education decision, parameterized by σ_d, and the noisy test result, parameterized by σ_t. As in the baseline model, the noisy test score does not directly influence the education decision.

Given these two sources of noise, the same estimation procedure as in the baseline model is followed. The parameter estimates are given in Table 21 and the moments in Table 22.

The parameter estimates are remarkably close to the baseline, reported in Table 3, including the noise in the test score, now denoted σ_t. For Germany and Italy, the estimated noise in the decision is about 10% of the noise in the test. For, Japan and the US, the noise in the decision plays a larger role relative to the test score. From Table 22, the treatment that eliminates the taste shocks entirely does not influence the results.[47]

To isolate the role of uncertainty in the test score, there is another treatment in which the test score noise is set to zero and the noise in the decision is estimated, "Estimated No Test Shock: $\sigma_t = 0$". The estimates of the noise in the signal for the education decision is much larger, as is the taste shock. The fit is not quite as good as the baseline, indicating that the noise in the test score does contribute to the fit of the model. Once again, though Germany and Italy are viewed as having systems with early tracking, the estimated noise in the decision alone is not obviously larger for those countries relative to, say, the US. This estimated model does produce asymmetry in mismatch, with the under-match rate higher in Italy and (modestly) in Germany compared to the over-match rate.

A model of education decisions under uncertainty is likely to fit better those institutional settings, such as Germany, in which the education decision is made at a relatively early age and is an administrative choice, rather than one made by an individual.[48] Along these lines, Brunello and Checchi (2007) provides an index of tracking,

[47] As households are risk neutral, studying the borrowing constraint is not of interest.
[48] See Döbert (2015) for a recent discussion of tracking in Germany.

Table 22: Imperfect Information: Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
				Baseline: Imperfect Information				
Ger.	0.345	0.104	0.087	-0.715	1.162	0.226	0.145	0.002
It.	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.081	0.123	0.240	0.859	0.189	0.108	0.003
US	0.413	0.069	0.076	-0.353	1.512	0.262	0.157	0.003
				No Taste Shock: $\bar{\varepsilon} = 0$				
Ger.	0.343	0.104	0.086	-0.725	1.177	0.226	0.144	0.002
It.	0.209	0.143	0.079	-1.511	0.894	0.136	0.065	0.001
Jap.	0.547	0.081	0.123	0.239	0.862	0.189	0.108	0.003
US	0.413	0.069	0.076	-0.355	1.514	0.262	0.157	0.003
				Estimated No Noise in Test: $\sigma_t = 0$				
Ger.	0.329	0.110	0.106	-0.710	1.162	0.190	0.175	0.007
It.	0.201	0.139	0.088	-1.505	0.891	0.115	0.084	0.002
Jap.	0.541	0.093	0.151	0.242	0.861	0.170	0.127	0.006
US	0.392	0.077	0.101	-0.351	1.516	0.229	0.176	0.011

This table reports data and simulated moments for the estimated models allowing a noisy signal of ability to influence the education choice.

indicating the extent to which countries make an early assignment of students to a college path or not. So, in Germany the index is relatively large, it is smaller in Italy and zero in the US. The correlation of that index with our measure of mismatch is 0.128, the correlation with the under-match rate is 0.581 and the correlation with the over-match rate is -0.413. So in countries where tracking is more extensive there is more mismatch, due to more under-matching. Interestingly, the effect is not symmetric as over-matching is negatively correlated with the measure of tracking. This is consistent with the over-match rate exceeding the under-match rate for both Germany and Italy.

Thus for these countries one might have conjectured that σ_d would be larger, at least relative to the noise in the test score. This is not the case. Apparently early tracking does not generate excessive noise in the educational choice.

These findings are consistent with those reported in Dustmann, Puhani, and Schönberg (2014), who also find, using a very different methodology, no significant long-term effects of tracking in Germany on labor market outcomes. They attribute this to *ex post* flexibility in the German system. In our model, these finding would imply that early tracking would not necessarily generate mismatch.

7.2 Reverse Causality: Allowing the Test Score to Depend on Education

The logistic regression used to estimate the response of the education decision to ability, (5), provides two moments that are central to the estimation. As noted earlier, this regression *per se* does not impose any causality. The structural model that frames the estimation gives an interpretation to that relationship: it reflects the influence of ability on the education decision and the dependence of the test score on ability.

The model imposes a restriction: education does not have a direct affect on the test score. In this sub-section we relax that restriction and allow the test score to reflect both education and ability.[49] We maintain the assumption

[49]Thanks to Marc Henry for discussions on this issue.

that the education choice of the individual also depends on a noisy measure of ability, thus building on sub-section 7.1.

Specifically, consider an extension of (18) in which the test score depends jointly on ability and education, ed_i, where α_{ed} parameterizes this dependence:

$$ts_i = \theta_i + \alpha_{ed} ed_i + \sigma\zeta_i. \tag{23}$$

Recall that this relationship is part of the structure of the model. As the model includes imperfect information on ability in the education decision, the education and ability measures in (23) are not perfectly correlated.

There are two approaches to estimate α_{ed} along with the other structural parameters. The first approach is simply to add α_{ed} to the set of parameters. The parameter estimates for this exercise are reported in the top block of Table 23. The associated moments are shown in Table 24.

A second approach enlarges the set of moments to include a coefficient, denoted ξ_{ed}, from a separate regression of the test score on education in the PIACC data, by country at the individual level. As ability is not observed, it is not possible to estimate (23). Thus, relative to (23), the regression run on the PIACC data suffers from omitted variable bias. Nonetheless, the inclusion of ξ_{ed} as an additional moment is informative about α_{ed} as the same regression is run on both the simulated and PIACC data sets. The resulting parameter estimates are reported in the bottom block of Table 23. The moments for this case are shown in Table 25, including the estimated dependence of the PIACC score on education.

The estimates and moments are quite close to those reported for the case of a noisy education decision, sub-section 7.1. Note too that our main findings of the irrelevance of both taste shocks and borrowing constraints remain. This is true whether ξ_{ed} is included as another moment or not.

As reported in Table 25, the estimates of ξ_{ed} are positive for all of the countries, indicating a positive correlation between test scores and education. The estimates of α_{ed} reported in Table 23 are less than the estimates of ξ_{ed} reflecting omitted variable bias: i.e. the relation between the test score and education in (23) includes both ability and education as covariates. The estimated value of α_{ed} is positive for all the countries in both treatments, except for Italy.[50] The effect of education on the test score is largest in Germany and the US.

The moments reported in the two tables under the "BC" and "No Taste Shock" sub-sections show the role of the borrowing constraints and taste shocks. As in our baseline case, neither the borrowing constraint nor the taste shocks are relevant for matching the moments. Relative to the baseline estimates reported in Table 3, the ordering of the returns to human capital accumulation is about the same, though the return in Italy is lower when ξ_{ed} is added as a moment.

8 Other Implications

This section explores other implications of the baseline model. We decompose the college premium into a selection effect and the return to higher education, quantify the loss in output due to mismatch and study the role of education mismatch for job mismatch.

[50]The estimate of α_{ed} in the bottom block of Table 23 for Italy is negative despite the estimate of $\xi_{ed} = 0.553$. This reflects the omitted variable bias in estimating of ξ_{ed} in the simulated data to match the estimate in the PIACC data.

Table 23: Reverse Causality: Parameter Estimates

	ϕ	$\bar{\varepsilon}$	σ_t	σ_d	$h(\bar{e})$	α_{ed}	$\bar{b}$
				Baseline			
Ger.	2.570	0.965	1.263	0.005	0.809	0.103	na
Italy	2.834	1.123	1.588	0.001	0.728	0.001	na
Jap.	4.126	0.692	0.574	0.001	1.222	0.033	na
US	3.167	1.730	0.706	0.019	1.061	0.179	na
				ξ_{ed} **added as a moment**			
Ger.	2.445	1.325	1.389	0.000	0.791	0.096	na
Italy	2.455	2.275	1.814	0.002	0.661	-0.262	na
Jap.	3.824	0.879	0.758	0.003	1.208	0.129	na
US	3.302	0.697	0.674	0.001	1.075	0.184	na

This table reports parameter estimates for models allowing education to affect test scores.

8.1 Return to Schooling

Table 26 provides an exact decomposition of the college wage premium during the late working phase.[51] The compensation of an agent of ability θ who went to college is $\omega_2 h(\bar{e})\theta$ while an agent not going to college receives only ω_2. The average college wage premium is $h(\bar{e})E(\theta|e = \bar{e})$. The first term is the return to higher education. The second reflects the process of selection into higher education.

Table 26 shows the overall premium and its components by country. These are all calculated from the simulated data. The college premium is highest in the US and lowest in Japan.[52] Interestingly, these relatively small differences in the premia mask larger differences in the returns to education and the selection by ability.

In Italy, at one extreme, the estimated return to education is less than one. But the selection process implies that the average ability of those going to college is much larger than those not attending. Though the return to education is higher, the same pattern emerges in Germany. At the other extreme, the wage premium is very high in the US but the selection on ability going to college is not as strong as in other countries.

It is interesting that the selection effect is strongest in Germany and Italy. Again, these are the countries with early tracking and one might have conjectured that the selection would be weaker.

Looking across the OECD countries, Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) report coefficient estimates on years of schooling in a wage regression that also controls for experience and numeracy score. This does not replicate the decomposition in Table 26. Still the estimated effects of schooling on wages in that study also finds a relatively low return for Italy and Germany and a high return for the US. The estimated return to education in Japan is, in the Hanushek, Schwerdt, Wiederhold, and Woessmann (2015) results, not as high as the US in contrast to the ordering in Table 26.

8.2 Output Maximizing Allocations

In this section, we study the output cost stemming from all sources of mismatch. While some of the apparent mismatch may be an efficient response to differences in tastes across agents, i.e. a high ability person has a strong

[51] The premium is the same as that during the early working phase.
[52] See Daiji and Yuko (2014) for a comparison of the college premium in the US and Japan.

Table 24: Reverse Causality: Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
				Baseline: Reverse				
Ger.	0.347	0.101	0.082	-0.714	1.162	0.222	0.137	0.001
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.082	0.120	0.240	0.860	0.189	0.102	0.003
US	0.418	0.064	0.066	-0.353	1.514	0.257	0.141	0.002
				BC				
Ger.	0.347	0.101	0.082	-0.714	1.162	0.222	0.137	0.001
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.082	0.120	0.240	0.860	0.189	0.102	0.003
US	0.418	0.064	0.066	-0.353	1.514	0.257	0.141	0.002
				No Taste Shocks: $\bar{\varepsilon} = 0$				
Ger.	0.347	0.102	0.082	-0.714	1.164	0.222	0.136	0.001
Italy	0.209	0.143	0.079	-1.507	0.891	0.135	0.064	0.001
Jap.	0.548	0.082	0.120	0.241	0.860	0.189	0.101	0.003
US	0.422	0.065	0.061	-0.353	1.509	0.275	0.147	0.002

This table reports simulated moments for the estimated models allowing education to affect test scores.

dislike for college and so works after high school, this exercise provides an upper bound on the output cost of mismatch.

The data and estimated models produce mismatch. The estimation finds that the mismatch is largely due to mismeasurement of ability through a noisy test score. Still taste shocks are present in some specifications and do influence education choice. Further, as in sub-section 7.1, education decisions may be made without knowing ability.

This section calculates the equivalent of the planner's solution in Section 2.1 for the three phase model and compares that outcome, measured in output net of education cost, to the outcome of two models where mismatch is inefficient. Note that calculating the output loss relative to the baseline is not of interest since the mismatch with those estimates is efficient.

The first comparison is with the model of normal taste shocks and no noise in the test score presented in sub-section 6.3. In this specification, there are taste shocks and thus measured mismatch which is taken, for the sake of this analysis, to be all associated with inefficient schooling choices. This means that the planner bases the education allocation ignoring taste shocks. Thus this is an upper bound on the output loss from mismatch.

These results are in reported the block labeled "Noise in Tastes" in Table 27. For all countries, the efficient solution produces an output gain in the first period indicating that the education rate is excessive in these countries. This reflects the taste shock. There are also output gains in the two working periods. In Japan, for example, output is about 3.3% higher in the second working period from the elimination of the mismatch.

The second comparison is with the specification of noisy education decisions presented in sub-section 7.1. Here mismatch is inefficient relative to a benchmark in which education decisions are made once ability is observed. Thus this exercise, labeled "Imperfect Information" estimates the output loss due to the information friction. For this case, the model has no noise in the test score, just imperfect information at the time of the education decision. In the alternative of perfect information, the noise at the time of the education decision is removed as is the taste shock.

Table 25: Reverse Causality: Moments including ξ_{ed}

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	ξ_{ed}	fit
				Data					
Ger.	0.373	0.104	0.062	-0.720	1.160	0.235	0.144	0.484	na
Italy	0.230	0.146	0.069	-1.510	0.890	0.132	0.071	0.553	na
Jap.	0.597	0.078	0.108	0.230	0.860	0.184	0.111	0.314	na
US	0.455	0.055	0.045	-0.360	1.510	0.279	0.149	0.568	na
				Baseline					
Ger.	0.346	0.104	0.085	-0.714	1.162	0.228	0.144	0.487	0.001
Italy	0.205	0.147	0.091	-1.505	0.877	0.157	0.089	0.583	0.003
Jap.	0.549	0.083	0.117	0.240	0.860	0.191	0.099	0.311	0.003
US	0.419	0.062	0.063	-0.358	1.517	0.250	0.132	0.544	0.003
				BC					
Ger.	0.346	0.104	0.085	-0.714	1.162	0.228	0.144	0.487	0.001
Italy	0.205	0.147	0.091	-1.505	0.877	0.157	0.089	0.583	0.003
Jap.	0.549	0.083	0.117	0.240	0.860	0.191	0.099	0.311	0.003
US	0.419	0.062	0.063	-0.358	1.517	0.250	0.132	0.544	0.003
				No Taste Shocks: $\bar{\varepsilon} = 0$					
Ger.	0.346	0.104	0.085	-0.714	1.161	0.229	0.143	0.486	0.001
Italy	0.206	0.147	0.089	-1.503	0.880	0.160	0.089	0.579	0.003
Jap.	0.549	0.084	0.115	0.240	0.860	0.193	0.098	0.313	0.003
US	0.425	0.062	0.055	-0.355	1.514	0.270	0.132	0.554	0.002

This table reports simulated moments for the estimated models allowing education to affect test scores, including the coefficient from a regression of test score on education.

Table 26: Return to College

	college prem.	$h(\bar{e})$	$E(\theta \| ed = 1)$	$E(\theta \| ed = 0)$
Germany	2.015	0.803	2.509	1.197
Italy	1.961	0.728	2.693	1.243
Japan	1.852	1.227	1.509	1.067
US	2.057	1.056	1.949	1.131

The college premium is the ratio of earnings in the late work phase for agents with college and without and $h(\bar{e})$ is the estimated return to college independent of ability.

These results are in reported the block labeled "Imperfect Information" in Table 27. Here the differences with the efficient solution are surprisingly small. First period output is higher in most countries, noticeably Japan. The second period output gains are negligible for the two countries in which tracking and thus the noisy education decision is most prominent, Germany and Italy.

This result of a small output effect is consistent with the parameter estimates for σ_d reported in Table 21 for Germany and Italy. This reinforces the earlier findings that tracking is not a key source of mismatch. The inefficiencies from these noisier education decisions are not large.

Table 27: Output Net of Education Cost

	Estimated Model			**Efficient Allocation**		
	Ed Phase	Early Work	Late Work	Ed Phase	Early Work	Late Work
	Noise in Tastes					
Germany	0.754	1.071	1.126	0.881	1.109	1.166
Italy	0.815	1.015	1.115	0.946	1.032	1.135
Japan	0.257	1.345	1.377	0.303	1.39	1.423
US	0.442	1.209	1.388	0.636	1.212	1.392
	Imperfect Information					
Germany	0.741	1.349	1.419	0.744	1.349	1.419
Italy	0.807	1.201	1.320	0.806	1.202	1.321
Japan	0.225	1.478	1.483	0.319	1.426	1.460
US	0.413	1.429	1.641	0.434	1.425	1.636

This table shows output loss in the two periods: baseline vs the efficient solution without taste shocks and perfect information about ability.

8.3 Jobs and Skill MisMatch

A recent study by McGowan and Andrews (2015) provides evidence on job mismatch across OECD countries.[53] It builds upon the self-assessment of workers and focuses on the numeracy score of those who are termed "well-matched".[54] From this group of so-called "well-matched", minimum and maximum test scores are determined at the 10% and 90% levels by occupation. Those with scores below the cut-off are deemed as over-matched and those above the cut-off are under-matched.

Our model lacks a job match component. Conditional on education, there are no additional labor market frictions that would create job mismatch independently of education mismatch. So, our framework is unable to independently characterize education and job related mismatch. It can shed some light on a related question: How much of the job mismatch is due to education mismatch?

The education and job mismatch rates for our four countries are presented in Table 28. Some interesting patterns emerge. The job under-match rate is quite high in Italy, both relative to other countries and as a fraction of its overall job mismatch rate. Of the four countries, Italy also has the highest level of education mismatch, again relative to other countries and as fraction of overall education mismatch. It seems that many individuals in Italy are under-matched in education and then in the workplace.

The US has the lowest mismatch rates in both education and in the workplace. In the US there is a slightly higher over-match rate in education and almost two-thirds of the job mismatch in the US is in the form of over-match. It seems that US individuals are under-placed in their jobs and to a lesser degree in the college outcome.

Overall, with the exception of Italy, the total job and education rates of mismatch are quite close. But, for the 21 countries common to both studies, the correlation of the education and skill mismatch measures is only 0.047. However, this masks a relationship between under- and over-matching and skill mismatch. The correlation of the over-match rate in education and skill mismatch is -0.34. And the correlation of the under-match rate in education and skill mismatch is 0.403. Finally, the correlation between the education rate and the skill mismatch rate is -0.444.

[53]The measurement of job mismatch is discussed in detail in their paper.

[54]McGowan and Andrews (2015) presents results for the literacy score in their text. Our analysis is of the numeracy score. The authors kindly provided the data for our calculations.

Table 28: Education and Job Mismatch

	Education			**Job**		
	Total	Under	Over	Total	Under	Over
Germany	0.185	0.094	0.091	0.186	0.016	0.179
Italy	0.222	0.079	0.143	0.198	0.068	0.13
Japan	0.199	0.079	0.120	0.119	0.039	0.08
US	0.135	0.060	0.075	0.133	0.025	0.102

This table shows mismatch in education and jobs. The education mismatch is reported in Table 2. The job mismatch is calculated from the data used in McGowan and Andrews (2015).

9 Conclusion

This paper studies mismatch in education and ability across 21 OECD countries. The education choice highlighted in the analysis is the college decision. Mismatch, in the form of both under- and over-matching occurs across a broad range of countries.

The structural estimation allows for multiple explanations of mismatch: (i) taste heterogeneity, (ii) borrowing restrictions and (iii) noise in the test score. A main finding is that imperfect capital markets are **not** a source of mismatch. Further the contribution of differences in tastes for higher education is minimal. Instead, noise in the test score seems sufficient to explain not only the measured mismatch across countries but other moments that link the education decision to the test score and compensation to the same score. By matching these additional moments, the noise in the test score is “over-identified”. This finding is robust across a number of alternative specifications and moments.

The estimated model is used to study a number of additional issues. The college wage premium is decomposed into the return to higher education and the selection, by ability, into college. Though countries have essentially the same premium, such as the US and Germany, they may differ substantially in the return to college and the selection into higher education. The mismatch in the college choice, as it is due largely to noisy test scores, does not reflect any inefficiency. This is also the case for education systems where college choice is made early in life. Finally, across the 21 OECD countries examined here, we do not find that education mismatch is the source of skill (job) mismatch though these measures are closely linked for three of our four key countries.

10 Appendix

These tables present results for all countries, for the baseline and many of the robustness exercises. These tables are intended as an online Appendix.

10.1 All Countries

Table A1: All Countries: Parameter Estimates

abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
Aus.	2.557	0.900	1.490	0.700	na
Bel	5.971	1.614	0.202	1.073	na
Can.	3.855	0.368	0.577	1.185	na
CzR.	4.786	0.049	0.388	0.966	na
Den.	4.581	0.681	0.516	1.032	na
Est.	3.358	0.693	0.861	1.016	na
Fin.	4.701	0.698	0.491	1.044	na
Fra.	5.583	1.372	0.218	1.063	na
Ger.	2.545	1.354	1.186	0.803	na
Ire.	3.389	0.605	0.626	1.138	na
It.	2.835	0.893	1.586	0.728	na
Jap.	4.243	0.464	0.511	1.227	na
Kor.	4.292	0.164	0.469	1.280	na
Net.	4.169	0.292	0.432	1.020	na
Nor.	5.445	0.716	0.374	1.059	na
Pol.	4.090	0.151	0.507	1.047	na
Slo.	2.833	1.288	1.196	0.824	na
Sp.	3.369	0.440	0.629	0.934	na
Swe.	5.466	1.014	0.338	1.034	na
UK	2.715	0.732	1.240	1.039	na
US	3.137	0.498	0.583	1.056	na
			BC		
Aus.	2.555	1.047	1.492	0.699	3.369
Bel	6.000	1.614	0.200	1.074	3.085
Can.	3.855	0.368	0.577	1.185	2.743
CzR.	4.786	0.049	0.388	0.966	2.622
Den.	4.584	0.681	0.515	1.032	3.403
Est.	3.358	0.693	0.861	1.016	1.038
Fin.	4.701	0.698	0.491	1.044	2.622
Fra.	5.482	1.189	0.225	1.060	1.733
Ger.	2.545	1.354	1.186	0.803	2.634
Ire.	3.389	0.605	0.626	1.138	2.616
It.	2.836	0.893	1.586	0.729	1.638
Jap.	4.243	0.464	0.511	1.227	2.622
Kor.	4.290	0.164	0.469	1.280	2.391
Net.	4.168	0.292	0.432	1.020	2.610
Nor.	5.444	0.716	0.375	1.059	2.685

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
Pol.	4.088	0.152	0.507	1.047	3.306
Slo.	2.833	1.291	1.196	0.824	1.722
Sp.	3.369	0.440	0.629	0.934	2.657
Swe.	5.466	1.014	0.338	1.034	2.633
UK	2.715	0.732	1.240	1.039	2.637
US	3.137	0.498	0.583	1.056	2.622

Table A2: All Countries: Data Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2
Aus.	0.280	0.121	0.065	-1.440	1.120	0.179	0.114
Bel	0.421	0.067	0.048	-0.290	1.570	0.149	0.085
Can.	0.548	0.083	0.098	0.300	0.860	0.193	0.127
CzR.	0.304	0.111	0.040	-1.420	1.460	0.124	0.088
Den.	0.523	0.093	0.105	-0.140	0.840	0.137	0.084
Est.	0.445	0.105	0.087	-0.260	0.820	0.179	0.118
Fin.	0.523	0.102	0.093	-0.040	0.830	0.142	0.075
Fra.	0.456	0.047	0.051	-0.490	1.710	0.174	0.094
Ger.	0.373	0.104	0.062	-0.720	1.160	0.235	0.144
Ire.	0.474	0.079	0.068	-0.270	1.150	0.241	0.134
It.	0.230	0.146	0.069	-1.510	0.890	0.132	0.071
Jap.	0.597	0.078	0.108	0.230	0.860	0.184	0.111
Kor.	0.648	0.083	0.109	0.460	0.880	0.217	0.092
Net.	0.412	0.084	0.046	-0.480	1.290	0.183	0.105
Nor.	0.494	0.081	0.098	-0.150	0.920	0.127	0.073
Pol.	0.429	0.101	0.078	-0.430	1.100	0.191	0.083
Slo.	0.250	0.139	0.039	-1.300	1.090	0.179	0.101
Sp.	0.399	0.079	0.048	-0.660	1.290	0.228	0.098
Swe.	0.491	0.085	0.091	-0.480	1.100	0.121	0.086
UK	0.462	0.088	0.076	-0.110	0.790	0.225	0.158
US	0.455	0.055	0.045	-0.360	1.510	0.279	0.149

Table A3: All Countries: Simulated Moments

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Baseline					
Aus.	0.227	0.129	0.070	-1.431	1.123	0.178	0.097	0.003
Bel	0.428	0.052	0.065	-0.291	1.570	0.157	0.087	0.001
Can.	0.560	0.081	0.125	0.297	0.860	0.200	0.119	0.001
CzR.	0.250	0.100	0.042	-1.413	1.461	0.141	0.059	0.004
Den.	0.468	0.094	0.111	-0.128	0.840	0.140	0.078	0.003
Est.	0.439	0.103	0.112	-0.259	0.822	0.185	0.106	0.001
Fin.	0.489	0.092	0.114	-0.034	0.830	0.139	0.079	0.002
Fra.	0.395	0.053	0.054	-0.480	1.710	0.166	0.089	0.004
Ger.	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
Ire.	0.436	0.080	0.093	-0.263	1.151	0.239	0.129	0.002
It.	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
Kor.	0.596	0.073	0.129	0.470	0.879	0.197	0.113	0.004
Net.	0.395	0.077	0.074	-0.479	1.291	0.184	0.100	0.001
Nor.	0.465	0.087	0.104	-0.143	0.920	0.129	0.068	0.001
Pol.	0.405	0.086	0.087	-0.427	1.100	0.180	0.094	0.001
Slo.	0.247	0.124	0.071	-1.300	1.092	0.183	0.091	0.001
Sp.	0.361	0.086	0.073	-0.654	1.292	0.204	0.117	0.003
Swe.	0.400	0.085	0.082	-0.462	1.100	0.134	0.067	0.009
UK	0.470	0.105	0.123	-0.112	0.796	0.230	0.141	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003
			BC					
Aus.	0.227	0.129	0.070	-1.431	1.124	0.178	0.098	0.003
Bel	0.428	0.052	0.065	-0.292	1.570	0.156	0.086	0.001
Can.	0.560	0.081	0.125	0.296	0.861	0.200	0.118	0.001
CzR.	0.250	0.100	0.042	-1.413	1.461	0.141	0.059	0.004
Den.	0.468	0.094	0.111	-0.128	0.840	0.140	0.078	0.003
Est.	0.439	0.103	0.112	-0.259	0.822	0.185	0.106	0.001
Fin.	0.489	0.092	0.114	-0.034	0.830	0.139	0.079	0.002
Fra.	0.395	0.053	0.053	-0.479	1.710	0.168	0.090	0.004
Ger.	0.345	0.104	0.087	-0.715	1.161	0.225	0.143	0.002
Ire.	0.436	0.080	0.093	-0.263	1.151	0.239	0.129	0.002
It.	0.209	0.143	0.079	-1.507	0.891	0.135	0.065	0.001
Jap.	0.548	0.082	0.122	0.240	0.860	0.189	0.103	0.003
Kor.	0.596	0.073	0.129	0.470	0.879	0.197	0.113	0.004

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Net.	0.395	0.077	0.074	-0.479	1.291	0.184	0.100	0.001
Nor.	0.465	0.087	0.104	-0.144	0.920	0.129	0.068	0.001
Pol.	0.405	0.086	0.087	-0.427	1.100	0.180	0.094	0.001
Slo.	0.247	0.124	0.071	-1.300	1.092	0.183	0.091	0.001
Sp.	0.361	0.086	0.073	-0.654	1.292	0.204	0.117	0.003
Swe.	0.400	0.085	0.082	-0.462	1.100	0.134	0.067	0.009
UK	0.470	0.105	0.123	-0.112	0.796	0.230	0.141	0.003
US	0.414	0.068	0.075	-0.353	1.512	0.261	0.153	0.003

Table A4: All Countries: Perturbations

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
No Borrowing								
Aus.	0.062	0.162	0.009	-3.933	2.042	0.154	0.046	7.122
Bel	0.027	0.178	0.000	-7.458	3.016	0.077	0.009	53.655
Can.	0.253	0.107	0.052	-1.346	1.340	0.211	0.073	3.033
CzR.	0.010	0.192	0.000	-9.512	2.979	0.049	0.004	67.898
Den.	0.088	0.151	0.011	-3.324	1.754	0.116	0.023	11.181
Est.	0.049	0.167	0.003	-4.595	2.162	0.135	0.024	20.776
Fin.	0.087	0.151	0.011	-3.365	1.777	0.115	0.022	12.155
Fra.	0.020	0.184	0.000	-8.533	3.091	0.071	0.007	66.818
Ger.	0.139	0.133	0.025	-2.451	1.852	0.217	0.089	3.534
Ire.	0.436	0.080	0.093	-0.263	1.151	0.239	0.129	0.002
It.	0.030	0.180	0.005	-5.101	1.998	0.094	0.018	14.176
Jap.	0.370	0.094	0.082	-0.612	1.082	0.200	0.081	0.812
Kor.	0.520	0.076	0.111	0.122	0.959	0.205	0.104	0.137
Net.	0.001	0.199	0.000	-22.114	3.214	0.017	0.001	471.957
Nor.	0.064	0.158	0.004	-4.142	2.055	0.094	0.014	17.431
Pol.	0.029	0.177	0.000	-6.483	2.745	0.101	0.013	39.532
Slo.	0.061	0.163	0.008	-4.040	2.036	0.151	0.038	8.444
Sp.	0.052	0.161	0.001	-5.174	2.772	0.153	0.033	22.715
Swe.	0.057	0.160	0.002	-4.617	2.289	0.093	0.014	18.733
UK	0.283	0.119	0.081	-1.063	1.062	0.235	0.107	1.017
US	0.328	0.077	0.053	-0.856	1.717	0.264	0.136	0.306
No Taste Shocks								
Aus.	0.226	0.129	0.070	-1.439	1.127	0.178	0.097	0.003
Bel***	0.417	0.049	0.055	-0.356	1.706	0.158	0.083	0.023

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Can.	0.560	0.081	0.125	0.296	0.862	0.200	0.118	0.001
CzR.	0.250	0.100	0.042	-1.415	1.462	0.141	0.058	0.004
Den.	0.466	0.094	0.110	-0.138	0.845	0.141	0.078	0.003
Est.	0.438	0.103	0.111	-0.263	0.825	0.185	0.105	0.001
Fin.	0.487	0.091	0.113	-0.043	0.839	0.140	0.078	0.002
Fra. **	0.388	0.051	0.049	-0.521	1.788	0.167	0.087	0.012
Ger.	0.343	0.103	0.086	-0.724	1.175	0.226	0.142	0.002
Ire.	0.435	0.080	0.092	-0.264	1.155	0.239	0.129	0.002
It.	0.209	0.143	0.079	-1.509	0.892	0.135	0.064	0.001
Jap.	0.547	0.082	0.122	0.237	0.859	0.189	0.103	0.003
Kor.	0.596	0.073	0.129	0.469	0.879	0.197	0.113	0.004
Net.	0.395	0.076	0.074	-0.481	1.294	0.184	0.100	0.001
Nor.	0.462	0.087	0.102	-0.157	0.931	0.130	0.068	0.001
Pol.	0.405	0.086	0.087	-0.426	1.098	0.180	0.094	0.001
Slo.	0.246	0.124	0.070	-1.309	1.096	0.183	0.091	0.001
Sp.	0.361	0.086	0.073	-0.656	1.291	0.204	0.117	0.003
Swe.	0.396	0.084	0.080	-0.482	1.123	0.134	0.066	0.010
UK	0.469	0.106	0.122	-0.117	0.796	0.230	0.141	0.003
US	0.413	0.068	0.075	-0.354	1.516	0.262	0.153	0.003
			No Noise in Test					
Aus.	0.227	0.000	0.000	-6.558	58.917	0.270	0.184	3366.697
Bel	0.428	0.000	0.000	1.699	9.864	0.201	0.139	72.761
Can.	0.560	0.000	0.000	36.903	98.572	0.301	0.211	10887.436
CzR.	0.250	0.000	0.000	-81.945	403.622	0.211	0.137	168218.746
Den.	0.468	0.000	0.000	8.890	33.488	0.238	0.177	1147.477
Est.	0.439	0.000	0.000	12.455	56.368	0.301	0.214	3247.328
Fin.	0.489	0.000	0.000	9.442	31.928	0.235	0.175	1057.016
Fra.	0.395	0.000	0.000	1.533	12.631	0.211	0.144	123.372
Ger.	0.345	0.000	0.000	4.466	46.215	0.302	0.217	2056.900
Ire.	0.436	0.000	0.000	15.520	71.438	0.329	0.212	5189.735
It.	0.209	0.002	0.000	-9.074	46.242	0.255	0.174	2114.073
Jap.	0.548	0.000	0.000	24.699	66.936	0.293	0.195	4964.765
Kor.	0.596	0.000	0.000	76.987	181.517	0.295	0.198	38486.079
Net.	0.395	0.000	0.000	13.915	91.273	0.257	0.180	8304.258
Nor.	0.465	0.000	0.000	6.424	24.871	0.213	0.154	616.910
Pol.	0.405	0.000	0.000	32.767	192.817	0.269	0.184	37857.426
Slo.	0.247	0.000	0.000	-3.438	36.258	0.287	0.186	1241.404

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Sp.	0.361	0.000	0.000	8.908	83.644	0.281	0.200	6873.699
Swe.	0.400	0.000	0.000	2.371	17.004	0.207	0.147	261.088
UK	0.470	0.000	0.000	22.434	92.955	0.344	0.240	9002.707
US	0.414	0.000	0.000	18.079	96.689	0.326	0.217	9398.981

10.2 Young Cohort

Table A5: Young Cohort: Parameter Estimates

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
		Young Baseline			
Aus.	2.690	0.805	1.103	0.667	
Bel	5.991	1.770	0.200	1.083	
Can.	3.596	0.115	0.706	1.150	
CzR.	5.357	0.609	0.311	1.015	
Den.	3.230	1.163	1.369	0.956	
Est.	3.421	0.694	0.768	1.002	
Fin.	3.600	0.744	0.936	0.935	
Fra.	5.581	1.057	0.204	1.071	
Ger.	2.085	0.019	2.536	0.703	
Ire.	3.456	0.636	0.552	1.118	
It.	2.844	1.195	1.646	0.761	
Jap.	4.296	0.678	0.484	1.244	
Kor.	3.052	0.173	1.251	1.233	
Net.	4.272	0.237	0.438	1.015	
Nor.	5.252	0.431	0.455	1.037	
Pol.	3.928	0.240	0.594	1.065	
Slo.	3.264	2.095	0.824	0.914	
Sp.	3.115	0.417	0.772	0.895	
Swe.	5.284	0.951	0.365	1.015	
UK	2.824	0.639	1.130	1.029	
US	3.157	0.556	0.603	1.068	
		Young, BC			
Aus.	2.684	0.999	1.108	0.666	2.577
Bel	5.999	1.747	0.200	1.083	3.622
Can.	3.596	0.115	0.706	1.150	2.661
CzR.	5.362	0.581	0.311	1.015	2.763

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
Den.	3.230	1.157	1.369	0.956	2.874
Est.	3.422	0.654	0.767	1.002	5.121
Fin.	3.600	0.742	0.936	0.935	2.727
Fra.	5.654	0.965	0.201	1.074	7.842
Ger.	2.085	0.020	2.536	0.703	6.788
Ire.	3.456	0.636	0.552	1.118	2.685
It.	2.850	1.284	1.638	0.762	2.978
Jap.	4.301	0.563	0.484	1.244	3.074
Kor.	3.051	0.171	1.251	1.233	2.857
Net.	4.064	0.002	0.470	1.003	7.931
Nor.	4.767	1.907	0.507	1.015	6.875
Pol.	3.936	0.297	0.592	1.066	3.008
Slo.	3.264	2.096	0.824	0.914	3.558
Sp.	3.114	0.216	0.773	0.895	5.262
Swe.	5.307	1.032	0.363	1.015	2.847
UK	2.757	0.373	1.179	1.023	3.758
US	3.139	0.252	0.611	1.066	6.196

Table A6: Young Cohort Moments: Data and Simulated

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
			Data					
Aus.	0.248	0.120	0.034	-1.890	1.540	0.179	0.114	na
Bel	0.441	0.086	0.034	-0.150	1.510	0.149	0.085	na
Can.	0.486	0.084	0.080	0.130	0.810	0.193	0.127	na
CzR.	0.360	0.110	0.035	-1.240	1.490	0.124	0.088	na
Den.	0.507	0.135	0.129	-0.180	0.530	0.137	0.084	na
Est.	0.417	0.115	0.077	-0.380	0.930	0.179	0.118	na
Fin.	0.418	0.122	0.107	-0.490	0.720	0.142	0.075	na
Fra.	0.479	0.043	0.046	-0.390	1.830	0.174	0.094	na
Ger.	0.332	0.146	0.072	-0.840	0.880	0.235	0.144	na
Ire.	0.429	0.070	0.064	-0.430	1.340	0.241	0.134	na
It.	0.241	0.147	0.096	-1.310	0.800	0.132	0.071	na
Jap.	0.621	0.104	0.115	0.370	0.860	0.184	0.111	na
Kor.	0.685	0.144	0.142	0.500	0.530	0.217	0.092	na
Net.	0.397	0.100	0.033	-0.580	1.250	0.183	0.105	na
Nor.	0.448	0.099	0.108	-0.330	0.820	0.127	0.073	na

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Pol.	0.452	0.106	0.085	-0.210	0.930	0.191	0.083	na
Slo.	0.299	0.125	0.038	-1.090	1.130	0.179	0.101	na
Sp.	0.364	0.104	0.051	-0.760	1.230	0.228	0.098	na
Swe.	0.454	0.093	0.096	-0.620	1.100	0.121	0.086	na
UK	0.420	0.094	0.065	-0.220	0.830	0.225	0.158	na
US	0.474	0.070	0.044	-0.300	1.400	0.279	0.149	na
				Baseline				
Aus.	0.184	0.124	0.040	-1.883	1.544	0.176	0.094	0.005
Bel	0.453	0.050	0.072	-0.155	1.510	0.157	0.090	0.003
Can.	0.522	0.090	0.123	0.120	0.815	0.199	0.115	0.004
CzR.	0.278	0.091	0.044	-1.229	1.492	0.141	0.058	0.009
Den.	0.459	0.129	0.139	-0.169	0.530	0.138	0.082	0.002
Est.	0.413	0.099	0.101	-0.380	0.929	0.190	0.106	0.001
Fin.	0.390	0.120	0.113	-0.484	0.720	0.140	0.077	0.001
Fra.	0.410	0.045	0.052	-0.379	1.829	0.172	0.094	0.005
Ger.	0.316	0.131	0.108	-0.836	0.884	0.213	0.137	0.002
Ire.	0.403	0.075	0.077	-0.425	1.340	0.243	0.128	0.001
It.	0.236	0.143	0.091	-1.309	0.800	0.136	0.067	0.000
Jap.	0.578	0.077	0.127	0.379	0.857	0.190	0.108	0.003
Kor.	0.618	0.113	0.156	0.513	0.525	0.191	0.123	0.008
Net.	0.378	0.082	0.073	-0.576	1.252	0.175	0.091	0.003
Nor.	0.426	0.102	0.105	-0.326	0.830	0.119	0.059	0.001
Pol.	0.449	0.091	0.105	-0.210	0.930	0.179	0.098	0.001
Slo.	0.283	0.112	0.073	-1.089	1.132	0.184	0.091	0.002
Sp.	0.341	0.095	0.075	-0.757	1.231	0.206	0.118	0.002
Swe.	0.373	0.090	0.079	-0.602	1.101	0.132	0.065	0.008
UK	0.445	0.104	0.116	-0.224	0.835	0.224	0.133	0.004
US	0.427	0.070	0.082	-0.289	1.403	0.258	0.151	0.004
				BC				
Aus.	0.184	0.124	0.040	-1.882	1.543	0.176	0.094	0.005
Bel	0.453	0.050	0.072	-0.157	1.510	0.157	0.089	0.003
Can.	0.522	0.090	0.123	0.120	0.815	0.199	0.115	0.004
CzR.	0.278	0.091	0.044	-1.228	1.492	0.141	0.058	0.009
Den.	0.459	0.129	0.139	-0.169	0.531	0.138	0.082	0.002
Est.	0.413	0.099	0.101	-0.380	0.930	0.190	0.106	0.001
Fin.	0.390	0.120	0.113	-0.484	0.720	0.140	0.077	0.001
Fra.	0.410	0.045	0.051	-0.379	1.830	0.171	0.092	0.005

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Ger.	0.316	0.131	0.108	-0.836	0.884	0.213	0.137	0.002
Ire.	0.403	0.075	0.077	-0.425	1.341	0.243	0.128	0.001
It.	0.236	0.143	0.091	-1.309	0.800	0.136	0.067	0.000
Jap.	0.577	0.077	0.127	0.377	0.859	0.190	0.108	0.003
Kor.	0.618	0.113	0.156	0.513	0.525	0.191	0.123	0.008
Net.	0.377	0.082	0.074	-0.575	1.253	0.181	0.097	0.003
Nor.	0.426	0.103	0.108	-0.325	0.820	0.129	0.070	0.001
Pol.	0.449	0.091	0.104	-0.211	0.931	0.178	0.097	0.001
Slo.	0.283	0.112	0.073	-1.089	1.132	0.184	0.091	0.002
Sp.	0.341	0.095	0.075	-0.756	1.231	0.206	0.118	0.002
Swe.	0.373	0.090	0.079	-0.604	1.100	0.131	0.064	0.008
UK	0.446	0.105	0.117	-0.223	0.837	0.228	0.137	0.004
US	0.426	0.070	0.082	-0.291	1.402	0.258	0.151	0.004
			No Taste; $\bar{\varepsilon} = 0$					
Aus.	0.184	0.124	0.040	-1.888	1.553	0.176	0.094	0.005
Bel	0.438	0.046	0.060	-0.234	1.665	0.160	0.086	0.033
Can.	0.523	0.090	0.123	0.124	0.816	0.199	0.115	0.004
CzR.	0.276	0.091	0.042	-1.246	1.508	0.141	0.058	0.009
Den.	0.457	0.129	0.138	-0.179	0.535	0.138	0.081	0.003
Est.	0.412	0.099	0.101	-0.384	0.935	0.190	0.106	0.001
Fin.	0.389	0.119	0.112	-0.492	0.725	0.140	0.076	0.001
Fra.	0.407	0.044	0.049	-0.399	1.880	0.173	0.093	0.008
Ger.	0.316	0.131	0.108	-0.836	0.885	0.213	0.137	0.002
Ire.	0.402	0.075	0.077	-0.428	1.343	0.243	0.128	0.001
It.	0.235	0.143	0.091	-1.313	0.801	0.136	0.067	0.000
Jap.	0.576	0.077	0.126	0.372	0.865	0.190	0.108	0.003
Kor.	0.618	0.113	0.156	0.513	0.525	0.191	0.123	0.008
Net.	0.377	0.082	0.073	-0.577	1.250	0.175	0.091	0.003
Nor.	0.426	0.101	0.104	-0.328	0.840	0.119	0.059	0.001
Pol.	0.450	0.091	0.104	-0.210	0.931	0.179	0.098	0.001
Slo.	0.279	0.111	0.068	-1.123	1.167	0.185	0.089	0.004
Sp.	0.341	0.095	0.075	-0.757	1.234	0.206	0.118	0.002
Swe.	0.369	0.090	0.077	-0.623	1.120	0.132	0.064	0.009
UK	0.445	0.104	0.116	-0.226	0.838	0.224	0.133	0.004
US	0.427	0.070	0.081	-0.288	1.412	0.258	0.151	0.004
			No Noise; $\sigma = 0$					
Aus.	0.184	0.019	0.000	-13.651	54.143	0.246	0.167	2905.440

	college	under-match	over-match	α_0	α_1	ν_1	ν_2	fit
Bel	0.453	0.000	0.000	1.904	8.988	0.202	0.141	60.155
Can.	0.522	0.000	0.000	110.729	338.883	0.312	0.217	126525.574
CzR.	0.278	0.000	0.000	-3.941	28.095	0.205	0.131	715.124
Den.	0.459	0.000	0.000	8.625	35.449	0.293	0.222	1296.931
Est.	0.413	0.000	0.000	9.939	53.469	0.295	0.207	2866.903
Fin.	0.390	0.000	0.000	6.411	42.867	0.265	0.198	1824.067
Fra.	0.410	0.000	0.000	2.598	16.505	0.214	0.146	224.288
Ger.	0.316	0.000	0.000	421.804	5947.227	0.295	0.206	35537671.469
Ire.	0.403	0.000	0.000	10.960	64.077	0.323	0.204	4065.666
It.	0.236	0.000	0.000	-4.642	37.289	0.266	0.182	1342.616
Jap.	0.578	0.000	0.000	18.108	44.933	0.290	0.196	2257.142
Kor.	0.618	0.000	0.000	130.867	337.293	0.346	0.251	130404.853
Net.	0.378	0.000	0.000	13.070	108.837	0.251	0.174	11761.282
Nor.	0.426	0.000	0.000	8.399	42.727	0.215	0.155	1832.390
Pol.	0.449	0.000	0.000	29.870	125.080	0.280	0.195	16318.186
Slo.	0.283	0.000	0.000	-0.914	18.142	0.282	0.186	289.463
Sp.	0.341	0.000	0.000	7.496	94.731	0.287	0.203	8810.699
Swe.	0.373	0.000	0.000	1.642	18.404	0.208	0.148	304.579
UK	0.445	0.000	0.000	21.195	96.201	0.338	0.233	9554.334
US	0.427	0.000	0.000	17.674	86.497	0.327	0.219	7564.661

10.3 Parent's Education

Table A7: Parental Education: Parameter Estimates

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
			Baseline		
Aus.	2.645	5.449	1.368	0.736	
Bel	6.094	1.672	0.213	1.091	
Can.	3.701	2.055	0.660	1.179	
CzR.	5.535	3.740	0.285	1.038	
Den.	4.309	2.554	0.621	1.023	
Est.	3.178	2.505	1.022	1.006	
Fin.	4.739	1.173	0.498	1.057	
Fra.	5.657	2.434	0.213	1.097	
Ger.	2.509	5.588	1.309	0.799	
Ire.	3.429	3.591	0.621	1.172	

	abil	$\bar{\varepsilon}$	σ	$h(\bar{e})$	$\bar{b}$
It.	4.163	6.442	0.635	0.942	
Jap.	4.091	3.146	0.572	1.225	
Kor.	4.261	1.721	0.483	1.302	
Net.	4.240	2.658	0.420	1.037	
Nor.	5.218	1.494	0.422	1.055	
Pol.	4.433	3.371	0.443	1.106	
Slo.	3.408	6.870	0.771	0.952	
Sp.	3.619	2.734	0.542	0.990	
Swe.	5.242	1.916	0.381	1.027	
UK	2.852	5.272	1.248	1.080	
US	3.114	3.439	0.619	1.055	
			BC		
Aus.	2.645	5.449	1.368	0.736	2.622
Bel	6.109	1.667	0.212	1.091	2.017
Can.	3.701	2.055	0.660	1.179	2.642
CzR.	5.535	3.740	0.285	1.038	5.708
Den.	4.304	2.557	0.623	1.023	1.064
Est.	3.178	2.505	1.022	1.006	2.800
Fin.	4.739	1.173	0.498	1.057	2.646
Fra.	5.657	2.434	0.213	1.097	2.988
Ger.	2.504	5.617	1.316	0.798	2.946
Ire.	3.429	3.591	0.621	1.172	2.637
It.	4.163	6.442	0.635	0.942	3.062
Jap.	4.091	3.146	0.572	1.225	2.622
Kor.	4.261	1.721	0.483	1.302	2.620
Net.	4.240	2.658	0.420	1.037	2.622
Nor.	5.218	1.494	0.422	1.055	1.609
Pol.	4.433	3.371	0.443	1.106	3.435
Slo.	3.408	6.870	0.771	0.952	2.651
Sp.	3.619	2.734	0.542	0.990	2.584
Swe.	5.242	1.916	0.381	1.027	0.996
UK	2.570	6.050	1.558	1.058	5.576
US	3.114	3.439	0.619	1.055	4.520

Table A8: PE Moments: Data and Simulated

	college	under-match	over-match	α_0	α_1	α_2	ν_1	ν_2	fit
				Data					
Aus.	0.280	0.107	0.051	-1.720	1.030	0.970	0.179	0.114	na
Bel	0.421	0.051	0.043	-0.640	1.440	1.090	0.149	0.085	na
Can.	0.548	0.080	0.084	-0.040	0.800	0.780	0.193	0.127	na
CzR.	0.304	0.079	0.035	-1.740	1.360	1.700	0.124	0.088	na
Den.	0.523	0.073	0.093	-0.570	0.740	1.050	0.137	0.084	na
Est.	0.445	0.097	0.087	-0.540	0.740	0.680	0.179	0.118	na
Fin.	0.523	0.088	0.085	-0.190	0.780	0.580	0.142	0.075	na
Fra.	0.456	0.037	0.047	-0.750	1.560	1.540	0.174	0.094	na
Ger.	0.373	0.101	0.042	-1.200	1.030	1.120	0.235	0.144	na
Ire.	0.474	0.074	0.064	-0.520	1.060	1.060	0.241	0.134	na
It.	0.230	0.131	0.053	-1.730	0.860	1.970	0.132	0.071	na
Jap.	0.597	0.064	0.086	-0.270	0.800	1.310	0.184	0.111	na
Kor.	0.648	0.079	0.106	0.300	0.830	0.890	0.217	0.092	na
Net.	0.412	0.074	0.049	-0.820	1.220	1.030	0.183	0.105	na
Nor.	0.494	0.063	0.089	-0.470	0.850	0.750	0.127	0.073	na
Pol.	0.429	0.082	0.067	-0.660	0.990	1.520	0.191	0.083	na
Slo.	0.250	0.109	0.035	-1.600	0.990	1.670	0.179	0.101	na
Sp.	0.399	0.078	0.042	-0.780	1.210	0.910	0.228	0.098	na
Swe.	0.491	0.077	0.090	-0.890	1.020	0.890	0.121	0.086	na
UK	0.462	0.069	0.063	-0.480	0.650	1.430	0.225	0.158	na
US	0.455	0.049	0.026	-0.760	1.390	0.890	0.279	0.149	na
				Baseline					
Aus.	0.262	0.121	0.056	-1.717	1.032	0.973	0.179	0.108	0.001
Bel	0.467	0.050	0.061	-0.649	1.442	1.086	0.151	0.085	0.002
Can.	0.570	0.081	0.115	-0.045	0.803	0.778	0.197	0.121	0.002
CzR.	0.340	0.076	0.029	-1.745	1.359	1.697	0.140	0.073	0.002
Den.	0.490	0.093	0.093	-0.562	0.741	1.054	0.137	0.083	0.002
Est.	0.453	0.105	0.105	-0.541	0.742	0.680	0.184	0.110	0.001
Fin.	0.519	0.089	0.113	-0.191	0.782	0.581	0.137	0.080	0.001
Fra.	0.486	0.041	0.050	-0.756	1.561	1.537	0.167	0.101	0.001
Ger.	0.366	0.101	0.065	-1.198	1.033	1.121	0.219	0.150	0.001
Ire.	0.493	0.072	0.084	-0.524	1.061	1.058	0.235	0.139	0.001
It.	0.358	0.097	0.033	-1.753	0.856	1.955	0.129	0.081	0.019
Jap.	0.577	0.072	0.096	-0.266	0.800	1.313	0.184	0.110	0.001
Kor.	0.650	0.065	0.124	0.298	0.830	0.890	0.192	0.119	0.002

	college	under-match	over-match	α_0	α_1	α_2	ν_1	ν_2	fit
Net.	0.430	0.071	0.066	-0.825	1.220	1.028	0.182	0.107	0.001
Nor.	0.478	0.087	0.098	-0.466	0.853	0.753	0.128	0.070	0.001
Pol.	0.510	0.068	0.070	-0.677	0.991	1.512	0.173	0.106	0.008
Slo.	0.350	0.097	0.040	-1.618	0.989	1.659	0.178	0.107	0.011
Sp.	0.425	0.075	0.072	-0.786	1.212	0.906	0.202	0.124	0.003
Swe.	0.410	0.086	0.073	-0.874	1.020	0.899	0.132	0.069	0.008
UK	0.535	0.089	0.093	-0.506	0.677	1.363	0.209	0.140	0.013
US	0.426	0.069	0.069	-0.754	1.394	0.893	0.257	0.155	0.004
				BC					
Aus.	0.262	0.121	0.056	-1.717	1.032	0.973	0.179	0.108	0.001
Bel	0.467	0.050	0.061	-0.649	1.441	1.086	0.151	0.084	0.002
Can.	0.570	0.081	0.115	-0.045	0.803	0.778	0.197	0.121	0.002
CzR.	0.340	0.076	0.029	-1.745	1.359	1.697	0.140	0.073	0.002
Den.	0.490	0.093	0.093	-0.562	0.742	1.054	0.137	0.083	0.002
Est.	0.453	0.105	0.105	-0.541	0.742	0.680	0.184	0.110	0.001
Fin.	0.519	0.089	0.113	-0.191	0.782	0.581	0.137	0.080	0.001
Fra.	0.486	0.041	0.050	-0.755	1.561	1.537	0.167	0.101	0.001
Ger.	0.366	0.101	0.065	-1.199	1.032	1.122	0.220	0.150	0.001
Ire.	0.493	0.072	0.084	-0.524	1.061	1.058	0.235	0.139	0.001
It.	0.358	0.097	0.033	-1.753	0.856	1.955	0.129	0.081	0.019
Jap.	0.577	0.072	0.096	-0.266	0.800	1.313	0.184	0.110	0.001
Kor.	0.650	0.065	0.124	0.298	0.830	0.890	0.192	0.119	0.002
Net.	0.430	0.071	0.066	-0.825	1.220	1.028	0.182	0.107	0.001
Nor.	0.478	0.087	0.098	-0.466	0.853	0.753	0.128	0.070	0.001
Pol.	0.510	0.068	0.070	-0.677	0.991	1.512	0.173	0.106	0.008
Slo.	0.350	0.097	0.040	-1.618	0.989	1.659	0.178	0.107	0.011
Sp.	0.425	0.075	0.072	-0.786	1.212	0.906	0.202	0.124	0.003
Swe.	0.410	0.086	0.073	-0.874	1.020	0.899	0.132	0.069	0.008
UK	0.543	0.090	0.094	-0.497	0.655	1.423	0.223	0.156	0.008
US	0.426	0.069	0.069	-0.754	1.394	0.893	0.257	0.155	0.004
				No Taste; $\bar{\varepsilon} = 0$					
Aus.	0.245	0.126	0.073	-1.305	1.099	0.000	0.180	0.099	1.120
Bel	0.452	0.052	0.070	-0.172	1.485	0.000	0.153	0.081	1.410
Can.	0.562	0.087	0.130	0.297	0.794	0.000	0.197	0.118	0.724
CzR.	0.301	0.083	0.046	-1.077	1.516	0.000	0.145	0.061	3.355
Den.	0.470	0.103	0.117	-0.121	0.758	0.000	0.138	0.077	1.309
Est.	0.444	0.110	0.119	-0.235	0.750	0.000	0.184	0.107	0.557

	college	under-match	over-match	α_0	α_1	α_2	ν_1	ν_2	fit
Fin.	0.513	0.092	0.120	0.073	0.790	0.000	0.137	0.079	0.407
Fra.	0.458	0.043	0.065	-0.115	1.652	0.000	0.171	0.094	2.784
Ger.	0.344	0.109	0.091	-0.717	1.084	0.000	0.221	0.139	1.494
Ire.	0.476	0.078	0.102	-0.071	1.082	0.000	0.238	0.132	1.327
It.	0.291	0.115	0.076	-1.038	1.008	0.000	0.135	0.060	4.386
Jap.	0.556	0.086	0.128	0.269	0.798	0.000	0.187	0.104	2.011
Kor.	0.643	0.072	0.139	0.682	0.821	0.000	0.193	0.117	0.941
Net.	0.416	0.074	0.079	-0.369	1.261	0.000	0.184	0.101	1.267
Nor.	0.467	0.092	0.108	-0.135	0.866	0.000	0.129	0.068	0.677
Pol.	0.477	0.078	0.100	-0.077	1.028	0.000	0.177	0.096	2.656
Slo.	0.303	0.108	0.074	-0.971	1.109	0.000	0.183	0.089	3.203
Sp.	0.413	0.077	0.083	-0.376	1.245	0.000	0.203	0.119	0.996
Swe.	0.397	0.090	0.086	-0.472	1.045	0.000	0.133	0.066	0.977
UK	0.505	0.109	0.133	0.036	0.684	0.000	0.211	0.130	2.322
US	0.416	0.071	0.079	-0.346	1.428	0.000	0.258	0.150	0.970
			No Noise; $\sigma = 0$						
Aus.	0.262	0.001	0.000	-5.084	19.948	6.640	0.273	0.199	401.397
Bel	0.467	0.000	0.000	1.262	18.767	6.337	0.201	0.140	331.380
Can.	0.570	0.000	0.000	10.721	37.617	6.389	0.306	0.219	1502.780
CzR.	0.340	0.000	0.000	-4.136	9.363	6.519	0.200	0.138	93.033
Den.	0.490	0.000	0.000	2.141	19.936	6.355	0.245	0.189	404.028
Est.	0.453	0.000	0.000	4.679	34.839	6.277	0.309	0.224	1221.335
Fin.	0.519	0.000	0.000	9.453	38.303	6.460	0.235	0.177	1535.551
Fra.	0.486	0.000	0.000	0.354	14.256	6.301	0.212	0.151	185.092
Ger.	0.366	0.000	0.000	-1.140	24.397	6.585	0.302	0.230	575.896
Ire.	0.493	0.000	0.000	3.423	24.976	6.391	0.327	0.222	615.974
It.	0.358	0.001	0.000	-3.930	7.774	6.404	0.226	0.179	72.360
Jap.	0.577	0.000	0.000	4.481	20.753	6.325	0.293	0.206	445.884
Kor.	0.650	0.000	0.000	14.301	37.237	6.269	0.290	0.204	1550.485
Net.	0.430	0.000	0.000	0.436	20.171	6.405	0.255	0.184	389.631
Nor.	0.478	0.000	0.000	3.566	25.220	6.260	0.218	0.161	640.594
Pol.	0.510	0.000	0.000	0.906	14.800	6.264	0.258	0.189	215.705
Slo.	0.350	0.000	0.000	-3.599	11.056	6.593	0.277	0.204	129.610
Sp.	0.425	0.000	0.000	1.257	24.630	6.376	0.277	0.204	582.554
Swe.	0.410	0.000	0.000	-0.484	19.282	6.393	0.211	0.154	363.976
UK	0.535	0.000	0.000	3.339	23.757	6.329	0.339	0.254	572.541
US	0.426	0.000	0.000	2.174	29.026	6.356	0.326	0.223	802.258

	college	under-match	over-match	α_0	α_1	α_2	ν_1	ν_2	fit

References

Adda, J., and R. Cooper (2003): *Dynamic economics: quantitative methods and applications*. The MIT Press.

Arum, R., and J. Roksa (2011): "Limited Learning on College Campuses," *Society*, pp. 203–207.

Brunello, G., and D. Checchi (2007): "Does school tracking affect equality of opportunity? New international evidence," *Economic Policy*, 22(52), 782–861.

Daiji, K., and M. Yuko (2014): "Winning the Race against Technology," Discussion paper, Bank of Japan, Working Paper #14.E.5.

Dillon, E. W., and J. A. Smith (2013): "The determinants of mismatch between students and colleges," Discussion paper, National Bureau of Economic Research.

Döbert, H. (2015): "Germany," *The education systems of Europe*, pp. 305–333.

Dustmann, C. (2004): "Parental background, secondary school track choice, and wages," *Oxford Economic Papers*, 56(2), 209–230.

Dustmann, C., P. A. Puhani, and U. Schönberg (2014): "The long-term effects of school quality on labor market outcomes and educational attainment," Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London CReAM Discussion Paper Series, forthcoming Economic Journal.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015): "Returns to Skills around the World: Evidence from PIAAC," *European Economic Review*, 73, 103–130.

Jones, C. I. (2015): "Pareto and Piketty: The macroeconomics of top income and wealth inequality," *Journal of Economic Perspectives*, 29, 29–46.

Keane, M. P., and K. I. Wolpin (2001): "The effect of parental transfers and borrowing constraints on educational attainment," *International Economic Review*, 42(4), 1051–1103.

Kim, M. (2013): "Relaxing Occupational Misallocation: Are Education Subsidies Effective?," Penn. State University, Thesis Chapter.

Lochner, L. J., and A. Monge-Naranjo (2011): "The Nature of Credit Constraints and Human Capital," *American Economic Review*, 101(6), 2487–2529.

McGowan, M., and D. Andrews (2015): "Skill mismatch and Public Policy in OECD Countries," Discussion paper, OECD, ECO/WKP(2015)28.

Smith, J., M. Pender, and J. Howell (2013): "The full extent of student-college academic undermatch," *Economics of Education Review*, 32, 247–261.

Spence, M. (1973): "Job market signaling," *The quarterly journal of Economics*, pp. 355–374.

Weiss, A. (1983): "A sorting-cum-learning model of education," *The Journal of Political Economy*, pp. 420–442.