CES ifo Working Papers

12740
2026

June 2026

Imperfect Self-knowledge about Skills and Skill Mismatch

Daniel Goller, Enzo Brox, Stefan C. Wolter

CES ifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
https://www.cesifo.org

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge
· from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
· from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
· from the RePEc website: https://ideas.repec.org/s/ces/ceswps.html

Imperfect Self-knowledge about Skills and Skill Mismatch

Daniel Goller[1,2] ⓡ Enzo Brox[1,2] Stefan C. Wolter[1,2,3,4]

[1]University of Bern

[2]Swiss Leading House VPET-ECON

[3] Swiss Coordination Centre for Research in Education

[4] CESifo, IZA@LISER, RF Berlin

June 2, 2026

Abstract

Why do people sort into poorly fitting occupations? This paper shows that imperfect self-knowledge about skills is an important source of skill mismatch at labor market entry. We use unique data from standardized professional aptitude tests linked to administrative records on educational trajectories and early labor market outcomes in Switzerland. The data allow us to observe objective skills and subjective skill beliefs for many productivity-relevant skills in a high-stakes setting. We document large differences among individuals in how well their beliefs align with their skills. Imperfect self-knowledge predicts misaligned occupational aspirations, higher realized skill mismatch, and a higher probability of dropout. Guided by a Roy-style model of occupational choice with imperfect self-knowledge, we interpret these findings as evidence that distorted self-assessments at the school-to-work transition contribute to the misallocation of talent.

Keywords: Information frictions, Occupational choice, Skill mismatch, Self-knowledge

JEL: D83, J24, J41

*The order of the two first authors is randomized. Declarations of interest: none. We would like to express our sincere thanks to gateway.one, for providing us with their data and for the time and effort they put into preparing and explaining it. We also thank the Swiss State Secretariat for Education, Research and Innovation for financial support through its Leading House on the Economics of Education and the Federal Statistical Office for merging the test data with administrative records. Data availability: The data supporting the study's findings are confidential and subject to third-party restrictions.
Corresponding author: Daniel Goller, Ph.D., Centre for Research in Economics of Education, University of Bern, Schanzeneckstrasse 1, 3012 Bern, Switzerland. +41 31 684 3269. daniel.goller@unibe.ch. https://orcid.org/0000-0002-3879-0744

1 Introduction

The allocation of workers across jobs and occupations is central for both individual labor market trajectories and firm performance (Gautier & Teulings, 2015; Lise & Postel-Vinay, 2020). Skill mismatch – the gap between workers' skills and the skill requirements of their job – raises the probability of unemployment (Şahin, Song, Topa, & Violante, 2014) and creates substantial wage penalties over the life cycle (Fredriksson, Hensvik, & Skans, 2018; Eckardt, 2026). For firms, mismatches between worker skills and job requirements reduce productivity and lead to inefficient allocation of talent (Hsieh, Hurst, Jones, & Klenow, 2019; Guvenen, Kuruscu, Tanaka, & Wiczer, 2020). Understanding the sources of skill mismatch is therefore important for explaining labor market outcomes and aggregate productivity.

Skill mismatch can arise from several sources, including search and matching frictions (Şahin et al., 2014), institutional features and labor market rigidities (Moscarini & Postel-Vinay, 2012), technological change that alters task and skill demands (Autor, Levy, & Murnane, 2003; Acemoglu & Autor, 2011), and information frictions that affect learning and decision-making (Fredriksson et al., 2018). Information frictions can generate occupational mismatch through at least three distinct channels. First, employers may imperfectly observe workers' skills, giving rise to employer learning, screening, and signaling dynamics (Heller & Kessler, 2024). Second, workers may misperceive the skill requirements or returns associated with different occupations, leading to poorly informed occupational choices (Arcidiacono, Hotz, & Kang, 2012; Arcidiacono, Hotz, Maurel, & Romano, 2020; Wiswall & Zafar, 2015). Third, workers may have imperfect self-knowledge about their own productivity-relevant skills (Carranza, Garlick, Orkin, & Rankin, 2022).

The third channel, imperfect self-knowledge about one's own skills, has received comparatively little attention. Despite a large literature documenting systematic errors in self-assessments of ability and performance, often framed as overconfidence or miscalibration (Camerer & Lovallo, 1999; Benoît & Dubra, 2011; Santos-Pinto & de la Rosa, 2020; Bobba & Frisancho, 2022), there is little evidence on the extent of self-knowledge in productivity-relevant skills and whether imperfect self-knowledge affects high-stakes occupational choices and the resulting skill mismatch (Carranza et al., 2022).

In this study, we address this gap. While previous studies approximate information frictions or self-knowledge with age or tenure (Fredriksson et al., 2018; Eckardt, 2026), we construct a direct measure of self-knowledge about productivity skills at the time of labor market entry. To do so, we exploit information from standardized professional aptitude tests. The tests assess skills in multiple productivity-relevant skill domains that extend substantially beyond school knowledge. After each skill domain (e.g. *technical understanding*), students report their subjective beliefs about their performance. Objective test scores and subjective beliefs are recorded on the same scale. Leveraging repeated within-individual observations across skill domains, we construct an individual-level measure of imperfect self-knowledge that identifies persistent heterogeneity in self-knowledge net of domain-specific noise and measurement error. Moreover, because the assessed competencies are directly productivity-relevant, self-knowledge about these skills is directly relevant for occupational choice and sorting.

To study the role of imperfect self-knowledge during the school-to-work transition, we link this measure to several administrative data sources capturing school trajectories, occupational preferences, realized occupational choices, and early labor market outcomes. To measure skill mismatch, we combine test scores with expert-based information on occupational skill requirements within the same skill domains. The combination of test and administrative data offers us two unique advantages. First, matching the test data to administrative records avoids sample attrition to a large degree. Second, and most importantly, because the students' subjective beliefs are reported in the certificate, it guarantees that the self-assessment is not just guesswork on the part of the test-takers.

Our empirical setting is the transition from lower secondary school to either general upper-secondary education or vocational education and training (VET) in Switzerland. At age 14 or 15, two-thirds of adolescents choose one of roughly 250 occupations offered through the dual apprenticeship system, while just around 30% choose a general education track. These early occupational choices strongly shape subsequent human capital accumulation and labor market trajectories (Hensvik, Müller, & Skans, 2023; Bruhn et al., 2025). Recent evidence suggests that occupational mismatch is particularly costly in apprenticeship systems with highly specialized vocational training, where workers employed outside their training occupation experience substantial wage penalties (Eckardt, 2026). Hence, it is particularly important to understand the initial occupational choice.

At the same time, these decisions are made at a stage when individuals may still have imperfect knowledge about their own productivity-relevant skills and comparative advantage across occupations. The Swiss VET system, therefore, provides a natural setting to study whether imperfect self-knowledge contributes to occupational mismatch and the misallocation of talent.

We document four main findings. First, students with greater imperfect self-knowledge about their skills prefer occupations that are less aligned with their objective skill profiles. Second, imperfect self-knowledge not only shows up in students' preferences for occupations but also predicts realized labor market outcomes: students with greater imperfect self-knowledge are more likely to be employed in occupations that are less well aligned with their skills. The misaligned occupational preferences almost entirely explain this relationship. Third, individuals with greater imperfect self-knowledge are more likely to revise their occupational choices when they receive more precise signals about their abilities. They are less likely to enter the occupation they initially aspired to, and switching tends to improve match quality, especially for those with the highest levels of imperfect self-knowledge. Fourth, imperfect self-knowledge also predicts dropout – an indicator of early labor market performance. A meaningful part of this effect can be explained by skill mismatch; however, a substantial residual association remains unexplained by skill mismatch. This suggests that imperfect self-knowledge captures additional important components for the human capital formation process, beyond operating through information frictions in line with the literature on the importance of beliefs for human capital accumulation (Bénabou & Tirole, 2002; Santos-Pinto & Sobel, 2005; Wiswall & Zafar, 2021).

To interpret these patterns, we extend a Roy-style model of occupational choice to allow for imperfect self-knowledge of skills. In the canonical Roy framework (Roy, 1951; Heckman & Sedlacek, 1985; Borjas, 1987), individuals sort into occupations based on comparative advantage, implicitly assuming that workers know their own productivity. We relax this assumption by allowing individuals to hold noisy beliefs about their occupation-specific fit when making their initial choice. In the model, individuals first choose an occupation based on imperfect self-knowledge, then receive more precise signals about their abilities, and may revise their choice subject to switching costs. Heterogeneity in the precision of self-knowledge generates clear predictions that guide the empirical analysis: individuals with greater imperfect self-knowledge are more likely to enter mismatched occupations, more likely to revise their choices when additional information becomes available, and, because adjustment is incomplete, remain more mismatched on average in realized outcomes.

1.1 Related Literature

This study contributes to several strands of the literature on occupational sorting, belief formation, and information frictions in the labor market.

A growing literature studies the allocation of workers across occupations and the consequences of mismatch between workers' skills and job requirements (Lise & Postel-Vinay, 2020; Eckardt, 2026). Recent work combines direct skill measures with dynamic models to show that mismatch can generate persistent earnings losses and inefficient allocation of talent (Papageorgiou, 2014; Guvenen et al., 2020). Fredriksson et al. (2018) show that information frictions are an important source of skill mismatch by showing that mismatch substantially decreases with labor market experience or prior firm experience. However, directly measuring information frictions remains challenging. We complement this work by combining direct measures of skills and subjective beliefs across a broad set of productivity-relevant domains with expert-assessed occupational requirements. This allows us to provide a direct measure of one important source of information frictions, imperfect self-knowledge about one's own skills, and to study how it relates to a task-based measure of skill mismatch rather than relying on skill sets of incumbent workers.

Another strand of literature examines how subjective beliefs affect human capital investment and field-of-study choices. Arcidiacono et al. (2012) show that beliefs about ability and expected returns help explain college major choice, while Wiswall and Zafar (2015) show that students revise choices in response to information about major-specific outcomes. Other studies show that providing information about academic performance changes beliefs and affects schooling decisions (Bobba & Frisancho, 2022; Stinebrickner & Stinebrickner, 2014; Hakimov, Schmacker, & Terrier, 2023; Arcidiacono, Aucejo, Maurel, & Ransom, 2025; Brox, Davoli, & Strazzeri, 2025). These studies focus primarily on educational investments. We instead study imperfect self-knowledge about multidimensional skills at the transition from school to work, when individuals make occupational choices under substantial uncertainty and when early decisions have persistent effects on labor market outcomes (Bruhn et al., 2025; Hensvik et al., 2023; Eckardt, 2026).

Furthermore, a growing body of literature studies the role of information frictions in job search and hiring. Theoretical and empirical works show that limited information about skills, job characteristics, or match quality can distort search behavior and lead to inefficient matching (Jovanovic, 1979; Spinnewijn, 2015; Conlon, Pilossoph, Wiswall, & Zafar, 2018). Field experiments further show that providing information about workers' skills or labor market prospects affects beliefs, search behavior, and employment outcomes (Pallais, 2014; Abel, Burger, & Piraino, 2020; Groh, Krishnan, McKenzie, & Vishwanath, 2016; Belot, Kircher, & Muller, 2018; Carranza et al., 2022; Kiss, Garlick, Orkin, & Hensel, 2023). These studies focus on job search after labor market entry and study how new information changes search behavior and labor market outcomes. In contrast, we study imperfect self-knowledge about skills before the first occupational choice is made. Focusing on the school-to-work transition allows us to examine how belief errors affect initial sorting across occupations. Because we observe both occupational preferences and realized occupations, we can measure mismatch already at the preference stage and quantify how distortions in occupational preferences translate into realized mismatch in the labor market.

Finally, this study is also related to a broader literature on self-perceptions and belief formation. A large body of work documents systematic biases in self-assessments of ability, skills, and preferences, and studies how individuals update beliefs in response to new information (Bénabou & Tirole, 2002; Camerer & Lovallo, 1999; Falk, Kosse, Schildberg-Hörisch, & Zimmermann, 2023). However, only a few studies systematically measure imperfect self-knowledge and link it to realized economic outcomes.

2 Model and Hypotheses

2.1 Setting

We model individuals' first occupational choice as follows. Each individual observes noisy signals about their own fits to an occupation (match qualities) and aspires or chooses an occupation $j_i^{(1)}$. Thereafter, each individual receives a more precise signal about their fit. Given the updated information, individuals may revise their choice and select a (possibly different) occupation $j_i^{(2)}$, incurring a switching cost if $j_i^{(2)} \neq j_i^{(1)}$.

Our starting point is a Roy-type model of occupational choice (Roy, 1951), as formalized in parametric form by, among others, Borjas (1987) and Heckman and Sedlacek (1985). Each individual has a vector of occupation-specific potential outcomes $(\theta_{ij})_{j \in J}$ and selects the occupation that maximizes expected match quality given the available information. In contrast to the original Roy framework, where individuals are typically assumed to know their productivity, we explicitly model imperfect self-information about occupation-specific fit in the spirit of models with imperfect self-knowledge and learning about match quality (e.g., Bénabou and Tirole (2002); Jovanovic (1979); Neal (1999)).

Consider a population of risk-neutral individuals who choose among occupations. Each individual i selects an occupation j from a finite set $J = \{1, \ldots, \bar{J}\}$. The payoff from working in occupation j is given by an occupation-specific match-quality (or "fit") parameter $\theta_{ij} \in \mathbb{R}$.

Assumption 1 (Match qualities). For each occupation $j \in J$, the match quality of individual i in occupation j is

$$\theta_{ij} \sim \mathcal{N}(\mu_j, \sigma_\theta^2)$$

independently across i and j. The means μ_j, capturing the average fit in occupation j for a randomly drawn individual, and the common variance $\sigma_\theta^2 > 0$, capturing how dispersed individual fits are around that average, are known to all agents.

Assumption 2 (Types and Information Frictions). Individuals differ in how well they are informed about their own abilities before the first occupational choice. This heterogeneity in self-information is captured by the precision of the signals they observe. Without loss of generality, each individual i is of type $t_i \in \{H, L\}$, where H denotes a high-information type and L a low-information type. Types differ only in the precision of their initial self-signals: high-information types have more precise signals about their own abilities. Formally, there exist $0 < \sigma_H^2 < \sigma_L^2$ such that type t observes signals with noise variance σ_t^2 before the first occupational choice. We interpret a larger noise variance σ_t^2 as stronger information frictions about one's own abilities.[1]

Period 1: Initial Self-Information under Information Frictions and First Occupational Choice. Quantities with superscript (0) refer to information that is available before the first occupational choice in period 1. Before choosing an occupation, individual i of

[1]We follow the literature on imperfect self-knowledge and self-confidence in treating ability or fit as a latent state about which individuals hold noisy private beliefs (Bénabou & Tirole, 2002; Bodner & Prelec, 2003). Relative to this literature, we take signal precision as an exogenous individual characteristic and abstract from strategic manipulation of beliefs, focusing instead on how given information frictions shape early labor market outcomes.

type t_i observes, for each $j \in J$, a noisy signal $s_{ij}^{(0)}$ about their match quality θ_{ij}:

$$s_{ij}^{(0)} = \theta_{ij} + \varepsilon_{ij}^{(t_i)}, \qquad \varepsilon_{ij}^{(t_i)} \sim \mathcal{N}(0, \sigma_{t_i}^2), \tag{1}$$

independently across j. Conditional on $(\theta_{ij})_{j \in J}$, the noises $(\varepsilon_{ij}^{(t_i)})_{j \in J}$ are independent and identically distributed normal shocks.

Because both the underlying match quality θ_{ij} and the noise term are normally distributed, the individual's best estimate of θ_{ij} after observing $s_{ij}^{(0)}$ is a weighted average of the signal and the occupation-specific average fit μ_j:

$$\hat{\theta}_{ij}^{(0)}(t_i, s_{ij}^{(0)}) \; = \; \mathbb{E}[\theta_{ij} \mid s_{ij}^{(0)}, t_i] \; = \; w_{t_i} s_{ij}^{(0)} + (1 - w_{t_i})\mu_j, \tag{2}$$

where

$$w_t \; = \; \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_t^2} \in (0, 1) \tag{3}$$

is the weight placed on the individual-specific signal by type $t \in \{H, L\}$. Because $\sigma_H^2 < \sigma_L^2$, we have $w_H > w_L$: high-information types rely more on their self-signals and less on the common occupation-specific benchmark μ_j than low-information types do.

In period 1, individuals make their initial occupational decision. Since utility is linear in match quality and future payoffs are proportional to θ_{ij}, it is without loss of generality to assume that each individual chooses the occupation with the highest estimated match quality.[2] Hence, the first occupational choice satisfies

$$j_i^{(1)} \in \arg\max_{j \in J} \hat{\theta}_{ij}^{(0)}(t_i, s_{ij}^{(0)}). \tag{4}$$

Partial Elimination of Information Frictions Between the initial and the second occupational decision, the information frictions about occupation-specific fit are partially resolved. This step is modeled based on job-matching and dynamic career-choice models, in which match quality is gradually revealed through noisy performance signals.[3] In those models, individuals

[2]This is a static Roy-style selection decision, but based on noisy self-assessment of occupation-specific productivity rather than perfect knowledge of it, embedded into a multi-period learning framework with heterogeneous information frictions.

[3]These signals can be interpreted as feedback from performance, evaluations, and learning about one's comparative advantage that generalizes across occupations (for example, through a general-ability factor), as in dynamic models of job and career choice with learning about match quality and ability (Jovanovic, 1979; Neal, 1999; Arcidiacono, 2004) and in empirical work documenting mismatch and

typically learn about the fit to the currently chosen job or major. Here, we adopt a simplified reduced-form version in which individuals receive more precise signals about their fit across all occupations, induced by a generalized increase in self-knowledge of their abilities.

Following their initial choice, each individual observes, for every $j \in J$, an additional signal, e.g., after working in the initial occupation or testing one's fit,

$$s_{ij}^{(1)} = \theta_{ij} + \eta_{ij}, \qquad \eta_{ij} \sim \mathcal{N}(0, \sigma_1^2),$$

independently across j and independent of the initial noise. The variance σ_1^2 is strictly smaller than both σ_H^2 and σ_L^2, so these signals are strictly more precise than the initial signals for all types: $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$.

Because the underlying match qualities and both noise terms are normally distributed, combining the initial and the additional signal leads to an updated estimate of θ_{ij} given by

$$\hat{\theta}_{ij}^{(1)}(t_i, s_{ij}^{(0)}, s_{ij}^{(1)}) \;=\; \mathbb{E}[\theta_{ij} \mid s_{ij}^{(0)}, s_{ij}^{(1)}, t_i] \;=\; \frac{\mu_j/\sigma_\theta^2 + s_{ij}^{(0)}/\sigma_{t_i}^2 + s_{ij}^{(1)}/\sigma_1^2}{1/\sigma_\theta^2 + 1/\sigma_{t_i}^2 + 1/\sigma_1^2}. \tag{5}$$

The additional signal makes the estimates $\hat{\theta}_{ij}^{(1)}$ more precise for all types and thus reflects a reduction in information frictions.

Period 2: Re-Evaluating Occupational Choice In period 2, individuals may revise their occupational choice based on the improved estimates $(\hat{\theta}_{ij}^{(1)})_{j\in J}$. If individual i keeps the occupation chosen initially, $j_i^{(1)}$, the expected future payoff is proportional to the updated match quality in that occupation,

$$U_i^{\text{stay}} = \hat{\theta}_{ij_i^{(1)}}^{(1)}. \tag{6}$$

If instead the individual switches to occupation $k \neq j_i^{(1)}$, the expected payoff is

$$U_i^{\text{switch}}(k) = \hat{\theta}_{ik}^{(1)} - c, \tag{7}$$

where $c \geq 0$ captures psychological and social switching costs, such as having mentally committed to the initial choice, having already invested in occupation-specific preparation or skills, or facing family and network ties and contacts that make changing plans more difficult.

information frictions in the labor market (Fredriksson et al., 2018).

The second occupational choice $j_i^{(2)}$ (switching occupations if $j_i^{(2)} \neq j_i^{(1)}$) solves

$$j_i^{(2)} \in \arg\max_{k \in J} \left\{ \hat{\theta}_{ik}^{(1)} - c \cdot \mathbf{1}\{k \neq j_i^{(1)}\} \right\}. \tag{8}$$

2.2 Model Predictions and Hypotheses

To formalize the hypotheses, we first define mismatch and switching in the model. For each individual i, the true optimal occupation is

$$j_i^* \in \arg\max_{j \in J} \theta_{ij}.$$

The individual is initially (finally) mismatched if the first (second) occupational choice differs from the optimal occupation, $j_i^{(1)} \neq j_i^*$ ($j_i^{(2)} \neq j_i^*$). Define the mismatch and switching indicators as:

$$M_i = \mathbf{1}\{j_i^{(1)} \neq j_i^*\}, \qquad M_i^{final} = \mathbf{1}\{j_i^{(2)} \neq j_i^*\}, \qquad S_i = \mathbf{1}\{j_i^{(2)} \neq j_i^{(1)}\}.$$

In the empirical analysis, the switching indicator corresponds directly to this definition: individuals are coded as switchers if they move from their initial occupational choice to a different occupation (or any other pathway) after the information update. The mismatch indicators are not directly observed, but can be proxied by a skill-based mismatch index constructed from observed outcomes, such as the distance between occupational skill requirements and individual skills.[4] Formal lemmas and proofs are provided in Appendix B.1 - B.5. As long as the empirical index increases with the underlying degree of mismatch, the qualitative comparative statics derived carry over (see Appendix B.6).

The model yields four testable hypotheses.

Hypothesis 1 (Information frictions and initial mismatch). Individuals with higher information frictions are more likely to choose an initially mismatched occupation.

Hypothesis 2 (Information frictions and switching). Conditional on receiving more precise information about their abilities, individuals with higher initial information frictions are more likely to switch away from their initial occupational choice.

[4] For overviews of empirical approaches to measuring mismatch and overeducation, see, for example, McGuinness (2006) and Leuven and Oosterbeek (2011).

Hypothesis 3 (Information frictions and gains from switching). Among individuals who switch occupations, those with higher initial information frictions experience larger expected improvements in match quality.

Hypothesis 4 (Information frictions and realized mismatch). Even after the information update and the possibility to switch, individuals with higher initial information frictions remain more mismatched on average.

While the model and the hypotheses are stated in terms of information frictions more generally, the empirical application focuses on one specific and empirically identifiable source of such frictions, namely, imperfect self-knowledge about one's own skills. In our data, we observe both objective skill assessments and subjective beliefs across a broad set of productivity-relevant domains, which allows us to construct a direct measure of this friction. This provides a natural mapping from the theoretical framework to the empirical setting and enables a direct test of the model's predictions.

3 Empirical Setting

3.1 The Career Choice Process

We study the career-choice process of adolescents in Switzerland. Starting in lower secondary school, career orientation is integrated into the curriculum through lectures, time allocated for trial apprenticeships, visits to career fairs, and access to career guidance centers, among other activities. After grade 11, students face a pivotal decision at the transition from ISCED level 2 to 3, at which they either choose to pursue general education or continue their upper-secondary education by entering the labor market through vocational education in one of roughly 250 occupations within Switzerland's dual apprenticeship system. Thus, by age 14 or 15, virtually every Swiss student is actively engaged in the career choice process. Vocational education is highly recognized in Switzerland and provides well-paid career opportunities; about two-thirds of Swiss students, throughout all ability levels, choose the vocational education pathway (Hanushek, Schwerdt, Woessmann, & Zhang, 2017; OECD, 2025).

Guided by their preferences and beliefs about their own skills, individuals choose an occupation and apply to firms offering apprenticeships in that field. A common, though not mandatory, component of these applications is a certificate from a standardized professional aptitude test, which forms the core of our data (see Section 3.2). If a firm and an applicant agree on an apprenticeship contract, the individual is employed by the company and acquires occupation-specific skills on the job while attending vocational school for theoretical instruction 1-2 days per week. Upon successful completion, the individual receives a uniformly recognized and valued federal diploma certifying expert qualification in the occupation. At the same time, the diploma is also an educational upper-secondary qualification that allows apprentices to pursue further education at the tertiary level.

3.2 The Professional Aptitude Test

The professional aptitude test is a standardized (paid) assessment offered by a private company (gateway.one) and is comparable in format to the GMAT or GRE.[5] Unlike those exams, test-takers specify the occupation for which they wish to be assessed. It is important to note that these tests are usually not taken by students who are just beginning to collect information about occupations (at the beginning of the career choice process), but are taken once they have decided on an occupation and are considering applying (at an advanced stage of the career choice process). The test is organized into three areas – school knowledge, cognitive potential, and professional knowledge – each comprising multiple modules (e.g., mathematics or English within school knowledge; for a full list of modules by test-type, see Appendix Table A.1). Modules are further divided into submodules (e.g., geometry and approximate calculations within mathematics; for a full list of the submodules, see Appendix Table A.2). After completing a module, test-takers self-assess their performance and proceed to the next module. These self-assessments, together with the module and submodule scores, are printed on the certificate, creating strong incentives to report them accurately.

Besides testing general skills, such as school knowledge, tests are tailored to the requirements of the specific occupation. For example, while every test contains the *logic* or *concentration* modules, *imagination* is contained in the test for Information Technologist or Draughtsman/-woman and *basic graphic skills* is contained in tests for Graphic Designer or Premedia Specialist, for example. Thus, especially the professional knowledge part makes the test "occupation specific".

[5]gateway.one is by far the largest provider of such standardized tests in Switzerland. At about CHF 100 (about USD 130) per test, the fee is not prohibitively high, but it is certainly too high for simple trial-and-error.

In Section 3.4, we explain how we construct our individual specific measure of imperfect self-knowledge and how we adjust our measure of imperfect self-knowledge for the fact that students partly answer different modules.

The widespread adoption of this test is driven by employers' demand for comparable skill assessments. In the absence of such tests, employers must rely on school transcripts with teacher- and school-specific grades, rendering them non-comparable across applicants. To mitigate this lack of comparability, many, though not all, employers require this specific external test as a common evaluative signal. We argue that students who take the test constitute an approximately representative sample of the population of applicants for apprenticeship positions after grade 11, supporting its use for empirical analysis.[6]

3.3 Data

More formally, our primary data consists of standardized professional aptitude tests belonging to distinct test types k. Each individual i is assigned to exactly one test-type $k(i)$ (the occupation in which the test-taker is tested). We treat the test-type occupation as the initial occupational choice or the revealed occupational preference. Each test-type comprises a fixed bundle of modules $m \in \mathcal{M}_k$. Modules appear across multiple test-types, forming an overlapping bipartite network in which the individual-module incidence structure overlaps. For each individual-module observation (i, m), we observe a (module) result $score_{im} \in \{0, 1, .., 100\}$ and a self-assessed perception of the score $self_{im} \in \{0, 10, .., 100\}$. These two variables form the basis of our measure of imperfect self-knowledge about skills, the construction of which is formalized in Section 3.4.

Complementarily, we observe expert-assessed occupational skill requirements. For a test occupation j and skill module m, we observe a requirement measure $req_{mj} \in \{0, 1, \ldots, 100\}$ that describes the level of ability required in occupation j for the skill tested in module m. These requirement profiles were developed by the test provider in collaboration with domain experts, including vocational trainers, vocational school teachers, and career counselors. We use this information, along with student scores, to construct measures of skill mismatch in occupational preferences (test occupation) and in chosen occupations (realized mismatch). A

[6]We cannot directly test this, because the population of those applying to apprenticeship positions after grade 11 is unknown. However, we show in Table A.3 that our test sample is close in observable characteristics to the population of 11th graders.

limitation we face is that we do not have information on requirements in all occupations. We cover a substantial share of the market, but not the entire market. We therefore complement the analysis with a second source of information on occupational requirements. This alternative dataset provides requirement measures for all occupations in the domains of first language and mathematics and is publicly available at `anforderungsprofile.ch`. We use this measure to assess the robustness of our results (Table A.6). The construction of the mismatch measures based on these requirement profiles is described in Section 3.5.

The final dataset contains one observation for each individual who took a professional aptitude test between January 2021 and May 2025. The dataset is restricted to test-takers who were enrolled in 10th grade in one of the cohorts from 2019/20 to 2022/23, who consented to linking their test data to administrative records (approximately 85% of all test-takers), and for whom we can link occupational requirement information.[7] Restricting the sample to 10th-graders ensures that all individuals, to a greater or lesser extent, follow the regular career choice process described in Section 3.1. This restriction primarily excludes a small number of older test-takers, who may have different incentives to take the test and different information about their skills.

The administrative data contains information on individuals' prior (and subsequent) educational trajectories and personal characteristics, such as nationality, gender, first language, and age. All of this information was recorded before the test. Outcome-related information collected after the test date includes early labor market outcomes, such as the chosen occupation and whether the individual drops out of the apprenticeship. In total, the final sample consists of 25,412 observations.[8]

Table A.4 in the Appendix presents the key variables in the combined data set, along with descriptive statistics.

[7]The test date is independent of the year the individuals went to 10th grade, as it is an individual decision at what point to take the test. An individual's first test execution is used in the main sample if they have taken the test more than once. Using only the first test execution is because, after the first test, incentives may differ, and the individual has had a chance to update their beliefs about their own skills.

[8]To investigate specific hypotheses, the sample had to be restricted. For example, in analyses of early labor market outcomes, the sample is restricted to individuals who were observed in the labor market for at least 1 year (N=11,048). We mention any sampling restrictions where they occur.

3.4 Imperfect Self-Knowledge about own Skills

To quantify imperfect self-knowledge of skills, we exploit the joint observation of objective test scores and self-assessed perceptions within individuals across modules to construct two related measures of individual self-knowledge. Variant A is a summary measure based on absolute miscalibration. Variant B is obtained from a two-way fixed-effects specification in the spirit of the "AKM" decomposition of Abowd, Kramarz, and Margolis (1999).

Each module m shows an absolute difference in score and perception of individual i as

$$d_{im} \;=\; \big|\, score_{im} - self_{im} \,\big|, \tag{9}$$

where $score_{im}$ is the score, $self_{im}$, the self-assessed perception of the score of individual i in module m as described in Section 3.3. In Table A.5, we provide some module-level descriptive statistics of the raw data. In total, each individual completes, on average, 12 modules, resulting in almost 300,000 module-level observations. The average absolute deviation across all modules between $score_{im}$ and $self_{im}$ is roughly 16.

As a natural benchmark, we define individual self-knowledge as the mean absolute difference between perceived and realized scores across all modules completed by individual i. Let $\mathcal{M}_i$ denote the set of modules in which individual i is tested. Variant A is then

$$U_i^A \;=\; \frac{1}{\mathcal{M}_i} \sum_{m \in \mathcal{M}_i} d_{im}. \tag{10}$$

This measure is transparent and easy to interpret. However, it conflates an individual component with a composition effect: individuals select themselves into one of the test types, each comprising a different bundle of modules, and modules may differ systematically in how difficult they are to self-assess. As a result, U_i^A reflects both a person-specific component and the particular mix of modules in $\mathcal{M}_i$.

To isolate a person-specific component of imperfect self-knowledge, we estimate the following additive two-way fixed-effects model:

$$d_{im} \;=\; \alpha_i \;+\; \mu_m \;+\; \varepsilon_{im}, \tag{11}$$

where α_i denotes an individual fixed effect, μ_m a module fixed effect,[9] and ε_{im} is an idiosyncratic error term with zero mean. The parameter α_i captures individual i's average tendency to misjudge their performance across the modules they complete, after netting out systematic differences in how difficult each module is to assess. Conversely, μ_m captures how difficult module m is to assess accurately, holding the composition of test-takers fixed.

Equation (11) is formally analogous to two-way fixed-effects wage models in matched worker-firm data, which regress wages on worker and firm fixed effects to disentangle individual and workplace heterogeneity (Abowd et al., 1999; Card, Heining, & Kline, 2013). Similar specifications have been used to separate rater generosity from instructor quality in student evaluations of teaching (Ayllón, Lefgren, Patterson, Stoddard, & Urdaneta, 2025). Identification of α_i and μ_m relies on variation along both dimensions: α_i is identified from within-module comparisons across individuals who take the same module, while μ_m is identified from within-individual comparisons across the modules that a given individual completes.

We interpret the estimated individual effects $\hat{\alpha}_i$ from (11) as Variant B, our preferred measure of individual imperfect self-knowledge about one's own ability. Relative to the simple mean absolute miscalibration U_i^A, the AKM-style Variant B explicitly decomposes observed self-knowledge into an individual component α_i and a module component μ_m. It therefore corrects for the fact that individuals face different module bundles and isolates a persistent, module-adjusted trait of self-knowledge, rather than conflating this trait with the specific set of modules an individual happened to encounter. Higher values indicate that individual i exhibits greater imperfect self-knowledge.

Figure 1 plots the raw (non-standardized) distributions of our imperfect self-knowledge measures. Panel (a) shows the distribution of Variant B (AKM-type), while Panel (b) displays Variant A. In Panel (b), the blue histogram shows the overall distribution of Variant A, while the yellow and red histograms show the same measure computed separately for different skill domains. The yellow distribution corresponds to skills tested in school (school knowledge), whereas the red distribution corresponds to non-school skills (cognitive potential and professional knowledge).

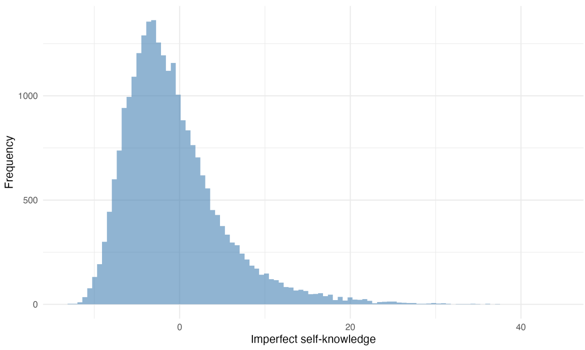
The domain-specific patterns are intuitive. Imperfect self-knowledge is substantially lower

[9] In fact, we regard modules themselves as different in the different test-types; the module fixed effect is, therefore, a module by test-type fixed effect. However, in practice, the resulting information frictions measures are highly correlated ($\approx .99$) for extracting module FE vs. module by test-type FE vs. module by test-type by year FE.

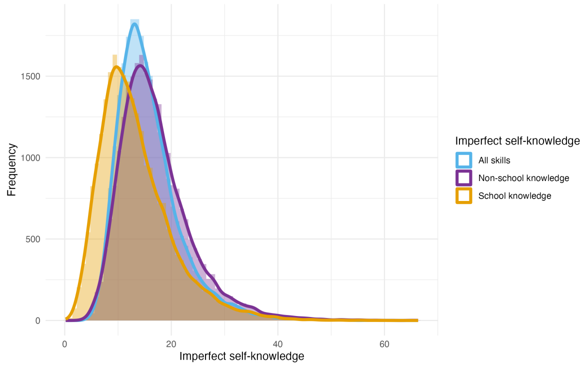
for school-related skills, for which students typically receive frequent and structured feedback through grades and exams, than for non-school skills, for which feedback is less frequent and less standardized. This pattern provides two pieces of validation for our measure. First, the differences across domains are consistent with the idea that imperfect self-knowledge reflects information frictions rather than noise, as bias is smaller in domains with more feedback. Second, the domain variation suggests that imperfect self-knowledge is not fixed but depends on the information environment, consistent with the view that self-knowledge is malleable and can improve when individuals receive more precise signals about their skills.

In the empirical analysis, we use standardized versions of the imperfect self-knowledge measures (mean 0, standard deviation 1).

Figure 1: Imperfect self-knowledge



(a) Distribution of the Variant B (AKM-type)



(b) Distribution of the Variant A, by skills

Notes: Distributions of the Variant A (on the right) and Variant B (on the left) measures for imperfect self-knowledge, as described in Section 3.4. Measures are not standardized. N=25,412. School and Non-school measures use all school-knowledge skills, respectively, non-school-knowledge skills in constructing the measures. In b) overlaid kernel density estimates (Gaussian kernel), scaled to histogram frequencies.

3.5 Skill Mismatch

For our empirical analysis, we measure skill mismatch at two stages. First, skill mismatch in the initial occupational choice (i.e., the model's period 1 choice). To measure occupational preferences, we elicit them through a revealed-preference mechanism and use occupation-specific test choices as our measure of the initial occupational choice. Second, realized skill mismatch (i.e., the model's period 2 choice). We measure this using the realized apprenticeship occupation from the administrative labor market data.

We define the mismatch between an individual and an occupation in module m as the absolute difference between the individual's raw score in module m and the required score in the occupation j.

$$m_{im}^{j} = |score_{im} - req_{m}^{j}| \tag{12}$$

m_{im}^{j} is the mismatch score for the individual i to the occupation j in the module m. $score_{im}$ is the score, req_{m}^{j} is the occupation-specific requirement in the module m. In the module-specific mismatch measure, we use the absolute difference between the achieved score and the requirements. Higher values imply a stronger mismatch in this module.

The mismatch m_{i}^{j} to an occupation j is the weighted sum of the matching scores of the modules m_{im}^{j}, and the higher the mismatch score, the worse the fit to the occupation, and vice versa for lower values.

$$m_{i}^{j} = \sum_{m=1}^{M} w_{m}^{j} \cdot m_{im}^{j} \tag{13}$$

$$\sum_{m=1}^{M} w_{m}^{j} = 1 \tag{14}$$

The weights w_{m}^{j} represent the importance of a skill tested in module m for occupation j and should sum to 1. For example, if all skills are equally important in the occupation j^{*}, then $w_{m}^{j^{*}} = \frac{1}{M} \;\; \forall \;\; m = 1, ..., M$. For our constructions, we have expert-based weights ranging from 1 (lowest importance) to 5 (highest importance) for each occupation, normalized to sum to 1. For the realized or hypothetical skill mismatch, only overlapping skills (modules) are used; i.e., skills that are not tested but would have been in the realized occupation are omitted from the calculation. In the case of missing skills, the weights are scaled proportionally so that they sum to 1. For the analysis, we use the mismatch variables in a standardized form, with a mean of zero and a standard deviation of one (see Table A.3).

3.6 Empirical Analysis

To test our four hypotheses from Section 2.2, we investigate the impact of imperfect self-knowledge of one's own skills on various career decision outcomes. Formally, we estimate the following function:

$$Y_i = \beta \cdot D_i + g(X_i) + \epsilon_i, \tag{15}$$

where D_i is the measure for imperfect self-knowledge about own skills (in the primary analyses: the AKM-based Variant B, compare Section 3.4), and Y_i is the outcome variable. In Hypothesis 1 and 4, the outcome is the initial ($m_i^{j_i^{(1)}}$) and realized mismatch ($m_i^{j_i^{(2)}}$); in Hypothesis 3, the difference in initial and realized mismatch ($m_i^{j_i^{(1)}} - m_i^{j_i^{(2)}}$); in Hypothesis 2, whether the individual changes away from the test occupation ($\mathbf{1}\{j_i^{(2)} \neq j_i^{(1)}\}$), and in additional analyses an indicator of an adverse early labor-market outcome in the first employment spell ($Y_i^{j_i^{(2)}} = 1$ if contract is being dissolved within the first year, and 0 otherwise).

$g(X_i)$ controls for potentially confounding characteristics of individual i that may be correlated with both treatment and the outcome(s). Most importantly, we flexibly control for the individual's ability using the test score information. First, it helps net out differences in ability. Higher-ability individuals may be better informed about their own skills. They may also navigate the career-choice process more effectively (e.g., by understanding occupational requirements or considering a broader set of occupations), which could reduce mismatches independently of imperfect information about their own skills. Second, controlling for the score addresses mechanical features of our variable construction that would otherwise bias our estimates. Because scores are bounded between 0 and 100, over- or underestimation is constrained near the extremes, which can mechanically reduce the accuracy of self-knowledge. Likewise, very high or very low scores may mechanically increase measured mismatch if most occupational requirements cluster away from the extremes.

Besides an ability measure, across all main specifications, $g(X_i)$ includes an indicator for gender, as men and women differ in their perceptions of their skills (Bordalo, Coffman, Gennaioli, & Shleifer, 2019; Brox, Goller, & Wolter, 2026) and may systematically choose different occupations. Similarly, we include age because older individuals may know more about their own skills and labor-market prospects. Some test takers have their test paid for by someone else – using a company voucher – which might imply different incentives for taking the test, thereby influencing the treatment and outcome variables. Being raised in different labor market contexts may influence when you become interested in different occupations and how well you can assess your skills. These factors are controlled for using four variables: being Swiss or having been born in Switzerland to capture the general context, and whether one's language is the locally spoken language or one of the four national languages to capture the cultural differences induced by language barriers.

In our baseline specifications, we include the test score in linear and quadratic form. In robustness checks, we alternatively control for fine score-bin indicators (0-1, 1-2, ..., 99-100) or for subtest-area scores (school knowledge, potential, and professional knowledge). Individuals might select into potentially slightly different test versions across cohorts (a test cohort always runs from May to April), and test composition also varies by occupation (see Section 3.2). While we account for these features when constructing the AKM-based Variant B of our imperfect self-knowledge measure, we include cohort and occupation fixed effects in most specifications. In robustness checks, we test the sensitivity of the estimates with regard to the inclusion of additional fixed effects for the canton (Swiss regions) in which the test taker lives, because the school curricula are partly determined in cantons, respectively, school fixed effects, both potentially carrying unobserved heterogeneity in the content and scope of the career information lectures and the intensity of feedback about one's own skills.

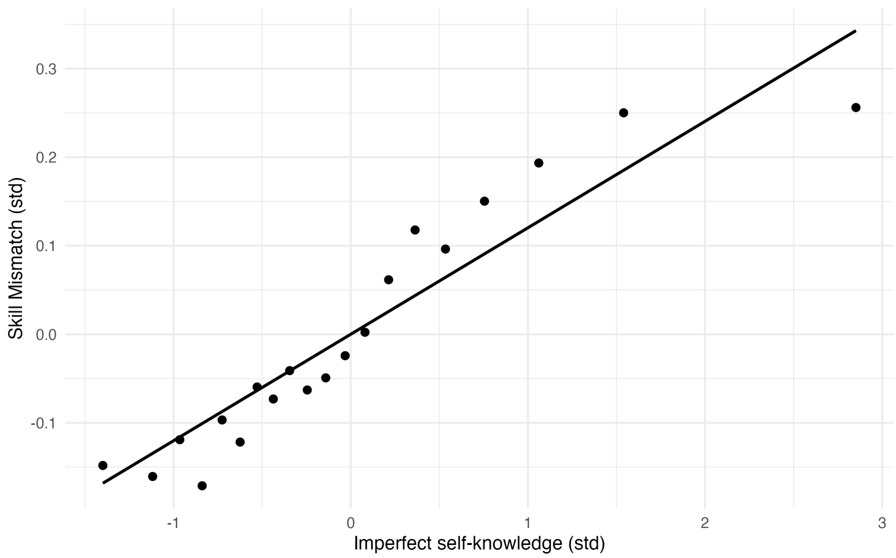
4 Results

In the following, we present empirical evidence for the hypotheses discussed in Section 2.2 on the association between imperfect self-knowledge about one's own skills and skill mismatch. We organize the results around three questions: how imperfect self-knowledge affects initial sorting, how it translates into early labor market outcomes, and how individuals adjust when information is revealed. We use the AKM-type imperfect self-knowledge variant B for the analyses in the main section.

4.1 Imperfect Self-Knowledge about Skills and Preferences for Occupations

We start by investigating whether information frictions distort occupational preferences. Occupational preferences are not directly observed but are revealed through the choice of an occupation-specific aptitude test. Students actively select the occupation for which they wish to be tested, making this choice a salient and consequential indicator of occupational preferences. We measure the mismatch between the objective skills and the requirements of the test occupation, as explained in Section 3.5, and link it to our individual measure of imperfect self-knowledge of one's own skills.

Figure 2: Imperfect self-knowledge and skill mismatch



Notes: Binscatter plot with 20 equally sized bins showing the raw relationship between the standardized (std) skill mismatch, on the y-axis, and the imperfect self-knowledge Variant B (akm-type) on the x-axis. N = 25,412.

Figure 2 shows the unconditional relationship between imperfect self-knowledge and the skill mismatch in the test occupation. The pattern is monotonic and economically large: students with higher imperfect self-knowledge select test occupations that are substantially less aligned with their objective skills profiles. This relationship is present across the distribution and is not driven by a few extreme observations. The figure provides direct visual evidence that imperfect self-knowledge of one's own skills translates into systematically distorted occupational aspirations.

Table 1 quantifies this relationship in a regression framework. From column (1) to (5), we add control variables to show how the coefficient of imperfect self-knowledge on skill mismatch evolves when controlling for other important components in the job search process. Across all models, higher imperfect self-knowledge is associated with significantly worse alignment between individual skills and occupational requirements. The estimated magnitudes decrease when controlling for student ability but remain sizable and stable when controlling for demographics, cohort effects, and occupation fixed-effects. In column (5), our preferred specification, an increase of one standard deviation in imperfect self-knowledge about one's own skills increases the skill mismatch by 0.050 standard deviations.

Table 1: Imperfect self-knowledge and skill mismatch.

	(1)	(2)	(3)	(4)	(5)
Imperfect self-knowledge (std)	0.120***	0.119***	0.073***	0.046***	0.050***
	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)
Male			0.044***	0.050***	0.063***
			(0.012)	(0.009)	(0.010)
Test score			-0.034***	-0.495***	-0.498***
			(0.001)	(0.005)	(0.005)
Test score (sq)				0.004***	0.004***
				(0.000)	(0.000)
All controls				✓	✓
Occupation FE		✓			✓
Cohort FE		✓			✓
Observations	25,412	25,412	25,412	25,412	25,412

Notes: Outcome: skill mismatch. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Taken together, the evidence establishes that imperfect self-knowledge about one's own skills is systematically associated with distorted occupational preferences. Students with noisier or lower self-knowledge about their own skills select occupations that are less well matched to their objective skill profiles, even before institutional constraints, employer screening, or learning dynamics come into play. This preference distortion at the point of aspiration is the first step in the misallocation process we study. On note of these results, we can conclude that Hypothesis 1, *Individuals with higher information frictions are more likely to choose an initially mismatched occupation*, cannot be rejected.

Having established that imperfect self-knowledge about skills is systematically associated with distorted occupational preferences, a natural next question is about its relevance. In a hypothetical world, how much would reducing or removing these frictions improve the alignment between individual skills and the requirements of preferred occupations? To gauge the quantitative importance of this channel, we conduct a simple counterfactual exercise, assuming improvements in self-knowledge, and see how they map onto changes in average skill mismatch.

Table 2: Counterfactual average mismatch under selected self-knowledge scenarios

	Obs. mean	Counterf. mean	Abs. red.	Relative red. (%)	Irreduc. mismatch	Red. in reducible
Perfect self-knowledge (Imp. self-knowl.=0)	12.143	10.922	1.221	10.1	8.412	32.7
Reduce imperfect self-knowledge by 1 SD (all)	12.143	11.681	0.462	3.8	8.412	12.4
Replace self-knowledge in top 50% by mean(bottom 50%)	12.143	11.807	0.337	2.8	8.412	9.0

Notes: This table uses the (raw) imperfect self-knowledge variant A and the skill mismatch, both in a non-standardized version. N = 25,412. Counterfactual simulations based on the scenarios described in the text. The reduction in reducible mismatch (*Red. in reducible*) is calculated based on the average reducible component of the skill mismatch, i.e., Observed mean (*Obs. mean)* - irreducible mismatch (*Irreduc. mismatch*). The counterfactual mean (*Counterf. mean*) is the predicted mismatch with imperfect self-knowledge reduced according to the scenario and multiplied by $\hat{\beta} = 0.077$, estimated from the observed relationship: ($m_i^{(1)} = \alpha + \beta \cdot ImperfectSelfKnowledge_i + u_i$). From this, the absolute reduction (*Abs. red*) and relative reduction in percent (*Relativ red. (%)*) are calculated.

The counterfactual simulations reported in Table 2 are based on a linear projection of (non-standardized) skill mismatch in the initial occupation on the Variant A measure of imperfect self-knowledge. We use this specification because Variant A has a natural lower bound at zero (i.e., perfect self-knowledge), allowing us to consider well-defined reductions in imperfect self-knowledge. Results are similar when using the AKM-based measure of imperfect self-knowledge. The observed average mismatch in the regression sample is $\overline{m}_{\text{obs}}^{(1)} = 12.143$. To distinguish between reducible and irreducible skill mismatch, we compute, for each individual, the minimum hypothetical skill mismatch across all observed occupations, i.e., the individual's best fit among all occupations. In this sample, the irreducible mismatch is $\overline{m}_{\text{irr}}^{(1)} = 8.412$, implying an average reducible component of 3.731.

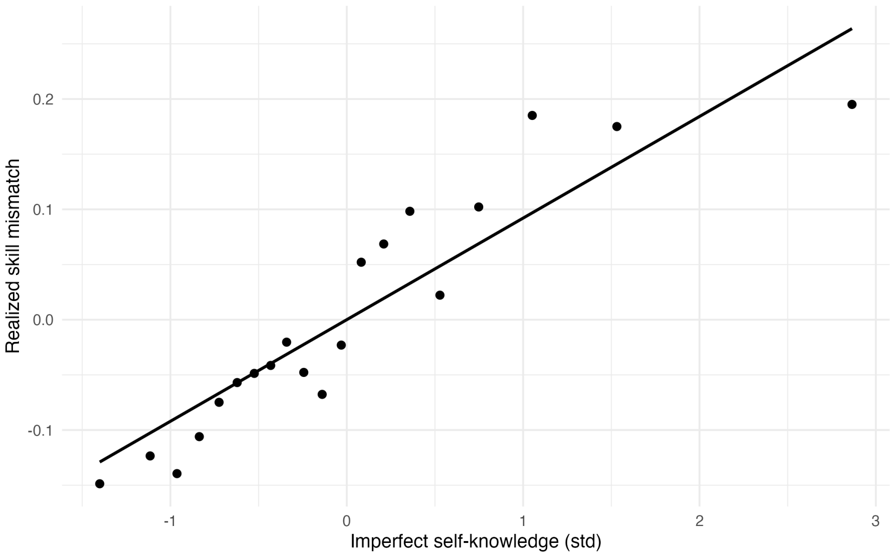
Scenario (1) of Table 2 considers a stylized lower-bound scenario in which imperfect self-knowledge is set to zero for all individuals. Under this extreme assumption, average mismatch falls by 1.221 points, corresponding to a reduction of about 10 percent relative to the observed mean. Importantly, this translates into a nearly one-third reduction in the reducible mismatch component, illustrating the potential importance of self-knowledge for occupational sorting.

Scenarios (2) and (3) report more conservative counterfactuals. A uniform one standard deviation reduction in imperfect self-knowledge lowers average mismatch by 0.462 points, eliminating about 12 percent of the reducible component. Replacing the imperfect self-knowledge of the top half of the distribution with the mean imperfect self-knowledge of the bottom half yields a slightly smaller but still non-negligible reduction of 0.337 points, corresponding to roughly 9 percent of the reducible mismatch.

While these simulations are purely mechanical and do not account for general equilibrium effects or behavioral responses, they help contextualize the magnitude of the estimates in Table 1. Even modest improvements in self-knowledge about one's own skills would translate into economically meaningful improvements in the alignment between individual skills and the requirements of the aspired occupations.

4.2 Imperfect Self-Knowledge about Skills and Realized Skill Mismatch

Figure 3: Imperfect self-knowledge and realized skill mismatch



Notes: Binscatter plot with 20 equally sized bins showing the raw relationship between the standardized (std) realized skill mismatch, on the y-axis, and the imperfect self-knowledge Variant B (akm-type) on the x-axis. The sample is restricted to those who started an apprenticeship in either occupation for which we have information on occupational skill requirements; N=16,083.

We now turn to the relationship between imperfect self-knowledge and realized mismatch in the occupation ultimately entered. While Section 4.1 shows that imperfect self-knowledge about one's own skills distorts occupational preferences at the point of aspiration, realized outcomes reflect both initial sorting and subsequent adjustment. Changing from the initially preferred occupation can be costly: individuals may have already mentally committed to their choice, informed peers and family, or invested time and effort in occupation-specific preparation. As a result, initial misallocation need not be fully undone, even when better information becomes available. Hypothesis 4, therefore, states that individuals with higher imperfect self-knowledge about skills remain more mismatched on average, even after seeing test results and receiving feedback on their skills and the possibility of changing their occupation.

Figure 3 presents the unconditional relationship between our measure of imperfect self-knowledge about skills and realized skill mismatch in the finally chosen occupation. The pattern closely mirrors that observed for mismatch in occupational preferences in Figure 2. Realized mismatch increases monotonically with imperfect self-knowledge about one's own skills, and the relationship is economically meaningful across the distribution.

Table 3: Imperfect self-knowledge and realized skill mismatch.

	(1)	(2)	(3)	(4)	(5)
Imperfect self-knowledge (std)	0.092***	0.084***	0.065***	0.048***	0.047***
	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)
Male			0.018	0.028**	0.066***
			(0.016)	(0.013)	(0.015)
Test score			-0.024***	-0.505***	-0.503***
			(0.001)	(0.007)	(0.007)
Test score (sq)				0.004***	0.004***
				(0.000)	(0.000)
All controls				✓	✓
Occupation FE		✓			✓
Cohort FE		✓			✓
Observations	16,083	16,083	16,083	16,083	16,083

Notes: Outcome is the realized skill mismatch. The sample is restricted to those who started an apprenticeship in either occupation for which we have information on the occupational skill requirements ('Worker'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

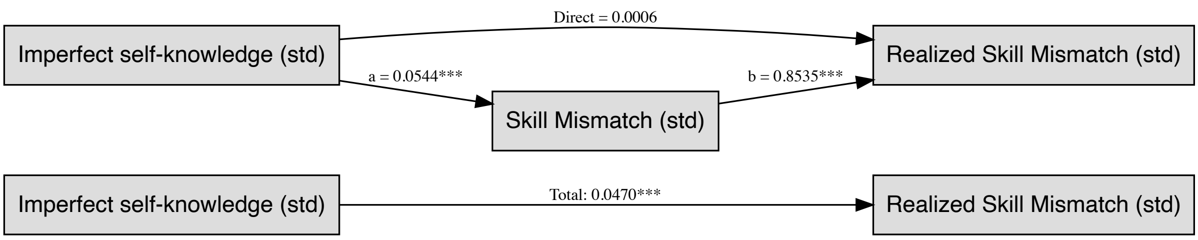
Table 3 quantifies this relationship. Across all specifications, higher information frictions are associated with significantly worse alignment between individual skills and the requirements of the realized occupation. The estimated coefficients are remarkably stable after controlling for ability. In our preferred specification in column (5), a one standard deviation increase in imperfect self-knowledge raises realized skill mismatch by 0.047 standard deviations, an effect size very similar to that obtained for initial mismatch in Table 1. This similarity in magnitude indicates that a substantial share of initial mis-sorting persists into realized outcomes.

The robustness of the coefficient across specifications suggests that observable background characteristics, differences in overall ability, or institutional sorting across occupations and cohorts do not drive the association. Instead, the results are consistent with the model's prediction that individuals with stronger information frictions start from a worse allocation and remain more mismatched on average even after receiving better information.

The results above establish a robust positive relationship between imperfect knowledge about one's own skills and realized skill mismatch. To understand through which channels this relationship arises, it is important to distinguish between the direct effects of imperfect self-knowledge on final outcomes and the indirect effects operating through earlier stages of the career-choice process. In particular, imperfect self-knowledge may impair realized occupational fit primarily by distorting initial occupational preferences, which persist despite subsequent adjustment.

To disentangle these mechanisms, we conduct a mediation analysis in which skill mismatch in initial occupational preferences serves as the mediator between imperfect self-knowledge and realized skill mismatch (Imai, Keele, & Tingley, 2010). This framework allows us to decompose the total effect of imperfect self-knowledge into an indirect component operating through initial skill mismatch and a direct component capturing any remaining association with realized skill mismatch.

Figure 4: Causal mediation analysis of imperfect self-knowledge on realized skill mismatch, mediator: initial skill mismatch.



Notes: The sample is restricted to those who started an apprenticeship in either occupation for which we have information on the occupational skill requirements. N = 16,083. We controlled for the variables: test score, test score squared, male, occupation, realized occupation, and cohort fixed effects, an indicator whether the individual changed the job, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators.

Table 4: Causal mediation analysis of imperfect self-knowledge on realized skill mismatch, mediator: initial skill mismatch.

Effect	Mean	90% CI Lower	90% CI Upper
Mediation effect	0.0464	0.0378	0.0549
Direct effect	0.0006	-0.0058	0.0068
Total effect	0.0470	0.0367	0.0573
% mediated	98.731	87.018	113.628

Notes: Outcome is the realized skill mismatch. The sample is restricted to those who started an apprenticeship in either occupation for which we have information on the occupational skill requirements ('Worker'). N = 16,083. 90% confidence intervals calculated in 999 bootstrap replications. We controlled for the variables: test score, test score squared, male, occupation, realized occupation, and cohort fixed effects, an indicator whether the individual changed the job, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators.

Figure 4 and Table 4 summarize the results. The total effect of self-knowledge about one's own skills on realized skill mismatch mirrors the estimates reported in Table 3. Almost the entire effect operates through the indirect channel: Lower self-knowledge leads to worse alignment with the requirements of the initially preferred occupation, which, in turn, strongly predicts skill mismatch in the realized occupation. Once the indirect channel through initial skill mismatch is accounted for, the direct effect of information frictions on realized mismatch is close to zero (0.0006) and statistically insignificant (90% confidence interval: [-0.0058; 0.0068], see Table 4).

These findings highlight the central role of initial occupational preferences in shaping final allocation outcomes. Imperfect self-knowledge does not appear to directly impair realized match quality beyond its impact on initial sorting. Instead, distorted self-assessments translate into a persistent mismatch because individuals start from a poorly aligned occupational choice.

4.3 Feedback and Adjustment

To examine whether individuals adjust initial career preferences after receiving information about their own skills through the aptitude test, we focus on two separate questions. First, whether individuals with lower self-knowledge about their own skills are more likely to revise their initial decision, either by choosing a different occupation or by switching to a different educational track. Second, whether this updating improves alignment between their skill set and occupational requirements.

Rational individuals update their choices after receiving feedback on their skills.[10] However, there are also costs involved with changing away from an initial choice. Table 5 shows evidence for changing behavior related to the extend of self-knowledge about their own skills. In column (1), a one standard deviation increase in imperfect self-knowledge is associated with a 1.5 percentage point (3.4%) increase in the likelihood of changing away from the original choice. For a graphical illustration of the raw relationship, see Figure A.1.

Table 5: Imperfect self-knowledge about own skills and changing the initial decision

Change to...	(1) anything	(2) job	(3) high school	(4) other school	(5) other
Imp. self-knowl. (std)	0.015***	0.008**	0.001	0.001	0.004***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Male	0.065***	0.073***	-0.003	-0.004	-0.001
	(0.008)	(0.008)	(0.002)	(0.003)	(0.003)
Test score	-0.021***	-0.019***	-0.003***	0.002**	-0.001
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Test score (sq)	0.000***	0.000***	0.000***	-0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	0.435	0.357	0.020	0.031	0.027
Effect as % of mean	3.4	2.3	5.8	4.5	13.5
Sample	2019-2021	2019-2021	2019-2021	2019-2021	2019-2021
Observations	17,159	17,159	17,159	17,159	17,159

Notes: The sample is restricted to 2019-2021 10th-grader cohorts who successfully finished 11th grade and took their test before the end of 2023. This restriction is because, for the 2022 cohort and some repeaters, we do not observe all the options for change. Outcomes: *anything* is 0 for those starting in the tested occupation, and 1 otherwise; *job* refers to another occupation than the test occupation; *high school* is the academic high track, *other school* formal education paths that are more specialized or technical, *other* includes transitional education programs from secondary level I to secondary level II and motivational semesters. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Moreover, we can distinguish four alternative educational or career options. Columns (3) and (4) in Table 5 show that changers are not choosing an educational or academic path, with insignificant coefficients close to zero. Positive statistically significant coefficients are found for

[10]Feedback can influence a person's behavior in various ways. Beyond realizing that they may not be a good fit, individuals may also respond to feedback by increasing their efforts (Brade, Himmler, & Jäckle, 2026; Goller & Späth, 2024; Goulas & Megalokonomou, 2021), thereby improving their suitability for the occupation.

changing to another occupation, in column (2), and the residual category (such as a transitional education program) in column (5). Hypothesis 2, therefore, cannot be rejected: The evidence provided is in line with the idea that individuals with greater imperfect self-knowledge about their own skills are more likely to switch away from their initial occupation.

So far, we have seen that greater imperfect self-knowledge of one's abilities leads to a poorer initial fit and increases the likelihood of revising the initial choice. Therefore, those who do not choose their first preference may have greater potential to improve their fit. In Table 6, we investigate the relationship between imperfect self-knowledge and the difference in skill mismatch between the initial choice and the realized occupation. First, note that the mean of this outcome variable is negative (-0.095), so we do not observe a general increase in fit for those who re-decide. One reason for the observed pattern may be that not everyone gets a contract in the occupation they want, because vacancies are limited and employment is contingent on both the employee and the employer agreeing to a contract.

Table 6: Imperfect self-knowledge and difference in mismatch.

	(1)	(2)	(3)	(4)	(5)
Imperfect self-knowledge (std)	0.225***	0.241***	0.130*	0.130*	0.150**
	(0.075)	(0.074)	(0.070)	(0.070)	(0.069)
Male			-0.350**	-0.301*	-0.105
			(0.178)	(0.179)	(0.185)
Test score			-0.302***	-0.318***	-0.333***
			(0.070)	(0.071)	(0.068)
Test score (sq)			0.002**	0.002***	0.002***
			(0.001)	(0.001)	(0.001)
All controls				✓	✓
Occupation FE		✓			✓
Realized occupation FE			✓	✓	✓
Cohort FE		✓			✓
Mean of outcome	-0.095	-0.095	-0.095	-0.095	-0.095
Observations	2,961	2,961	2,961	2,961	2,961

Notes: Outcome: skill mismatch - realized skill mismatch. The sample is restricted to those who started an apprenticeship in an occupation other than the initial occupation ('Changer'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

We find evidence in line with Hypothesis 3 that *among individuals who switch occupations, those with higher initial information frictions experience larger expected improvements in match* *quality.* With an estimate of 0.150 in our preferred specification in Table 6, column (5), there is a solid and statistically significant relationship. For a graphical representation of the raw relationship, see Appendix Figure A.2.

An open question regarding the results in this section is whether the observed adjustments reflect individuals updating their information and revising occupational choices, or whether market mechanisms drive them. In particular, students with high levels of imperfect self-knowledge may not voluntarily adjust their choices; instead, they may be forced to do so when they fail to obtain an apprenticeship in their preferred occupation. Ideally, one would observe application behavior directly. As this information is not available in our data, we exploit an additional feature of the aptitude test to study individual adjustment behavior.

Students can do more than one aptitude test.[11] Retaking the test requires an active decision by the student and therefore provides a setting in which adjustments are likely to reflect individual choices rather than general market effects. We use this feature to examine whether students with higher levels of imperfect self-knowledge are more likely to revise their occupational choices.

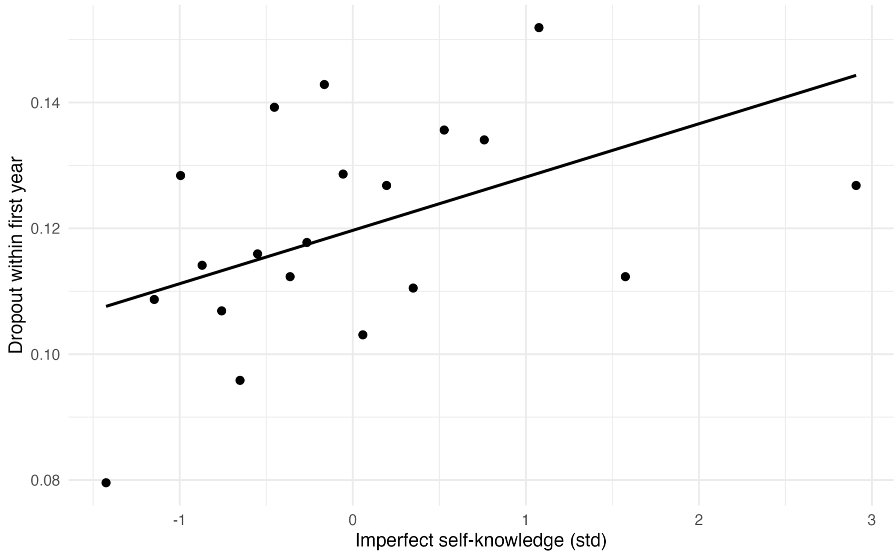
Column (1) of Appendix Table A.7 shows that students with higher imperfect self-knowledge are significantly more likely to register for another test (holding constant the performance in the initial test). Column (2) shows that these students are also more likely to take the second test in a different occupation than in the first test. These findings are consistent with the earlier evidence, showing that students with higher imperfect-self knowledge are less likely to begin an apprenticeship in the test occupation and more likely to begin in another occupation. Moreover, we see that re-taking (a test in a different occupation) is associated with improvements in performance and occupational fit. Students who retake the test in a different occupation perform better on average (Column (3)), and the occupation chosen in the second test exhibits a better match to the individual's skill profile (Column (4)). Two mechanisms drive the reduction in the mismatch between the first and second tests. First, students do better on the test in the newly chosen occupation. Second, students choose occupations that are more closely aligned with their skill profile. Together, these two channels explain why skill mismatch in the second test is lower than in the first test (Column (5)). Finally, we do not find evidence that self-assessments become more accurate between the first and the second test (Column (6)).

[11]In general, students can register at any time for the same test type or a different one than the one they have already taken, provided they pay the test fee again.

4.4 Early Labor Market Outcomes

So far, our results show that imperfect self-knowledge of one's own skills is associated with a higher realized skill mismatch, which, in turn, is associated with poor labor market outcomes (Brunello & Wruuck, 2021; Fredriksson et al., 2018; Guvenen et al., 2020). The question that naturally follows is whether imperfect self-knowledge itself affects labor-market success, or only indirectly through occupational choice. In this section, we use an early indicator of labor market performance, namely dropout within the first year, to study this question.

Figure 5: Imperfect self-knowledge and dropout from apprenticeship within the first year



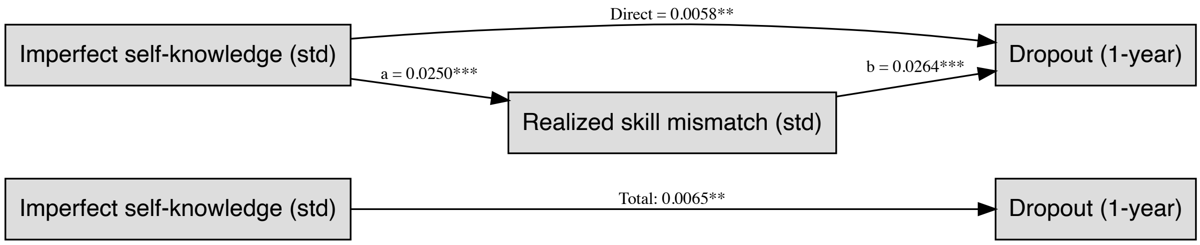
Notes: Binscatter plot with 20 equally sized bins showing the raw relationship between dropout within the first year, on the y-axis, and the standardized (std) imperfect self-knowledge Variant B (akm-type) on the x-axis. The sample is restricted to those who started an apprenticeship and who we can observe for at least 1 year. N = 11,048.

Figure 5 shows the raw relationship between imperfect self-knowledge about one's own skills and a binary indicator showing whether an individual's working contract is dissolved within the first year. The graph shows a positive association between imperfect self-knowledge and dropout. This relationship holds when controlling for various potential confounders (see Appendix Table A.8).

While this pattern is informative, it does not distinguish between direct effects of imperfect self-knowledge and indirect effects operating through occupational choice. To separate these channels, we conduct a second mediation analysis (Imai et al., 2010) to decompose the total effect into a direct effect and an indirect effect operating through skill mismatch.

Figure 6 and Appendix Table A.9 show the results of the mediation analysis. The total effect of imperfect self-knowledge on dropout is 0.65 percentage points (5% relative to the sample mean). Figure 6 shows the relationship between imperfect self-knowledge and the mediator (realized skill mismatch), with a statistically significant coefficient of 0.0250. Similarly, the relationship between the mediator and the outcome (dropout) is of similar magnitude (0.0264). In sum, this mediation effect is a robust effect that accounts for approximately 10 percent of the total effect (see Appendix Table A.9).

Figure 6: Causal mediation analysis of imperfect self-knowledge on dropout via realized skill mismatch.



Notes: The sample is restricted to those who started an apprenticeship and who we can observe for at least 1 year. N = 11,048. We controlled for the variables: test score, test score squared, male, occupation, realized occupation, and cohort fixed effects, an indicator whether the individual changed the job, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. The actual confidence intervals for the point estimates can be found in Appendix Table A.9.

At the same time, once the indirect channel through skill mismatch is accounted for, the direct effect of imperfect self-knowledge on dropout is positive and statistically significant. This direct channel might include residual associations of imperfect self-knowledge about one's own skills through other channels, such as meta-cognitive ability, decision-making quality, or adjustment costs, which matter for early labor market success beyond acting as information frictions in the occupational choice process (Bénabou & Tirole, 2002; Falk et al., 2023).

4.5 Robustness

We conduct a series of robustness checks. First, we assess the robustness of our results to alternative ways of controlling for skills. In Table A.10 - Table A.14, we show that controlling for ability in a non-parametric way using score fixed effects does not change the result. Second, we assess the robustness of our results towards the alternative measure of imperfect self-knowledge described in Section 3.6. The results summarized in Table A.15 - Table A.19 show that our results are not affected by the alternative measure of imperfect self-knowledge.

Third, we examine the sensitivity of the estimates to the tested cohorts. Although the baseline specifications already include cohort fixed effects, defined by the testing season the test-taker took the aptitude test, the results could still be influenced by cohort-specific labor market conditions or institutional features at the time of VET search. To assess this possibility, Figure A.3 reports the estimated relationship separately by cohort. The results are highly similar across cohorts, indicating that the main findings are not driven by any particular year.

Fourth, we examine whether the results are robust to regional heterogeneity. In Table A.20, we add canton fixed effects to the baseline specification to account for differences in cantonal education systems, regional labor market conditions, and institutional environments that may affect both self-knowledge and occupational sorting. The estimated coefficients remain very close to the baseline results, suggesting that the main findings are not explained by regional variation.

Fifth, we include school fixed effects to control for unobserved differences in school-level inputs. Schools may differ in the extent of career guidance, job-search training, performance feedback, and other resources that help students learn about their own skills and make occupational choices. By including school fixed effects, the estimates are identified from comparisons within the same school, holding constant any unobserved school-level characteristics. The results remain stable, indicating that differences across schools are not driving the main results (Table A.21).

Finally, we assess the robustness of the results with respect to the construction of the imperfect self-knowledge measure. Objective test scores are recorded on a continuous scale from 0 to 100, whereas subjective beliefs are reported on a discrete 10-point grid. The difference in measurement scales may mechanically induce deviations between scores and beliefs at the module level, even in the absence of true misperception, and may therefore affect the measured level of imperfect self-knowledge. To address this concern, we construct an alternative measure of skills by discretizing test scores to the same 10-point grid used for belief elicitation. Using this discretized skill measure, we recompute the imperfect self-knowledge index and re-estimate the main specifications. The results remain quantitatively and statistically very similar to the baseline estimates, indicating that the findings are not driven by the difference in measurement scales (Table A.22).

5 Conclusion

This study investigates a simple yet largely neglected mechanism that explains early occupational misallocation: imperfect self-knowledge. Using uniquely rich data from standardized occupation-specific aptitude tests linked to administrative education and early labor-market records in Switzerland, we can observe and measure objective performance and subjective self-assessments, and construct individual-level measures of self-knowledge about one's own skills.

Our findings yield three main insights. First, imperfect self-knowledge systematically distorts occupational aspirations at the outset of the career choice process. Individuals with poorer self-knowledge select occupations that are less aligned with their objective skill profiles. Counterfactual simulations indicate that improving self-knowledge would lead to economically meaningful gains in allocative efficiency.

Second, these initial distortions persist into realized outcomes. Even after opportunities to revise initial choices, individuals with greater misperception remain more mismatched in the occupations they ultimately enter. Mediation analysis shows that this persistence operates almost entirely through the channel of initial distorted occupational preferences, underscoring the central role of self-beliefs in shaping career trajectories at a very early stage.

Third, inaccurate self-assessment has tangible consequences for early labor-market outcomes. Individuals with poorer self-knowledge are significantly more likely to experience early dissolution of apprenticeship contracts. While a substantial share of this relationship is mediated by realized mismatch, a direct effect remains, consistent with self-assessment errors influencing human capital accumulation beyond their role as an informational friction. At the same time, individuals exhibit active adjustment behavior: those with poorer self-knowledge are more likely to revise their initial choices – primarily by switching occupations – and conditional on switching, these adjustments tend to improve occupational fit.

Our results suggest the broader insight that a non-trivial share of occupational misallocation originates before labor-market entry, driven by systematic errors in self-assessment rather than solely search costs or institutional constraints. This perspective shifts attention toward the formation of self-beliefs during adolescence. It highlights the potential role of policies and institutions that aim to improve self-knowledge through structured feedback, repeated evaluation, and targeted guidance in enhancing allocative efficiency. An important avenue for future research is to identify which types of feedback and counseling interventions are most effective in improving self-assessment accuracy, and whether such improvements translate into persistent reductions in mismatch and long-run gains in educational and labor-market outcomes.

References

Abel, M., Burger, R., & Piraino, P. (2020). The value of reference letters: Experimental evidence from south africa. *American Economic Journal: Applied Economics*, *12*(3), 40–71.

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–334.

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4B, pp. 1043–1171). Elsevier.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, *121*(1–2), 343–375.

Arcidiacono, P., Aucejo, E., Maurel, A., & Ransom, T. (2025). College attrition and the dynamics of information revelation. *Journal of Political Economy*, *133*(1), 53–110.

Arcidiacono, P., Hotz, V. J., & Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, *166*(1), 3-16.

Arcidiacono, P., Hotz, V. J., Maurel, A., & Romano, T. (2020). Ex ante returns and occupational choice. *Journal of Political Economy*, *128*(12), 4475-4522.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333.

Ayllón, S., Lefgren, L., Patterson, R., Stoddard, O. B., & Urdaneta, N. (2025). *'Sorting' Out Gender Discrimination and Disadvantage: Evidence from Student Evaluations of Teaching* (IZA Discussion Paper No. 18040). Institute of Labor Economics (IZA).

Baley, I., Figueiredo, A., & Ulbricht, R. (2022). Mismatch cycles. *Journal of Political Economy*, *130*(11), 2943–2984.

Belot, M., Kircher, P., & Muller, P. (2018). Providing advice to jobseekers at low cost: An experimental study on online advice. *The Review of Economic Studies*, *86*(4), 1411-1447.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, *117*(3), 871–915.

Benoît, J.-P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica*, *79*(5), 1591–1625.

Bobba, M., & Frisancho, V. (2022). Self-perceptions about academic achievement: Evidence from mexico city. *Journal of Econometrics*, *231*(1), 58-73.

Bodner, R., & Prelec, D. (2003). Self-signaling and diagnostic utility in everyday decision making. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions: Rationality and well-being* (pp. 105–123). Oxford: Oxford University Press.

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, *109*(3), 739–773.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review*, *77*(4), 531–553.

Brade, R., Himmler, O., & Jäckle, R. (2026). No student left behind? Relative feedback and university completion. *Journal of Economic Behavior & Organization*, *241*, 107383.

Brox, E., Davoli, M., & Strazzeri, M. (2025). *Classroom rank in math, occupational choices and labor market outcomes* (Tech. Rep.). LH Working Paper Seeries.

Brox, E., Goller, D., & Wolter, S. C. (2026). Gender differences in beliefs about productivity relevant skills. *Mimeo*.

Bruhn, J., Fabian, J., Gallagher, L., Gudgeon, M., Isen, A., & Phipps, A. R. (2025). *Path dependence in the labor market: The long-run effects of early career occupational experience* (Tech. Rep.). National Bureau of Economic Research.

Brunello, G., & Wruuck, P. (2021). Skill shortages and skill mismatch: A review of the literature. *Journal of Economic Surveys*, *35*(4), 1145–1167.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*(1), 306–318.

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics*, *128*(3), 967–1015.

Carranza, E., Garlick, R., Orkin, K., & Rankin, N. (2022). Job search and hiring with limited information about workseekers' skills. *American Economic Review*, *112*(11), 3547–3583.

Conlon, J. J., Pilossoph, L., Wiswall, M., & Zafar, B. (2018). Labor market search with imperfect information and learning. *Journal of Political Economy*, *126*(1), 110–156.

Eckardt, D. (2026). Training specificity and occupational mobility: evidence from german apprenticeships. *Econometrica*, *forthcoming*.

Falk, A., Kosse, F., Schildberg-Hörisch, H., & Zimmermann, F. (2023). Self-assessment: The role of the social environment. *Journal of Public Economics*, *223*, 104908.

Fredriksson, P., Hensvik, L., & Skans, O. N. (2018). Mismatch of talent: Evidence on match quality, entry wages, and job mobility. *American Economic Review*, *108*(11), 3303–3338.

Gautier, P. A., & Teulings, C. N. (2015). Sorting and the output loss due to search frictions. *Journal of the European Economic Association*, *13*(6), 1136-1166.

Gibbons, R., Katz, L. F., Lemieux, T., & Parent, D. (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics*, *23*(4), 681–724. doi: 10.1086/491606

Goller, D., & Späth, M. (2024). 'Good job!' The impact of positive and negative feedback on performance. *Sports Economics Review*, *8*, 100045.

Goulas, S., & Megalokonomou, R. (2021). Knowing who you actually are: The effect of feedback on short-and longer-term outcomes. *Journal of Economic Behavior & Organization*, *183*, 589–615.

Groh, M., Krishnan, N., McKenzie, D., & Vishwanath, T. (2016). Do wage subsidies provide a stepping-stone to employment for recent college graduates? Evidence from a randomized experiment in Jordan. *The Review of Economics and Statistics*, *98*(3), 488–502.

Guvenen, F., Kuruscu, B., Tanaka, S., & Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, *12*(1), 210–44.

Hakimov, R., Schmacker, R., & Terrier, C. (2023). *Confidence and college applications: Evidence from a randomized intervention* (Tech. Rep.). Working Paper.

Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017). General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of Human Resources*, *52*(1), 48–87.

Heckman, J. J., & Sedlacek, G. (1985). Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy*, *93*(6), 1077–1125.

Heller, S. B., & Kessler, J. B. (2024). Information frictions and skill signaling in the youth labor market. *American Economic Journal: Economic Policy*, *16*(4), 1–33.

Hensvik, L., Müller, D., & Skans, O. N. (2023). Connecting the young: High school graduates' matching to first jobs in booms and great recessions. *The Economic Journal*, *133*(652), 1466–1509.

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, *87*(5), 1439-1474.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309-334.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, *87*(5), 972–990.

Kiss, A., Garlick, R., Orkin, K., & Hensel, L. (2023). Jobseekers' beliefs about comparative advantage and (mis) directed search. *IZA Discussion Paper*.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 4, pp. 283–326). Elsevier.

Lise, J., & Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, *110*(8), 2328–2376.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, *20*(3), 387–418.

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political Economy*, *92*(6), 1086–1120.

Moscarini, G., & Postel-Vinay, F. (2012). The contribution of large and small employers to job creation in times of high and low unemployment. *American Economic Review*, *102*(6), 2509–2539.

Neal, D. (1999). The complexity of job mobility among young men. *Journal of Labor Economics*, *17*(2), 237–261.

OECD. (2025). Vocational education and training in Switzerland. In *Vocational Education and Training Systems in Nine Countries* (pp. 190–205). Paris: OECD Publishing. (Chapter 10)

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, *104*(11), 3565–3599.

Papageorgiou, T. (2014). Learning your comparative advantages. *The Review of Economic Studies*, *81*(3), 1263–1295.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Santos-Pinto, L., & de la Rosa, L. E. (2020). Overconfidence in labor markets. In K. F. Zimmermann (Ed.), *Handbook of Labor, Human Resources and Population Economics* (pp. 1–42). Cham: Springer.

Santos-Pinto, L., & Sobel, J. (2005). A model of positive self-image in subjective assessments. *American Economic Review*, *95*(5), 1386–1402.

Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, *13*(1), 130–167.

Stinebrickner, R., & Stinebrickner, T. (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, *32*(3), 601–644.

Topel, R. H., & Ward, M. P. (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, *107*(2), 439–479.

Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, *82*(2), 791-824.

Wiswall, M., & Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy*, *129*(5), 1361-1424.

Yamaguchi, S. (2010). The effect of match quality and specific experience on career decisions and wage growth. *Labour Economics*, *17*(2), 407–423.

Şahin, A., Song, J., Topa, G., & Violante, G. L. (2014). Mismatch unemployment. *American Economic Review*, *104*(11), 3529–64.

Appendix A: Descriptive Statistics and Additional Results

A.1 The Professional Aptitude Test

Table A.2: Submodules contained in each module.

Module	Submodules
Panel A: School knowledge	
German (language)	Match words (assign words), Communicate, Reading comprehension, Word order / arrange words, Incorrect letter, Formulating, Grammar, Key statements, Keywords, Synonyms, Spelling, Sentence structure, Vocabulary
French (language)	Match words (assign words), Communicate, Reading comprehension, Word order / arrange words, Incorrect letter, Formulating, Grammar, Communicating, Key statements, Keywords, Synonyms, Spelling, Syntax, Vocabulary
Italian (language)	Match words (assign words), Communicate, Complete words, Reading comprehension, Word order / Arrange words, Incorrect letter, Grammar, Communicating, Spelling, Vocabulary
English (language)	Match words (assign words), Communicate, Reading comprehension, Word order / Arrange words, Incorrect letter, Grammar, Communicating, Reading comprehension, Vocabulary
Mathematics	Calculations, Approximate calculations (estimation), Mental arithmetic and conversions, Units, Geometry, Reasoning exercises, Word problems, With calculator, Without calculator
Panel B: Cognitive potential	
Concentration	Count images, Direction changes, Compare words, Compare numbers, Compare number sequences, Coordinates, Figure series, Identify symbols
Digital competences	Computer and internet

Table A.2 (continued)

Module	Submodules
Imagination	Spatial visualization / 3D vision, Unfolding (nets), Assemble photos, Completions, Understand 3D representation
Logic	Figural analogies, Verbal analogies, Train lines, Picture stories, Missing figure, Word groups, Letter-sequence logic, Process logic, Syllogisms, Numeric processing capacity, Linear, Vignette
Memory	Recall objects, Recall clothing, Recall people, Recall signs, Recall text, Recall pictograms (learning phase), Recall pictograms (response phase)
Perception	Assemble photos, Recognize portraits, Spatial visualization / 3D vision
Short-term memory	Color combinations, Character/sign combinations, Remember shape sequences, Numeric processing capacity
Panel C: Professional knowledge	
Connected thinking	Company history, Connect information, Connect information 1, Connect information 2
Graphical foundations	Image mood, Color palettes, Color mood
IT basics	Analysis, Nature and technology, Programming
Organizational skills	Scheduling, Restock shelves
Practical knowledge	Practical basic knowledge, Everyday practical knowledge, Sorting methods
Science	Basic natural science knowledge
Technical understanding	Nature and technology, Physics and technology

Notes: Submodules for the Modules (translated to English).

Table A.1: Overview of skills by field-based test-type

Module	Skilled trades	Media and Design	Health and Social Services	Technical	Retail and Service	Business and Administration	ICT
			Panel A: School knowledge				
German (language)	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
English (language)	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
French (language)	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
Italian (language)	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]
Mathematics	✓	✓	✓	✓	✓	✓	✓
			Panel B: Cognitive potential				
Concentration	✓	✓	✓	✓	✓	✓	✓
Digital competences						✓	
Imagination	✓	✓		✓		✓	✓
Logic	✓	✓	✓	✓	✓	✓	✓
Memory	✓	✓	✓	✓	✓	✓	✓
Perception			✓				
Short-term memory		✓	✓	✓		✓	✓
			Panel C: Professional knowledge				
Graphical foundations		✓					
IT basics							✓
Organizational skills			✓	✓	✓	✓	✓
Connected thinking		✓	✓		✓	✓	✓
Practical knowledge	✓						
Sciences			✓				
Technical understanding				✓			

Notes: This table shows the skills that are tested in each of the field-based test-types. Each of those 7 test types contains several occupation-specific tests. [a] The language that is tested can be chosen by the test-taker. It is usually two of the Swiss national languages (German, French, Italian) plus English.

A.2 Descriptive Statistics and Figures

Table A.3: Descriptive comparisons of sample with two populations

Variable	(1) Test-taker	(2) 11th grader	(3) VET starter	(4) Std. Diff (1) vs. (2)	(5) Std. Diff (1) vs. (3)
Male	0.53 (0.50)	0.51 (0.50)	0.57 (0.50)	0.04	-0.08
Age	15.21 (1.00)	14.76 (0.67)	17.91 (3.87)	0.53	-0.95
Born in CH	0.87 (0.33)	0.84 (0.36)	0.80 (0.40)	0.09	0.20
Swiss	0.75 (0.43)	0.74 (0.44)	0.71 (0.46)	0.03	0.10
First = Local lang.	0.58 (0.49)	0.68 (0.47)	0.64 (0.48)	-0.21	-0.12
First = Nat lang.	0.63 (0.48)	0.72 (0.45)	0.69 (0.46)	-0.20	-0.12
Prior School in CH	1.00 (0.00)	1.00 (0.00)	0.88 (0.33)	0.00	0.52
Observations	25,412	325,790	231,294		

Notes: Mean and standard deviations in parentheses. Columns 5-6 report standardized differences (Cohen's d; pooled SD). 'Test-taker' is the sample used in this study, which is restricted to 10th-graders from 2019-2022. The population of 11th graders covers the full cohorts in all Swiss schools from 2020-2023. VET starters are those starting a dual apprenticeship in Switzerland from 2020-2023 (only the first appearance for those starting multiple apprenticeships in this time frame).

Table A.4: Descriptive Statistics

Variable	Lower imperfect self-knowledge	Higher imperfect self-knowledge	Overall	N
Panel A: Outcomes				
Skill mismatch	11.74 (3.59)	12.55 (3.99)	12.14 (3.82)	25,412
Skill mismatch (std)	-0.11 (0.94)	0.11 (1.05)	0.00 (1.00)	25,412
Realized skill mismatch	11.53 (3.47)	12.11 (3.70)	11.82 (3.59)	16,083
Realized skill mismatch (std)	-0.08 (0.96)	0.08 (1.03)	-0.00 (1.00)	16,083
Change anything	0.41 (0.49)	0.46 (0.50)	0.44 (0.50)	17,159
Change job	0.33 (0.47)	0.38 (0.49)	0.36 (0.48)	17,159
Other sek II	0.03 (0.18)	0.03 (0.17)	0.03 (0.17)	17,159
Gymnasium	0.02 (0.15)	0.02 (0.14)	0.02 (0.14)	17,159
Dropout (1-year)	0.11 (0.32)	0.13 (0.33)	0.12 (0.32)	11,048
Panel B: Treatments				
Imperfect self-knowledge (A)	11.63 (2.30)	20.06 (5.59)	15.84 (6.00)	25,412
Imperfect self-knowledge (B)	-0.72 (0.35)	0.72 (0.91)	0.00 (1.00)	25,412
Panel C: Covariates				
Male	0.55 (0.50)	0.51 (0.50)	0.53 (0.50)	25,412
Age	15.20 (1.00)	15.22 (0.99)	15.21 (1.00)	25,412
Test score	54.94 (9.49)	51.83 (9.83)	53.38 (9.79)	25,412
Test score (sq)	3108 (1045)	2783 (1030)	2946 (1050)	25,412
Born Swiss	0.77 (0.42)	0.73 (0.44)	0.75 (0.43)	25,412
Born in Switzerland	0.88 (0.33)	0.87 (0.34)	0.87 (0.33)	25,412
First lang. = national lang.	0.65 (0.48)	0.61 (0.49)	0.63 (0.48)	25,412
First lang. = local lang.	0.60 (0.49)	0.56 (0.50)	0.58 (0.49)	25,412
Company voucher	0.03 (0.18)	0.03 (0.18)	0.03 (0.18)	25,412

Notes: Higher/Lower imperfect self-knowledge divided at the median value of the imperfect self-knowledge (Variant B). The number of observations in column 'N' refers to the 'Overall' column. (std) = standardized, (sq) = squared. Mean values. Standard deviations in parentheses.

Table A.5: Descriptive insights on the module level

Variable	Mean	SD	Min	Max	N
Score	55.406	(18.091)	0	100	295,671
Self	53.801	(20.963)	0	100	295,671
d_{im}	15.841	(12.495)	0	100	295,671

Notes: Mean and standard deviations (SD; in parentheses). The sample covers the full sample of N=25,412 tests with an average of about 11.6 modules per test/individual.

A.3 Additional Results

Table A.6: Imperfect self-knowledge, alternative skill mismatch measures.

	(1) Skill Mismatch	(2) Change Anything	(3) Difference in Mismatch	(4) Realized Mismatch	(5) Dropout first yr.
Imp. self-knowl. (school, std)	0.046***	0.014***	0.303**	0.034***	0.003
	(0.006)	(0.004)	(0.118)	(0.007)	(0.003)
Male	0.184***	0.064***	-0.914***	0.164***	0.036***
	(0.013)	(0.008)	(0.304)	(0.016)	(0.007)
Test score	-0.239***	-0.021***	-0.799***	-0.167***	-0.016***
	(0.005)	(0.003)	(0.112)	(0.007)	(0.003)
Test score (sq)	0.002***	0.000***	0.005***	0.001***	0.000***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Realized mismatch (std)					0.003
					(0.004)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Realized occupation FE			✓		✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	0.000	0.435	1.507	-0.000	0.133
Observations	24,282	17,159	4,500	17,990	12,326

Notes: Outcomes are those from Hypotheses 1 to 4 and the analyses using early labor market outcomes, but constructed using alternative occupational skill requirements for math and first language (see Section 3.5, for the construction of the mismatch measures). The samples are larger because the occupational skill requirements are available for a larger set of occupations. For consistency reasons, imperfect self-knowledge (akm-type variant B) is measured in school knowledge skills (and not in potential and/or professional knowledge skills). (std) = standardized, (sq) = squared. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. Robust standard errors in parentheses. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.7: Repeater table.

	(1) Repeat test	(2) Repeat diff. test	(3) Test Score	(4) Skill Mismatch	(5) Skill Mismatch	(6) Imp. self- knowl.
Imp. self-knowl.	0.005**	0.003**				
	(0.002)	(0.002)				
Second execution			4.400***	-1.937***	-1.555***	0.010
			(0.219)	(0.156)	(0.171)	(0.026)
Test score	0.004**	-0.001			-0.087***	
	(0.002)	(0.001)			(0.020)	
All controls	✓	✓				
Occupation FE	✓	✓				
Cohort FE	✓	✓				
Individual FE			✓	✓	✓	✓
Sample	Full	Full	Repeater	Repeater	Repeater	Repeater
Observations	25,412	25,412	3,246	3,246	3,246	3,246

Notes: Columns (1) and (2) use the full sample, and the outcomes are binary indicators for whether the individual did a second aptitude test afterward, in any test (1), respectively, in another test type (2). In columns (3) - (6), we use two observations per individual, the first and the second test execution. The sample is restricted to those who took the second test in a different occupation than the first ('Repeater'). Imperfect self-knowledge is standardized. Robust standard errors in parentheses. 'All controls' includes the variables: test score (sq), male, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.8: Imperfect self-knowledge and dropout, Sensitivity.

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.008***	0.006**	0.006**	0.006**	0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Male		0.041***	0.038***	0.034***	0.037***
		(0.007)	(0.007)	(0.006)	(0.007)
Test score		-0.017***	-0.016***	-0.016***	-0.008**
		(0.003)	(0.003)	(0.003)	(0.004)
Test score (sq)		0.000***	0.000***	0.000***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Realized Mismatch (std)					0.026***
					(0.007)
All controls				✓	✓
Occupation FE		✓	✓		✓
Realized occupation FE		✓	✓		✓
Cohort FE		✓	✓		✓
Mean of outcome	0.120	0.120	0.120	0.120	0.120
Sample	1yr-Work	1yr-Work	1yr-Work	1yr-Work	1yr-Work
Observations	11,048	11,048	11,048	11,048	11,048

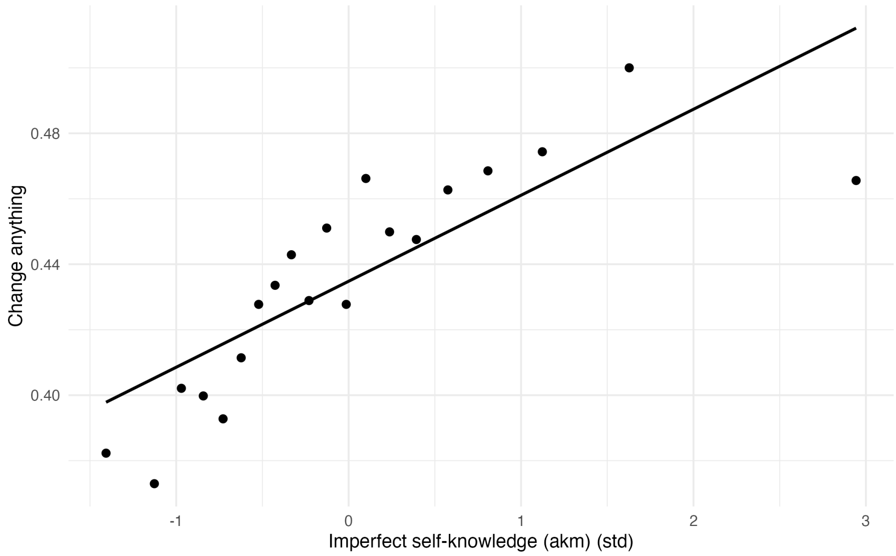
Notes: Outcome: Dropout within 1 year. The sample is restricted to those who started an apprenticeship and can be observed for at least 1 year ('1yr-Work'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.9: Causal mediation analysis of imperfect self-knowledge on dropout via realized skill mismatch.

Effect	Mean	90% CI Lower	90% CI Upper
Mediation effect	0.0007	0.00043	0.0010
Direct effect	0.0058	0.0012	0.0112
Total effect	0.0065	0.0019	0.0119
% mediated	10.224	3.882	34.515

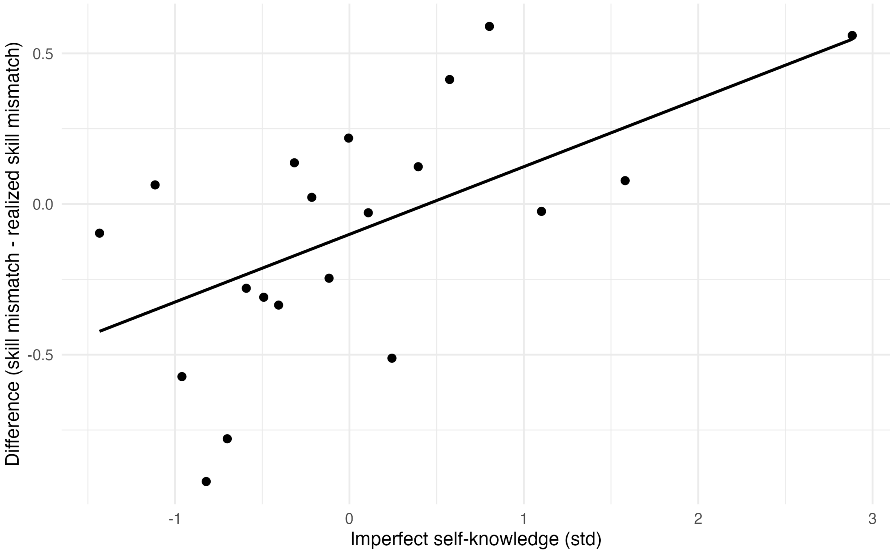
Notes: Outcome is the dropout within the first year. The sample is restricted to those who started an apprenticeship and who we can observe for at least 1 year. N = 11,048. 90% confidence intervals calculated in 999 bootstrap replications. We controlled for the variables: test score, test score squared, male, occupation, realized occupation, and cohort fixed effects, an indicator whether the individual changed the job, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators.

Figure A.1: Imperfect self-knowledge and change anything, unconditional



Notes: Binscatter plot with 20 equally sized bins showing the raw relationship between decision to change away from the initial occupational choice ("change anything"), on the y-axis, and the standardized (std) imperfect self-knowledge Variant B (akm-type) on the x-axis. The sample is restricted to 2019-2021 10th-grade cohorts who completed 9th grade and took the test before the end of 2023. This restriction is because, for the 2022 cohort and some repeaters, we do not observe all the options for change; N=17,159.

Figure A.2: Imperfect self-knowledge and improvement in fit for changers, unconditional



Notes: Binscatter plot with 20 equally sized bins showing the difference in skill mismatch and realized skill mismatch (not standardized), on the y-axis, and the standardized (std) imperfect self-knowledge Variant B (akm-type) on the x-axis. The sample is restricted to those who started an apprenticeship in an occupation other than the initial occupation; N=2,961.

A.4 Robustness checks

Table A.10: Imperfect self-knowledge and Mismatch, Robustness

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.074***	0.050***	0.070***	0.050***	0.068***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Imp. self-knowledge (std, sq)			-0.013***		-0.012***
			(0.002)		(0.002)
Male	0.084***	0.063***	0.062***	0.063***	0.062***
	(0.013)	(0.010)	(0.010)	(0.010)	(0.010)
Test score	-0.035***	-0.498***	-0.498***		
	(0.001)	(0.005)	(0.005)		
Test score (sq)		0.004***	0.004***		
		(0.000)	(0.000)		
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Test score FE				✓	✓
Observations	25,412	25,412	25,412	25,412	25,412

Notes: Outcome: Skill mismatch. (std) = standardized, (sq) = squared. Variant B Imperfect self-knowledge (akm-type). Test score FE discretizes the raw test score variable into 100 percentile indicators. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.11: Imperfect self-knowledge and realized skill mismatch, Robustness

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.062***	0.047***	0.068***	0.047***	0.067***
	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Imp. self-knowledge (sq, std)			-0.013***		-0.013***
			(0.003)		(0.003)
Male	0.076***	0.066***	0.065***	0.067***	0.066***
	(0.018)	(0.015)	(0.015)	(0.015)	(0.015)
Test score	-0.022***	-0.503***	-0.503***		
	(0.001)	(0.007)	(0.007)		
Test score (sq)		0.004***	0.004***		
		(0.000)	(0.000)		
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Test score FE				✓	✓
Sample	Worker	Worker	Worker	Worker	Worker
Observations	16,083	16,083	16,083	16,083	16,083

Notes: Outcome: Skill mismatch to realized occupation. The sample is restricted to those who started an apprenticeship in either occupation for which we have information on the occupational skill requirements ('Worker'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.12: Imperfect self-knowledge and change anything, Robustness

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.015***	0.015***	0.017***	0.014***	0.017***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Imp. self-knowledge (std, sq)			-0.002		-0.002
			(0.002)		(0.002)
Male	0.065***	0.065***	0.065***	0.065***	0.065***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Test score	-0.010***	-0.021***	-0.021***		
	(0.000)	(0.003)	(0.003)		
Test score (sq)		0.000***	0.000***		
		(0.000)	(0.000)		
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Test score FE				✓	✓
Mean of outcome	0.435	0.435	0.435	0.435	0.435
Sample	2019-2021	2019-2021	2019-2021	2019-2021	2019-2021
Observations	17,159	17,159	17,159	17,159	17,159

Notes: Outcome: Changed away from initial occupation (anywhere). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. The sample is restricted to 2019-2021 8th-grade cohorts who completed 9th grade and took the test before the end of 2023. This restriction is because, for the 2022 cohort and some repeaters, we do not observe all the options for change. Test score FE discretizes the raw test score variable into 100 percentile indicators. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.13: Imperfect self-knowledge and difference in mismatch, Robustness

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.154**	0.145**	0.144	0.167**	0.172*
	(0.070)	(0.070)	(0.089)	(0.071)	(0.088)
Imp. self-knowledge (sq, std)			0.001		-0.004
			(0.038)		(0.038)
Male	-0.279	-0.303*	-0.303*	-0.383**	-0.383**
	(0.178)	(0.178)	(0.178)	(0.177)	(0.177)
Test score	-0.118***	-0.327***	-0.327***		
	(0.010)	(0.069)	(0.069)		
Test score (sq)		0.002***	0.002***		
		(0.001)	(0.001)		
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Test score FE				✓	✓
Mean of outcome	-0.095	-0.095	-0.095	-0.095	-0.095
Sample	Changer	Changer	Changer	Changer	Changer
Observations	2,961	2,961	2,961	2,961	2,961

Notes: Outcome: difference in skill mismatch vs realized skill mismatch. The sample is restricted to those who started an apprenticeship in an occupation other than the initial occupation ('Changer'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.14: Imperfect self-knowledge and Dropout, Robustness

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (std)	0.007**	0.006**	0.004	0.006**	0.004
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Imp. self-knowledge (sq ,std)			0.001		0.001
			(0.002)		(0.002)
Male	0.039***	0.038***	0.038***	0.038***	0.038***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Test score	-0.004***	-0.016***	-0.016***		
	(0.000)	(0.003)	(0.003)		
Test score (sq)		0.000***	0.000***		
		(0.000)	(0.000)		
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Realized occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Test score FE				✓	✓
Mean of outcome	0.120	0.120	0.120	0.120	0.120
Sample	1yr-Work	1yr-Work	1yr-Work	1yr-Work	1yr-Work
Observations	11,048	11,048	11,048	11,048	11,048

Notes: Outcome: Dropout within 1 year. The sample is restricted to those who started an apprenticeship and who we can observe for at least 1 year in the apprenticeship ('1yr-Work'). (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.15: Imperfect self-knowledge (Variant A) and skill mismatch

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowledge (MAE, std)	0.121***	0.120***	0.072***	0.047***	0.049***
	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)
Male			0.042***	0.048***	0.061***
			(0.012)	(0.009)	(0.010)
Test score			-0.034***	-0.495***	-0.498***
			(0.001)	(0.005)	(0.005)
Test score (sq)				0.004***	0.004***
				(0.000)	(0.000)
All controls				✓	✓
Occupation FE		✓			✓
Cohort FE		✓			✓
Observations	25,412	25,412	25,412	25,412	25,412

Notes: Outcome: Skill mismatch. (MAE) = mean absolute error; refers to imperfect self-knowledge variant A. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.16: Imperfect self-knowledge (Variant A) and realized skill mismatch

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowl. (MAE, std)	0.090***	0.084***	0.062***	0.047***	0.046***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.006)
Male			0.016	0.027***	0.065***
			(0.016)	(0.013)	(0.015)
Test score			-0.024***	-0.505***	-0.503***
			(0.001)	(0.007)	(0.007)
Test score (sq)				0.004***	0.004***
				(0.000)	(0.000)
All controls				✓	✓
Occupation FE		✓			✓
Cohort FE		✓			✓
Sample	Worker	Worker	Worker	Worker	Worker
Observations	16,083	16,083	16,083	16,083	16,083

Notes: Outcome: Realized skill mismatch. The sample is restricted to those who started an apprenticeship in either occupation for which we have information on the occupational skill requirements ('Worker'). (MAE) = mean absolute error; refers to imperfect self-knowledge variant A. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.17: Imperfect self-knowledge (Variant A) and changing away from initial choice

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowl. (MAE, std)	0.012***	0.007**	0.001	0.001	0.003**
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Male	0.064***	0.073***	-0.003	-0.004	-0.001
	(0.008)	(0.008)	(0.002)	(0.003)	(0.003)
Test score	-0.021***	-0.019***	-0.003***	0.002**	-0.001
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Test score (sq)	0.000***	0.000***	0.000***	-0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	0.435	0.357	0.020	0.031	0.027
Effect as % of mean	2.9	2.1	4.3	3.2	11.8
Sample	2019-2021	2019-2021	2019-2021	2019-2021	2019-2021
Observations	17,159	17,159	17,159	17,159	17,159

Notes: The sample is restricted to 2019-2021 10th-grader cohorts who completed 11th grade and took their test before the end of 2023. This restriction is because, for the 2022 cohort and some repeaters, we do not observe all the options for change. (MAE) = mean absolute error; refers to imperfect self-knowledge variant A. Outcomes: *anything* is 0 for those starting in the tested occupation, and 1 otherwise; *job* refers to another occupation than the test occupation; *high school* is the academic high track, *other school* formal education paths that are more specialized or technical, *other* includes transitional education programs from secondary level I to secondary level II and motivational semesters. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.18: Imperfect self-knowledge (Variant A) and difference in skill mismatch

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowl. (MAE, std)	0.287***	0.305***	0.153**	0.150**	0.170**
	(0.077)	(0.076)	(0.071)	(0.071)	(0.070)
Male			-0.352**	-0.303*	-0.107
			(0.178)	(0.179)	(0.185)
Test score			-0.301***	-0.317***	-0.332***
			(0.070)	(0.071)	(0.068)
Test score (sq)			0.002**	0.002***	0.002***
			(0.001)	(0.001)	(0.001)
All controls				✓	✓
Occupation FE		✓			✓
Realized occupation FE			✓	✓	✓
Cohort FE		✓			✓
Mean of outcome	-0.095	-0.095	-0.095	-0.095	-0.095
Sample	Changer	Changer	Changer	Changer	Changer
Observations	2,961	2,961	2,961	2,961	2,961

Notes: Outcome: Skill mismatch - Realized skill mismatch. The sample is restricted to those who started an apprenticeship in a different than the test occupation ('Changer'). (MAE) = mean absolute error; refers to imperfect self-knowledge variant A. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.19: Imperfect self-knowledge and dropout

	(1)	(2)	(3)	(4)	(5)
Imp. self-knowl. (MAE, std)	0.009***	0.006**	0.007**	0.006*	0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Male		0.041***	0.038***	0.034***	0.037***
		(0.007)	(0.007)	(0.006)	(0.007)
Test score		-0.017***	-0.016***	-0.016***	-0.008**
		(0.003)	(0.003)	(0.003)	(0.004)
Test score (sq)		0.000***	0.000***	0.000***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Realized Mismatch					0.026***
					(0.007)
All controls				✓	✓
Occupation FE		✓	✓		✓
Realized occupation FE		✓	✓		✓
Cohort FE		✓	✓		✓
Mean of outcome	0.120	0.120	0.120	0.120	0.120
Sample	1yr-Work	1yr-Work	1yr-Work	1yr-Work	1yr-Work
Observations	11,048	11,048	11,048	11,048	11,048

Notes: Outcome: Dropout within 1 year. The sample is restricted to those who started an apprenticeship and who we can observe for at least 1 year in the apprenticeship ('1yr-Work'). (MAE) = mean absolute error; refers to imperfect self-knowledge variant A. (std) = standardized, (sq) = squared. Robust standard errors in parentheses. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. FE-variables included as binary indicators. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.20: Robustness to canton fixed effects

	(1) Skill Mismatch	(2) Change Anything	(3) Difference in Mismatch	(4) Realized Mismatch	(5) Dropout first yr.
Imp. self-knowledge (std)	0.050***	0.015***	0.161**	0.046***	0.006**
	(0.005)	(0.003)	(0.069)	(0.006)	(0.003)
Male	0.062***	0.070***	-0.079	0.061***	0.037***
	(0.010)	(0.008)	(0.185)	(0.015)	(0.007)
Test score	-0.497***	-0.022***	-0.329***	-0.502***	-0.015***
	(0.005)	(0.003)	(0.068)	(0.007)	(0.003)
Test score (sq)	0.004***	0.000***	0.002***	0.004***	0.000***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Realized occupation FE			✓		✓
Canton FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	0.000	0.435	-0.095	0.000	0.120
Observations	25,412	17,159	2,961	16,083	11,048

Notes: Outcomes and samples are those from Hypotheses 1 to 4 and the analyses using early labor market outcomes. This table replicates the main results, but additionally controls for canton (i.e., the Swiss political regions) fixed effects. (std) = standardized, (sq) = squared. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. Robust standard errors in parentheses. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.21: Robustness to school fixed effects

	(1) Skill Mismatch	(2) Change Anything	(3) Difference in Mismatch	(4) Realized Mismatch	(5) Dropout first yr.
Imp. self-knowledge (std)	0.049***	0.009***	0.140	0.049***	0.006*
	(0.005)	(0.004)	(0.087)	(0.007)	(0.003)
Male	0.062***	0.070***	0.002	0.068***	0.025***
	(0.011)	(0.008)	(0.224)	(0.015)	(0.008)
Test score	-0.493***	-0.022***	-0.321***	-0.497***	-0.016***
	(0.005)	(0.003)	(0.084)	(0.007)	(0.003)
Test score (sq)	0.004***	0.000***	0.002**	0.004***	0.000***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Realized occupation FE			✓		✓
School FE	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	0.000	0.435	-0.095	0.000	0.120
Observations	25,412	17,159	2,961	16,083	11,048

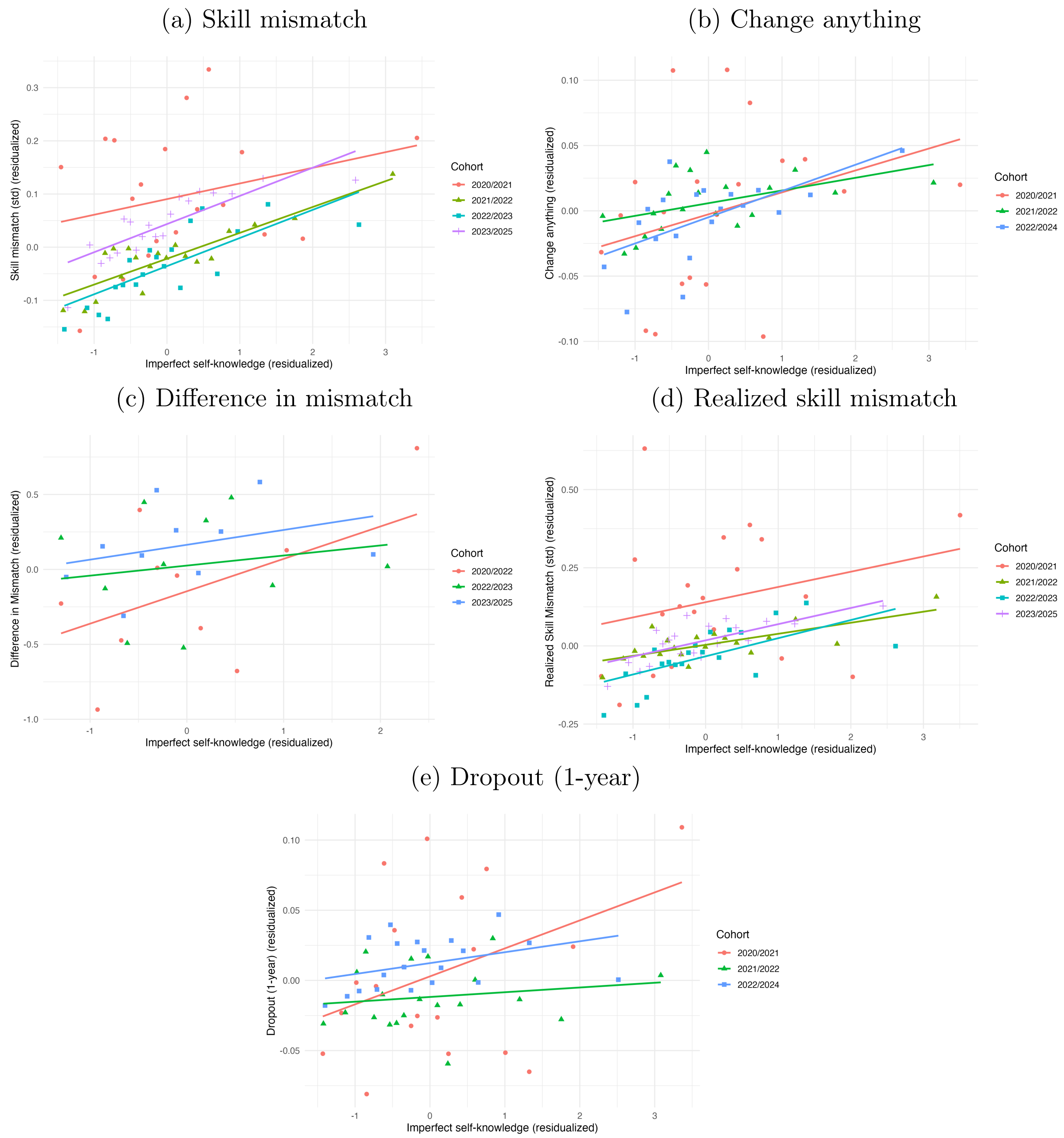
Notes: Outcomes and samples are those from Hypotheses 1 to 4 and the analyses using early labor market outcomes. This table replicates the main results, but additionally controls for school fixed effects. (std) = standardized, (sq) = squared. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. Robust standard errors in parentheses. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Table A.22: Robustness to treatment measurement

	(1) Skill Mismatch	(2) Change Anything	(3) Difference in Mismatch	(4) Realized Mismatch	(5) Dropout first yr.
Imp. self-knowl.§ (std)	0.049***	0.014***	0.142**	0.047***	0.006**
	(0.005)	(0.003)	(0.070)	(0.006)	(0.003)
Male	0.063***	0.065***	-0.105	0.066***	0.038***
	(0.010)	(0.008)	(0.185)	(0.015)	(0.007)
Test score	-0.498***	-0.021***	-0.334***	-0.503***	-0.016***
	(0.005)	(0.003)	(0.068)	(0.007)	(0.003)
Test score (sq)	0.004***	0.000***	0.002***	0.004***	0.000***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
All controls	✓	✓	✓	✓	✓
Occupation FE	✓	✓	✓	✓	✓
Realized occupation FE			✓		✓
Cohort FE	✓	✓	✓	✓	✓
Mean of outcome	-0.000	0.435	-0.095	-0.000	0.120
Observations	25,412	17,159	2,961	16,083	11,048

Notes: Outcomes and samples are those from Hypotheses 1 to 4 and the analyses using early labor market outcomes. This table replicates the main results using an alternative measure of imperfect self-knowledge. Specifically, imperfect self-knowledge§ is based on single-module test scores rounded to the nearest multiple of ten, i.e., $score^{\S}_{im} \in 0, 10, \ldots, 100$, to correspond to the scale on which self-beliefs are measured. (std) = standardized, (sq) = squared. 'All controls' includes the variables: born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language. Robust standard errors in parentheses. FE-variables included as binary indicators. * p<0.10, ** p<0.05, *** p<0.01.

Figure A.3: Effect of imperfect self-knowledge by cohort



Notes: Outcomes and samples are those from Hypotheses 1 to 4 and the analyses using the early labor market outcome. These subfigures replicate the main results, but separately for each cohort. Very small cohorts are combined with the subsequent or previous cohort. The 2020/2021 cohort in our data started in January 2021. Occupation fixed effects and in (c) and (e) realized occupation fixed effects, and control variables (test score, test score squared, male, born Swiss, age at test, born in Switzerland, company voucher, first language = local language, and first language = national language) are absorbed.

Appendix B: Lemmas and Proofs for Section 2

This appendix collects all formal lemmas and proof sketches underlying the model predictions presented in Section 2.2. All results are derived under the assumptions stated in Section 2 and are unchanged.

B.1 Definitions and Preliminaries

For each individual i, the true optimal occupation is

$$j_i^* \in \arg\max_{j \in J} \theta_{ij}.$$

The individual is mismatched after the first choice if $j_i^{(1)} \neq j_i^*$. Define the mismatch and switching indicators

$$M_i = \mathbf{1}\{j_i^{(1)} \neq j_i^*\}, \qquad S_i = \mathbf{1}\{j_i^{(2)} \neq j_i^{(1)}\}.$$

Define the final mismatch indicator:

$$M_i^{\text{final}} = \mathbf{1}\{j_i^{(2)} \neq j_i^*\}.$$

B.2 Information Frictions and Initial Mismatch

Lemma 1 (Information frictions and initial mismatch). Suppose that Assumptions 1–2 hold and that the distribution of $(\theta_{ij})_{j \in J}$ is non-degenerate. Then, for all parameter values with $\sigma_H^2 < \sigma_L^2$,

$$\Pr(M_i = 1 \mid t_i = L) > \Pr(M_i = 1 \mid t_i = H). \tag{16}$$

Proof sketch. From (2), the period-1 estimates of match quality are affine transformations of $(\theta_{ij}, \varepsilon_{ij}^{(t)})$:

$$\hat{\theta}_{ij}^{(0)}(t, s_{ij}^{(0)}) = \mu_j + w_t(\theta_{ij} - \mu_j) + w_t \varepsilon_{ij}^{(t)}.$$

Fix an individual i and two occupations j and k. The sign of the difference in estimates is

$$\hat{\theta}_{ij}^{(0)}(t, s_{ij}^{(0)}) - \hat{\theta}_{ik}^{(0)}(t, s_{ik}^{(0)}) = w_t\big[(\theta_{ij} - \theta_{ik}) + (\varepsilon_{ij}^{(t)} - \varepsilon_{ik}^{(t)})\big] + (1 - w_t)(\mu_j - \mu_k).$$

Conditional on $(\theta_{ij}, \theta_{ik})$, this difference is a normal random variable whose mean is increasing in $\theta_{ij} - \theta_{ik}$ and whose conditional variance is proportional to $w_t^2(\sigma_t^2 + \sigma_t^2)$. Since w_t is decreasing in σ_t^2, high-information types ($t = H$) have differences in estimated match quality that are more tightly concentrated around the actual differences $(\theta_{ij} - \theta_{ik})$ than low-information types.

In other words, the ranking of occupations induced by $(\hat{\theta}_{ij}^{(0)})_{j \in J}$ is more likely to coincide with the ranking induced by $(\theta_{ij})_{j \in J}$ when σ_t^2 is lower. Formally, for any pair of occupations (j, k), the probability that the estimate-based ordering agrees with the true ordering, $\Pr(\hat{\theta}_{ij}^{(0)} > \hat{\theta}_{ik}^{(0)} \mid \theta_{ij} > \theta_{ik}, t)$, is strictly decreasing in σ_t^2. Aggregating over all pairs and using the fact that $j_i^{(1)}$ maximizes $\hat{\theta}_{ij}^{(0)}$, we obtain (16). □

Lemma 1 states that Roy-style selection based on noisy self-information sorts individuals more efficiently when they face weaker information frictions.[12] In terms of observable outcomes, this leads to the following empirical hypothesis.

Hypothesis 1 (Information frictions and initial mismatch). Individuals with higher information frictions are more likely to choose an initially mismatched occupation.

B.3 Information Frictions and Occupational Switching

Lemma 2 (Information frictions and switching). *Suppose that Assumptions 1–2 hold, that $c \geq 0$ is not too large, and that $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$. Then there exists $\bar{c} > 0$ such that for all $c \in [0, \bar{c})$,*

$$\Pr(S_i = 1 \mid t_i = L) \; > \; \Pr(S_i = 1 \mid t_i = H). \tag{17}$$

Proof sketch. When σ_1^2 is small, the additional signals $s_{ij}^{(1)}$ are very precise, so the updated estimates $\hat{\theta}_{ij}^{(1)}$ are close to the true match qualities θ_{ij} for all j, regardless of t_i. In the limit as $\sigma_1^2 \to 0$, $\hat{\theta}_{ij}^{(1)}(t_i, s_{ij}^{(0)}, s_{ij}^{(1)}) \to \theta_{ij}$ almost surely. Thus, ignoring switching costs for the moment, the second-period optimal occupation $j_i^{(2)}$ converges to the true optimal occupation j_i^* for both types.

[12] This is consistent with the general insight that better information improves the allocation of workers to occupations or careers in models of learning about match quality and career choice (Jovanovic, 1979; Miller, 1984; Neal, 1999; Arcidiacono, 2004; Papageorgiou, 2014). Our setting emphasizes cross-sectional heterogeneity in information precision at the time of the first choice.

If $c = 0$, then, by the switching rule in period 2,

$$S_i = 1 \iff j_i^{(2)} \neq j_i^{(1)} \iff j_i^{(1)} \neq j_i^*,$$

so that

$$\Pr(S_i = 1 \mid t_i) = \Pr(M_i = 1 \mid t_i),$$

and Lemma 1 implies equation (17). For small positive c, only those individuals whose gain from switching,

$$G_i = \max_{k \in J} \hat{\theta}_{ik}^{(1)} - \hat{\theta}_{ij_i^{(1)}}^{(1)},$$

exceeds c will switch. Since initial choices are less accurate for low-information types, the distribution of G_i among initially mismatched individuals is more dispersed and has more mass on large values for type L than for type H. Hence, there exists $\bar{c} > 0$ such that, for any $c \in [0, \bar{c})$, the probability that an initially mismatched individual finds it profitable to switch remains higher for type L than for type H. Therefore, $\Pr(S_i = 1 \mid t_i = L) > \Pr(S_i = 1 \mid t_i = H)$ for all sufficiently small c. $\square$

Lemma 2 captures the idea that individuals who start from a poorer information base are more likely to discover, once better information becomes available, that their initial occupation was a mistake and therefore have stronger incentives to switch jobs.[13] This motivates the following empirical hypothesis.

Hypothesis 2 (Information frictions and switching after improved information). Conditional on receiving more precise information about their abilities, individuals with higher initial information frictions are more likely to switch away from the initial occupations.

[13]This logic resonates with empirical evidence on career and major changes and early-career job mobility in the presence of learning about comparative advantage and match quality. Some studies (Neal, 1999; Topel & Ward, 1992; Arcidiacono, 2004; Papageorgiou, 2014; Fredriksson et al., 2018) document that early career and major choices are frequently revised as individuals learn about their match quality and returns across occupations or jobs. Our contribution is to emphasize that heterogeneity in initial self-information about abilities generates systematic differences in both initial mismatch and subsequent switching behavior.

B.4 Information Frictions and Gains from Switching

Lemma 3 (Information frictions and gains from switching). Define the expected gain from switching for individual i as

$$G_i \;=\; \max_{k\in J}\hat{\theta}^{(1)}_{ik} \;-\; \hat{\theta}^{(1)}_{ij_i^{(1)}},$$

so that $G_i \geq 0$ and $S_i = 1$ if and only if $G_i > c$. Suppose that Assumptions 1–2 hold, that $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$, and that the switching cost $c \geq 0$ is not too large. Then there exists $\bar{c} > 0$ such that for all $c \in [0, \bar{c})$,

$$\mathbb{E}[G_i \mid S_i = 1, t_i = L] \;>\; \mathbb{E}[G_i \mid S_i = 1, t_i = H]. \tag{18}$$

Proof sketch. When σ_1^2 is small, the additional signals $s_{ij}^{(1)}$ are very precise, so the updated estimates $\hat{\theta}^{(1)}_{ij}$ are close to the true match qualities θ_{ij} for all j, regardless of t_i. In the limit as $\sigma_1^2 \to 0$,

$$\hat{\theta}^{(1)}_{ij}(t_i, s^{(0)}_{ij}, s^{(1)}_{ij}) \to \theta_{ij}$$

almost surely, so that G_i converges to the true gain from moving from $j_i^{(1)}$ to the best occupation j_i^*,

$$G_i \;\to\; \theta_{ij_i^*} - \theta_{ij_i^{(1)}} \;\geq\; 0.$$

Lemma 1 implies that low-information types are not only more likely to be initially mismatched ($j_i^{(1)} \neq j_i^*$), but also that their first choices deviate more from j_i^* in terms of match quality on average: the distribution of the gap $\theta_{ij_i^*} - \theta_{ij_i^{(1)}}$ among initially mismatched individuals is stochastically larger for type L than for type H. Intuitively, because low-information types start from a noisier ranking of occupations, the mistakes they make tend to be more severe.

For $c = 0$, we have $S_i = 1$ if and only if $j_i^{(1)} \neq j_i^*$, so that among switchers

$$G_i \;\to\; \theta_{ij_i^*} - \theta_{ij_i^{(1)}},$$

and the stochastic dominance of this gap for type L implies

$$\mathbb{E}[G_i \mid S_i = 1, t_i = L] \;>\; \mathbb{E}[G_i \mid S_i = 1, t_i = H]$$

in the limit as $\sigma_1^2 \to 0$. For small positive c, switchers are those with $G_i > c$. Since the distribution of G_i for type L among initially mismatched individuals puts more mass on large gains than for type H, the conditional expectation of G_i given $G_i > c$ remains larger for type L than for type H for all c in a neighborhood of zero. This yields (18). $\square$

Lemma 3 shows that, conditional on switching, individuals who started from a poorer information base tend to move further towards their optimal occupation than those who were initially better informed. Given the same underlying vector of occupation-specific talents, weaker initial self-information therefore implies larger expected gains from revising the occupational choice once better information becomes available.[14] This leads to the following empirical hypothesis.

Hypothesis 3 (Information frictions and gains from switching). Among individuals who switch occupations, those with higher initial information frictions experience larger expected improvements in match quality, i.e., larger reductions in mismatch.

B.5 Information Frictions and Residual Mismatch

Lemma 4 (Information frictions and residual mismatch). Define the final mismatch indicator

$$M_i^{\text{final}} = \mathbf{1}\{j_i^{(2)} \neq j_i^*\}.$$

Suppose that Assumptions 1–2 hold, that $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$, and that the switching cost $c \geq 0$ is not too large. Then there exists $\bar{c} > 0$ such that for all $c \in (0, \bar{c})$,

$$\Pr\left(M_i^{\text{final}} = 1 \mid t_i = L\right) \ > \ \Pr\left(M_i^{\text{final}} = 1 \mid t_i = H\right). \tag{19}$$

Proof sketch. By definition,

$$M_i^{\text{final}} = 1 \quad \Longleftrightarrow \quad j_i^{(2)} \neq j_i^* \quad \Longleftrightarrow \quad \text{the individual does not end up in } j_i^*.$$

[14]This aligns with models in which learning about worker skills and comparative advantage drives sectoral sorting and wage growth, and where improvements in information lead to more efficient allocation over the life cycle (Gibbons, Katz, Lemieux, & Parent, 2005; Yamaguchi, 2010, for example). Empirically, recent work documents that match quality and comparative advantage are important drivers of wage growth and job mobility (Fredriksson et al., 2018).

When σ_1^2 is small, the updated estimates $\hat{\theta}_{ij}^{(1)}$ are very precise, so that for any initially mismatched individual the gain from switching to j_i^* is close to the true gain $\theta_{ij_i^*} - \theta_{ij_i^{(1)}}$. If $c = 0$, all initially mismatched individuals switch to j_i^* in the second period and thus $M_i^{\text{final}} = 0$ almost surely for both types. For small positive c, however, some initially mismatched individuals will find that their gain from switching, G_i, does not justify paying the cost c and will therefore remain in a suboptimal occupation.

Lemma 1 implies that the probability of being initially mismatched, $\Pr(M_i = 1 \mid t_i)$, is strictly larger for type L than for type H. Lemma 2 implies that low-information types are also more likely to switch in response to the improved information. Combining these results, the probability of remaining mismatched after the second decision can be written as

$$\Pr(M_i^{\text{final}} = 1 \mid t_i) = \Pr\left(M_i = 1,\, G_i \leq c \mid t_i\right).$$

For small c, the event $\{G_i \leq c\}$ corresponds to a narrow band of initially mismatched individuals whose potential gains from switching are very small. Under mild regularity (continuous distributions of match qualities and signals), the probability of falling into this band is of order c for both types, whereas the overall mass of initially mismatched individuals is strictly larger for type L than for type H by Lemma 1. Hence, for all sufficiently small $c > 0$,

$$\Pr(M_i^{\text{final}} = 1 \mid t_i = L) \;>\; \Pr(M_i^{\text{final}} = 1 \mid t_i = H),$$

which yields equation (19). Intuitively, even though low-information types are more likely to correct their initial mistakes (Lemma 2), they start from a worse allocation and, as long as switching costs prevent full adjustment, they end up with a higher residual mismatch rate after the second decision. □

Lemma 4 emphasizes that heterogeneous information frictions generate persistent differences in realized match quality and mobility behavior, even after an information update. Individuals who start from a poorer information base remain more likely to be mismatched, both because they make more mistakes initially and because some of those mistakes are not corrected when switching is costly.[15] This leads to the following empirical hypothesis.

[15]This mechanism resonates with evidence that skill mismatch can persist over time, and that adjustment is incomplete, in particular when mobility is costly or constrained. On the persistence and macroeconomic relevance of mismatch in the presence of information and search frictions, see Baley,

Hypothesis 4 (Information frictions and realized mismatch). Even after the information update and the possibility to switch, individuals with higher initial information frictions remain more mismatched on average.

Figueiredo, and Ulbricht (2022). For broader surveys of skill mismatch, its measurement, and its persistence over the business cycle and over workers' careers, see Brunello and Wruuck (2021) and the references cited therein.

B.6 Comparative Statics with a Continuous Skill-Based Mismatch Index

In the main text, mismatch is defined as a binary indicator $M_i = \mathbf{1}\{j_i^{(1)} \neq j_i^*\}$, where j_i^* denotes the occupation that maximizes the individual's true match quality. In the empirical analysis, however, we use a continuous, skill-based mismatch index that captures the distance between individual skills and the skill requirements of the chosen occupation. This appendix shows that the comparative-statics results derived in the main text carry over when the mismatch is measured in this continuous way.

Definition B.1 (Skill-based mismatch index). Let x_i denote a (scalar) index of individual i's skills or competencies, and let q_j denote the corresponding skill requirement level in occupation j (for example, obtained from external data such as occupational skill surveys). For a given occupational choice $j_i^{(1)}$, define the skill-based mismatch index as

$$m_i \;=\; \big|x_i - q_{j_i^{(1)}}\big|.$$

More generally, one could replace the absolute value by any distance measure $d(x_i, q_{j_i^{(1)}})$ that increases with the discrepancy between individual skills and job requirements.

To link this empirical index to the theoretical match quality θ_{ij}, we impose the following mild monotonicity assumption.

Assumption B.1 (Match quality and skill distance). For each occupation $j \in J$, there exists a strictly decreasing function $g_j : [0, \infty) \to \mathbb{R}$ such that

$$\theta_{ij} = g_j\big(|x_i - q_j|\big).$$

Thus, match quality in occupation j is a strictly decreasing function of the distance between individual skills and occupation-specific skill requirements. In particular, for given x_i, the occupation that maximizes θ_{ij} is the one that minimizes $|x_i - q_j|$.

Under Assumption B.1, the binary theoretical notion of mismatch in the main text ($j_i^{(1)} \neq j_i^*$) is equivalent to choosing an occupation whose skill requirement is not closest to x_i, whereas the continuous index m_i measures how far the chosen occupation's skill requirement is from the individual's skills.

Lemma B.1 (Information frictions and expected skill mismatch). *Suppose that Assumptions 1-2 from the main text and Assumption B.1 hold, and that the distribution of* $(x_i)_i$ *and* $(q_j)_{j\in J}$ *is such that the distribution of* $(\theta_{ij})_{j\in J}$ *is non-degenerate. Then, for all parameter values with* $\sigma_H^2 < \sigma_L^2$*, we have*

$$\mathbb{E}[m_i \mid t_i = L] \;>\; \mathbb{E}[m_i \mid t_i = H]. \tag{20}$$

Proof sketch. By Assumption B.1, for each individual i, there is a one-to-one, strictly decreasing mapping between match quality θ_{ij} and skill distance $|x_i - q_j|$ for each occupation j. The true optimal occupation j_i^* minimizes the skill distance $|x_i - q_j|$ and maximizes θ_{ij}.

Lemma 1 in the main text establishes that, because high-information types have more precise initial self-signals (larger w_H), their first-period choices $j_i^{(1)}$ are more likely to coincide with j_i^* than those of low-information types. Equivalently, low-information types are more likely to choose occupations whose skill requirements are not closest to their skills.

Condition on the vector $(x_i, (q_j)_{j\in J})$. When $j_i^{(1)} = j_i^*$, the mismatch index is $m_i = |x_i - q_{j_i^*}|$, the minimal achievable distance for that individual. When $j_i^{(1)} \neq j_i^*$, the mismatch index is strictly larger: $|x_i - q_{j_i^{(1)}}| > |x_i - q_{j_i^*}|$, because j_i^* minimizes the distance. Since low-information types have a higher probability of choosing $j_i^{(1)} \neq j_i^*$ than high-information types, the distribution of m_i for type L puts more mass on larger distances than for type H. This implies that the conditional distribution of m_i given $t_i = H$ first-order stochastically dominates the distribution of m_i given $t_i = L$, and in particular Lemma B.1. $\square$

Lemma B.1 shows that the first comparative static in the main text – that stronger information frictions lead to a larger initial mismatch – extends naturally to the continuous, skill-based mismatch index m_i. Low-information types not only have a higher probability of being mismatched in the binary sense but also exhibit, on average, a larger distance between their skills and the skill requirements of their initially chosen occupation.

Lemma B.2 (Information frictions, skill mismatch, and switching). *Suppose that Assumptions 1-2 from the main text and Assumption B.1 hold, that* $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$*, and* *that the switching cost $c \geq 0$ is not too large. Then there exists $\bar{c} > 0$ such that for all $c \in [0, \bar{c})$,*

$$\Pr(S_i = 1 \mid t_i = L) \;>\; \Pr(S_i = 1 \mid t_i = H). \tag{21}$$

Proof sketch. The decision rule in period 2 in the main text implies that individual i switches occupations if and only if the gain from switching, $G_i = \max_{k \in J} \hat{\theta}_{ik}^{(1)} - \hat{\theta}_{ij_i^{(1)}}^{(1)}$, exceeds the switching cost c. When the additional signals $s_{ij}^{(1)}$ are very precise (σ_1^2 small), the updated estimates $\hat{\theta}_{ij}^{(1)}$ are close to the true match qualities θ_{ij} for all j, regardless of t_i. In the limit as $\sigma_1^2 \to 0$, G_i converges to the true gain from moving from $j_i^{(1)}$ to the best occupation j_i^*, which is increasing in how poorly $j_i^{(1)}$ fits relative to j_i^*.

Under Assumption B.1, the true gain from switching is increasing in the gap between the skill distance in the initially chosen occupation and the minimal achievable distance. Thus, G_i is monotonically related to the underlying degree of skill mismatch captured by m_i. Individuals with larger m_i tend to have larger G_i and are therefore more likely to find it optimal to switch for any given cost c.

Lemma B.1 implies that the distribution of m_i (and hence of G_i) for low-information types places more mass on high values than for high-information types. Consequently, there exists $\bar{c} > 0$ such that, for all $c \in [0, \bar{c})$, the fraction of individuals with $G_i > c$ is strictly larger among low-information types than among high-information types, which yields Lemma B.2. □

Lemma B.2 shows that the second main prediction of the model – that individuals facing stronger information frictions before their first choice are more likely to revise their occupational choice once better information becomes available – does not depend on whether mismatch is measured as a binary indicator or by a continuous, skill-based index m_i. In both cases, the mechanism is the same: lower initial information precision leads to worse initial sorting, which in turn increases the expected gains from switching after the information update and therefore raises the probability of observed switching.

Lemma B.3 (Information frictions and reductions in skill mismatch among switchers). Let $m_i = |x_i - q_{j_i^{(1)}}|$ denote the initial skill-based mismatch index from Definition B.1 and define the final skill-based mismatch index as

$$m_i^{\text{final}} \;=\; \left| x_i - q_{j_i^{(2)}} \right|.$$

The reduction in skill mismatch due to switching is

$$\Delta m_i \;=\; m_i - m_i^{\text{final}} \;\geq\; 0,$$

with $\Delta m_i > 0$ only if $S_i = 1$. Suppose that Assumptions 1–2 from the main text and Assumption B.1 hold, that $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$, and that the switching cost $c \geq 0$ is not too large. Then there exists $\bar{c} > 0$ such that for all $c \in [0, \bar{c})$,

$$\mathbb{E}[\Delta m_i \mid S_i = 1, t_i = L] \;>\; \mathbb{E}[\Delta m_i \mid S_i = 1, t_i = H]. \tag{22}$$

Proof sketch. Under Assumption B.1, match quality in occupation j can be written as a strictly decreasing function of the skill distance, $\theta_{ij} = g_j\big(|x_i - q_j|\big)$, so that a reduction in the distance $|x_i - q_j|$ is equivalent to an increase in match quality θ_{ij}. In particular, for any two occupations j and k,

$$|x_i - q_k| < |x_i - q_j| \quad \Longleftrightarrow \quad \theta_{ik} > \theta_{ij}.$$

When the additional signals $s_{ij}^{(1)}$ are very precise (σ_1^2 small), Lemma 3 in the main text implies that, among switchers, the expected gain in match quality ($G_i \;=\; \max_{k \in J} \hat{\theta}_{ik}^{(1)} \;-\; \hat{\theta}_{ij_i^{(1)}}^{(1)}$) is larger for low-information types than for high-information types:

$$\mathbb{E}[G_i \mid S_i = 1, t_i = L] \;>\; \mathbb{E}[G_i \mid S_i = 1, t_i = H].$$

Because of the one-to-one, strictly monotone relationship between match quality and skill distance in Assumption B.1, larger gains in match quality correspond to larger expected reductions in skill distance. Formally, there exists a strictly increasing transformation $h(\cdot)$ such that, for switchers,

$$\Delta m_i = h(G_i),$$

up to negligible approximation error when σ_1^2 is small. Hence the stochastic dominance of G_i for type L over type H among switchers carries over to Δm_i, implying

$$\mathbb{E}[\Delta m_i \mid S_i = 1, t_i = L] \;>\; \mathbb{E}[\Delta m_i \mid S_i = 1, t_i = H]$$

for all c in a neighborhood of zero. This yields (22). $\square$

Lemma B.3 shows that the prediction of Lemma 3 in the main text – that individuals with stronger initial information frictions realize larger gains in match quality when they switch – extends naturally to a continuous, skill-based measure of mismatch. Among switchers, not only do low-information types improve their match quality more, but they also reduce the distance between their skills and job requirements by more than high-information types do, on average.

Lemma B.4 (Information frictions and residual skill mismatch). Define the final skill-based mismatch index as above ($m_i^{\text{final}} = \left|x_i - q_{j_i^{(2)}}\right|$). Suppose that Assumptions 1-2 from the main text and Assumption B.1 hold, that $\sigma_1^2 < \sigma_H^2 < \sigma_L^2$, and that the switching cost $c \geq 0$ is not too large. Then there exists $\bar{c} > 0$ such that for all $c \in (0, \bar{c})$,

$$\mathbb{E}\big[m_i^{\text{final}} \mid t_i = L\big] \;>\; \mathbb{E}\big[m_i^{\text{final}} \mid t_i = H\big]. \tag{23}$$

Proof sketch. By Assumption B.1, the occupation j_i^* that maximizes θ_{ij} minimizes the skill distance $|x_i - q_j|$. Hence the minimal achievable skill mismatch for individual i is $m_i^* = \left|x_i - q_{j_i^*}\right|$. Initial mismatch in the continuous sense is $m_i - m_i^*$, and final mismatch is $m_i^{\text{final}} - m_i^*$.

Lemma B.1 establishes that the expected initial skill mismatch $\mathbb{E}[m_i \mid t_i]$ is larger for low-information types than for high-information types. Lemma B.2 shows that low-information types are more likely to switch occupations once better information becomes available, and Lemma B.3 shows that, among switchers, low-information types tend to achieve larger reductions in skill mismatch than high-information types. However, switching costs $c > 0$ prevent full adjustment: some initially mismatched individuals do not switch, and even among switchers, the final occupation need not coincide with j_i^*.

Formally, write the final skill mismatch as

$$m_i^{\text{final}} = m_i - \Delta m_i,$$

so that

$$\mathbb{E}[m_i^{\text{final}} \mid t_i] = \mathbb{E}[m_i \mid t_i] - \mathbb{E}[\Delta m_i \mid t_i].$$

Lemma B.1 implies $\mathbb{E}[m_i \mid t_i = L] > \mathbb{E}[m_i \mid t_i = H]$, while Lemmas B.2 and B.3 imply that $\mathbb{E}[\Delta m_i \mid t_i = L]$ is larger than $\mathbb{E}[\Delta m_i \mid t_i = H]$, but not so large as to offset the initial disadvantage when c is small but positive. Under mild regularity conditions (continuous distributions of skills and requirements), the difference in initial mismatch carries through to the final allocation for all c in a neighborhood of zero, yielding Lemma B.4. $\square$

Intuitively, low-information types start further away from their optimal occupations in skill space, and although they, on average, reduce their mismatch more when they switch, switching costs prevent full convergence to the optimum. As a result, they remain more mismatched in expectation even after the information update and the second occupational choice. Lemma B.4 therefore complements Lemma B.2 by showing that the higher residual mismatch rate for low-information types in the main text translates into a higher expected skill-based mismatch index in the continuous setting.