Human Capital and Labor Market Shocks in the Modern Economy*

Valerie Smeets
Aarhus University & CEPR

Lin Tian
INSEAD & CEPR

Sharon Traiberman
NYU & NBER

February 1, 2025

Abstract

We study how university students' field of study choices influence the economy's skill distribution and its adaptability to labor market disruptions. Our framework incorporates forward-looking students select majors that determine their comparative advantage across occupations. Using Danish administrative data on university applications and career outcomes, we account for unobserved heterogeneity in student and worker skills. Our estimated model finds that fields vary significantly in specificity, with the most generalized fields being nearly six times more responsive to shocks than the most specialized. Students cluster into STEM and non-STEM groups but show greater elasticity within clusters than previously estimated. Counterfactuals reveal that easing capacity constraints could boost GDP by 0.3% and reallocate majors. Simulating labor market shocks, such as declining foreign goods prices and AI-driven demand shifts, shows that incorporating field choice doubles occupational reallocation, amplifying trade gains, while AI shocks have smaller effects.

JEL: F14, F16, J24, I24

*Contact information: vas@econ.au.dk, lin.tian@insead.edu, st1012@nyu.edu.

1 Introduction

Economic shocks—such as globalization, automation, and immigration—transform labor markets, creating winners and losers. These disruptions raise fundamental questions: Who benefits, who suffers, and how can policies build resilience? Addressing these issues requires understanding the heterogeneity in workers' skills, their transferability across occupations, and, most importantly, how these skills are developed. The foundation of workers' skills lies in decisions made as students (Altonji et al., 2012a; Kirkeboen et al., 2016), yet the role of educational choices in shaping labor market outcomes remains underexplored. A unified understanding of how these choices interact with labor market dynamics is crucial for answering these pressing questions and designing effective policy responses.

This paper develops a comprehensive framework to examine how students' field of study choices influence skill formation and labor market dynamics. The model captures how individual preferences and abilities interact with field-specific pathways to shape workers' comparative advantage and career trajectories. Importantly, it identifies two distinct margins of adjustment to shocks: the reallocation of incumbent workers across occupations and the adaptation of new student cohorts through changes in field selection. Through this framework and the rich Danish administrative data, we quantify the responsiveness of different fields of study to shocks. The counterfactual analyses uncover the implications of capacity constraints in higher education, and evaluate the broader implications of labor market shocks under different policy scenarios.

The context of our analysis is the Danish tertiary education system, which we discuss in detail in Section 2. We use administrative data from Denmark, where we observe a long panel of workers' career histories between 1998 and 2018. This is merged to a panel of tertiary education application data. In addition to the coverage and length of the data, the Danish education system has two attractive features ideal for our study. First, the data allows us to follow every student through multiple layers of the education system, and into their subsequent career. Second, similar to many countries[1], Denmark employs a centralized admission process to allocate students to programs. Specifically, students who opt to enter the college system apply for admission by submitting a list of priorities over institutions and programs. The education ministry then allocates students according to a deferred acceptance algorithm, in which students are ranked by their application score and are assigned to places based on their preferred choices. We use the priority

[1] For example, Norway (Kirkeboen et al., 2016), Turkey (Balinski and Sonmez, 1999), and Chile (Hastings et al., 2013).

lists of students as a measure of their ranking over fields, which will allow us to recover students' unobserved preferences when estimating our model.

In Section 3 we document several facts that motivate the models' ingredients. First, we show that different fields of study prepare workers for a fairly specialized set of occupations. This motivates our focus on field of study, as it immediately suggests that this is likely an important margin of heterogeneity for understanding the incidence of labor market shocks. Second, we use the fixed effects decomposition proposed by Porzio et al. (2022) to decompose reallocation across occupations into the portion attributable to new cohorts and to incumbents switching jobs. We find that both are important margins, but which is more important differs across occupations. This fact motivates us to jointly model the labor supply decision of incumbents, and the education decision of new cohorts. Finally, to motivate how we model field choice, we use a Kitagawa-Blinder-Oaxaca decomposition to demonstrate that a significant fraction of occupational reallocation by new cohorts is driven by students' changing their field of study, rather than students of the same field of study changing their occupations.

Motivated by these facts, in Section 4 we develop a model of labor supply and educational choices. Our model can be thought of as comprising two blocks: the career decisions of incumbent workers, and the education decisions of students. In the career choice block of our model, workers choose their occupation in each period à la Artuç et al. (2010). Workers make different career choices for two reasons: switching requires paying a partially-stochastic fixed cost, and workers differ in their comparative advantages across occupations. Specifically, at the beginning of their careers, workers draw a vector of productivities across different occupations. This vector is drawn from a distribution that depends on workers' innate talents, measured by test scores, their field of study, and other characteristics. Abilities differ within and across educational backgrounds. For example, social science programs may generally provide better quantitative skills but weaker communication skills compared to humanities programs. However, within social sciences, some programs may have lower quantitative skills but stronger communication skills than others.

In the field choice block of our model, students choose which field of study to pursue. This block of the model builds on the large literature that models major choice (Kinsler and Pavan, 2015; Arcidiacono et al., 2020). Students do not know what their labor market type will be, but understand that it is mediated by their education choice. We allow for students to be heterogeneous in three dimensions: they can vary in their non-pecuniary preferences across fields, they can vary in their preferences for income relative to non-pecuniary factors, and the value of a field of study

can vary across students' observable characteristics, such as test scores.

In Section 5, we discuss identification and estimation. To estimate the labor supply block of the model we use a two-stage estimator, following the approach of Traiberman (2019), building on Arcidiacono and Miller (2011). This approach requires first using the Expectation-Maximization (EM) algorithm to recover a subset of parameters, including the distribution of workers' latent comparative advantage, and transition rates across occupations. In the second stage, we regress transition rates on wage differentials to recover the remaining labor supply parameters, as in Artuç et al. (2010). We use students' ranked lists to then estimate the field choice block of our model. We build on the literature that exploits the deferred acceptance algorithm (Luflade, 2019; Fack et al., 2019) to estimate preferences. Having information on students' *second choices* allows us to model latent preferences across fields of study. We use time variation in the model-implied returns to fields of study in order to estimate how responsive students are to shocks to future income.

In Section 6 we present estimates from the labor market and field choice blocks of our model. We construct a measure of fields' comparative advantage at different occupations. We find that some fields are highly specialized, matching strongly to a small handful of occupations. Other fields are much more general. To measure the economic content of specificity, we calculate the increase in the exit rate from an occupation given a 1% decline in earnings in that occupation. We have two main results. First, controlling for workers' human capital is crucial to estimating the responsiveness of workers to shocks: a 1% decline in the NPV of an occupation raises the exit rate by 11%, however this effect is four times smaller when we ignore worker and field heterogeneity. Second, we find substantial heterogeneity in workers' responsiveness by field. The most generalized fields are nearly six times more responsive to shocks than the most specialized.

Turning to students decision over fields, we focus on three key findings. First, we find that the average student is relatively responsive to shocks to future income. For example, a 1% increase in the net present value of a field increases entry into that field by .4 percentage points (an elasticity of 2.6). Second, we find an important role for heterogeneity across students—both in terms of responsiveness and preferences. For example, we find that students who wish to study education have a strong preference for that field, but are relatively elastic and willing to switch into care or health fields. On the other hand, we find that STEM-types are relatively inelastic, and mostly switch within STEM fields or into business. Finally, we show that modeling heterogeneity is crucial to capturing realistic substitution patterns between students. We show that the implied substitution matrix for students—which fields students switch into given a shock to a single

field—has similarities to the observed preferences we observe from students' ranked lists. On the other hand, estimating our model without this heterogeneity—akin to estimating a simple logit model featuring independence of irrelevant alternatives—leads to wildly unrealistic patterns.

In Section 7, we use our estimates to perform several counterfactual experiments. To do so, we first close our model by embedding the labor supply side into a small open economy model of labor demand. Firms in different industries face import competition and can choose to use offshored intermediates in production. We then simulate demand shocks across occupations, by counterfactually reducing the costs of foreign goods and offshored inputs. The trade shocks would in turn induce heterogeneous responses across occupations. We adopt the exact hat algebra approach to estimate the labor demand parameters and solve the counterfactuals under different policy scenarios. In particular, we shut down reallocation for all fields, we maintain the current Danish system (which has capacity constraints for some fields), and we allow students to freely choose fields and assume supply can adjust (the "American" system). We show that allowing students to reallocate under the current Danish system, with capacity constraints, leads to 12% higher welfare gains from offshoring. If institutional constraints (e.g., capacity constraints and high school tracks) are removed, welfare gains following the shock are 22% higher.

Our paper contributes to several strands of literature. First, we contribute to the literature that uses general equilibrium models to study the dynamic response of workers to shocks.

Our paper is most closely related to papers that try to understand the career decisions of workers who differ in their abilities, either due to mobility frictions, or due to horizontally differentiated skill sets. These models have been used to study trade shocks (Artuç et al., 2010; Dix-Carneiro, 2014; Traiberman, 2019), immigration (Llull, 2018; Burstein et al., 2020; Monras, 2020; Khanna and Morales, 2017), and to understand search and sorting patterns in the economy (Lise and Postel-Vinay, 2020).

We also contribute to the large literature on the the joint determination of human capital and career decisions. Numerous papers have documented variation in the returns to different fields of study (Kirkeboen et al., 2016; Bleemer and Mehta, 2022; Andrews et al., 2022), and also explored the importance of different features of majors, such as specificity (Leighton and Speer, 2020).[2] Within this literature, our paper is most closely related to the subset of the literature focused on estimating models of human capital accumulation in tertiary education—for example,

[2] Altonji et al. (2016) and Altonji et al. (2012b) are excellent surveys, containing many more references, as well as discussions of structural modeling issues.

Arcidiacono (2004), Adda and Dustmann (2023), Beffy et al. (2012), and Kinsler and Pavan (2015), who estimate both a labor market process alongside a human capital decision. We also build on the seminal work of Keane and Wolpin (1997), who model labor demand in addition to a college decision. Methodologically, we are close to the literature exploiting centralized assignment mechanisms as a source of identification for preferences over college majors (Kapor et al., 2024; Fack et al., 2019; Larroucau and Rios, 2022). Our contributions to these literature are several-fold. First, much of the prior literature on education and careers has focused on vertical education decisions—e.g., no college, vocational school, or college—while we focus on horizontally differentiated fields of study—Beffy et al. (2012) and Arcidiacono et al. (2016) being notable exceptions. Second, we tie students' education decisions to a model of the labor market that features many occupations and sectors, and allows for dynamic adjustment along these margins. This is novel to our approach, and we show that modeling such occupational heterogeneity is important for two reasons. First, we show that both incumbent and new cohort switching are key margins of adjustment in the economy. They also interact—as the reallocation of incumbents changes expected earnings for new cohorts, while the actions of new cohorts similarly affects the anticipated change in labor supply for incumbents. Second, a common feature of trade shocks, technological shocks, and immigration shocks are that they impact only a subset of occupations directly, and the remainder through equilibrium reallocation. For example, certain "STEM"-intensive occupations may be exposed to trade or artificial intelligence—e.g., those in R&D—while others less so—e.g., medicine. Our paper provides a unified framework that bridges a rich model of field choice, with a dynamic occupational choice model that can be used to understand reallocation to the kinds of targeted shocks that worry policy makers.

Finally, our paper contributes to the growing literature on the demand for fields of study (Robles and Krishna, 2012; Wiswall and Zafar, 2015; Kirkeboen et al., 2016; Agarwal and Somaini, 2018; Fack et al., 2019), by nesting field choice within a broader labor supply framework. Unlike much of the existing literature, our approach integrates educational decisions with labor market reallocation, which allows us to study how field choices interact with occupational reallocation in response to shocks, providing insights into aggregate outcomes and general equilibrium effects.

2 Data and Institutional Background

In this section, we describe the data sources and definitions used in the empirical analysis. We then provide a brief description of the Danish education system and show prominent features in students' applications and field choices.

2.1 Data Sources

We employ several administrative datasets covering the Danish economy from 1996 to 2019. The first is the *Integrated Database for Labour Market Research* (IDA)—a panel dataset with information on the universe of Danish workers over 15, both native and non-native. IDA includes information on employment status, earnings, ISCO-based occupation, age, sex at birth, and other demographic variables. Throughout the text "earnings" refers to the earnings concept in this register—which are total annual earnings in workers' main job. Workers are assigned an employer and occupation based on their employment status in November of each year.

Our main source of information regarding individuals' education comes from the *Education Register* (UDDA). The data include the type of education completed, the level of education and the year of graduation . The education register also contains references from all education programs to the International Standard Classification of Education (ISCED) which allows us to classify education programs into field of study. In the next subsection we discuss our classification of occupations and fields of study. We supplement the education register with information regarding student applications from the *Coordination Register* (KOT). The data are provided by the Coordinated Admissions System, which is run by the Ministry of Higher Education. For each student, we know the set of programs to which they applied, whether they received admission in a program, the system under which they received admission (i.e., through GPA, or an alternative channel), their test score, among other variables. We hand-link the education programs in KOT to UDDA and calculate admission cutoffs based on all applications data[3].

In addition to these datasets, we bring in standard firm level data from the *Firm Accounting Statistics Register* (FIRE). This allows us to identify workers' industries. Finally, we use the *Danish Foreign Trade Statistics Register* (UHDI). These data are product-firm-origin/destination level customs data on all international transactions.[4] They contain data on value of imports and exports at

[3]We also need to determine the effective GPA cutoff threshold for each program with a binding testing cutoff, as this is not available (digitally at least). The Data Appendix contains a complete description of how we concorded KOT to UDDA and calculated admission cutoffs.

[4]For EU countries, there is a minimum reporting threshold.

the CN8 level, and often also contain quantity and unit value information.

In addition to the Danish register data, we use the O*Net Database, which is maintained by the National Center for O*NET Development under sponsorship of the US Department of Labor. These data, which can be merged to the Danish occupation codes, are aggregated from survey responses pertaining to the importance of various tasks, skills, and types of knowledge to different occupations.

2.2 Defining Fields and Occupations

We define the level of education based on the ISCED system, and divide students into three levels: no tertiary education, short-cycle education, and medium-cycle or higher education. We refer to last group as level 3 education, and this is the primary focus of our analysis. Our definition of fields is based on the Danish variant of the ISCED-F classification of fields. The ISCED-F classification breaks fields into 27 broad categories. We aggregate several codes—especially in the natural sciences—due to their small size. We also break apart the health code into programs at the Professional Bachelor's level and the University level. Appendix table 10 contains the full list of educational fields.[5]

Occupations in Denmark are recorded according to the ISCO classification. There is a break in the recording of occupational codes beginning in 2009, as the ILO updated the codes. We use the concordance of Humlum (2021) to create a time-consistent series of occupations. For data quality issues, we focus on 2 digit occupational codes. After removing some small occupations, this yields 22 occupations. When the data breaks, there are several quality issues in the data. The Data Appendix contains a complete description of how we concord and aggregate occupations over time, how we deal with variation in the quality of occupational codes (e.g., some occupations are recorded according to an employee register, while others are imputed), and how we partially impute missing codes. Appendix Table 11 contains the list of occupations in our analysis.

2.3 The Danish Education System

The Higher Education System in Denmark

Our analysis focuses on tertiary or higher education in Denmark, which includes two main types of programs: Professional Bachelor's degrees and University Bachelor's degrees. Professional

[5]Note that not all of these fields have corresponding programs at the highest level of tertiary education in Denmark, but we include them as in counterfactuals we include *all* workers, who may have studied these fields at other schools.

Bachelor's degrees, often referred to as Academy Bachelor's degrees, typically require three years of study and are designed to provide students with practical, job-oriented skills tailored to meet labor market demands. These programs cover fields such as business, education, engineering, and various forms of vocational training, emphasizing technical expertise and immediate employability. University Bachelor's degrees, on the other hand, follow a more academically focused path, comparable to a standard American four-year degree but completed in three years . These programs are oriented toward theoretical knowledge and research, preparing students for further academic pursuits or careers requiring in-depth expertise in their chosen fields.

In Denmark, as in much of Europe, the higher education system emphasizes early specialization, differing significantly from the U.S. model of majors and minors. Students apply directly to specific programs, such as engineering or business, rather than entering with a broad area of study. Unlike the U.S. model, which offers general education requirements and flexibility in course selection, the curriculum in Denmark is more predefined, with less room for electives outside the chosen field.

High school tracks play a critical role in shaping students' options for higher education. Danish upper secondary education comprises several tracks, each influencing the types of tertiary programs students can pursue (see Appendix XXX for more information in high school tracks). While the system has become more fluid over time, with for example options for students to take supplementary courses to broaden their eligibility, these tracks still determine much of the access to higher education. To this end, in our estimation, we use upper secondary background, including specialization when available, to construct students' consideration sets for tertiary education. We also use high school tracks as a control when estimating students' returns and preferences over different fields of study.

The Application System

Entry into both Professional Bachelor's programs and University Bachelor's programs occurs through a centralized admissions process. Students can apply to up to eight programs (a combination of field and institution). Each program has its own GPA threshold for entry, which is determined by student demand and the program supply, regulated by the Danish Ministry of Higher Education and Science. Some programs also have specific prerequisites, such as sufficient mathematics preparation. Students in Denmark are given significant resources to help them make informed choices about their education and improve their information about alternatives

educational choice [6]. The UddannelsesGuiden (Education Guide[7]) provides detailed information on program-specific prerequisites, historical GPA cutoffs, admissions numbers, labor market outcomes, demand trends across fields of study, and subject matter expectations. Study guidance centers (Studievalg) further support students through counseling services. Nevertheless, while students inevitably face some gaps in information and cannot fully predict the returns to different fields of study, significant efforts are made to provide them with comprehensive guidance and resources.

The Danish admissions system uses a Deferred Acceptance mechanism, ensuring that if a student is ineligible for their first choice, it does not affect their chances of being considered for subsequent choices. This structure encourages students to truthfully rank their preferences (Fack et al., 2019). However, some students may still choose not to apply to programs where they believe they are ineligible or incur psychological costs from the process. In our estimation, we assume students apply within a "realistic" set of alternatives, consistent with the "score bounds" approach of Kapor et al., 2024. While the majority of admissions to tertiary education programs take place through GPA cutoffs (also referred to as the "Quota 1" system), a fraction of admissions take place based on supplementary exams or other criteria for students who were rejected from their Quota 1 application (the "Quota 2" system). In our analysis, we only focus on Quota 1 admissions. It is also not uncommon for students to wait between high school and college if they do not receive their top choice. Although the system allows for some fluidity, quotas remain binding for many programs, and delaying applications can result in penalties. As a result, most students apply to multiple programs and many are admitted to or attend their second (or higher) choice.

Figure 1 displays the unconditional distribution of first priorities for students across fields of study, as well as the growth in fields over time. The most popular fields are in Care and Medium Health, followed closely by Business, Social Sciences and Education. The least popular fields are those in the STEM fields, as well as Arts. In terms of shares, the education field is declining in popularity while the medium health is gaining. ICT, which begins on a very small base, has seen explosive growth, as has Business. The Humanities have shrunk in popularity. However, the distribution of applicants also changes over time. The number of students applying to tertiary education through the Coordinated Admissions system has grown 61% between 1996 and 2017, driven both by higher take up of higher education, and an expansion of those programs covered

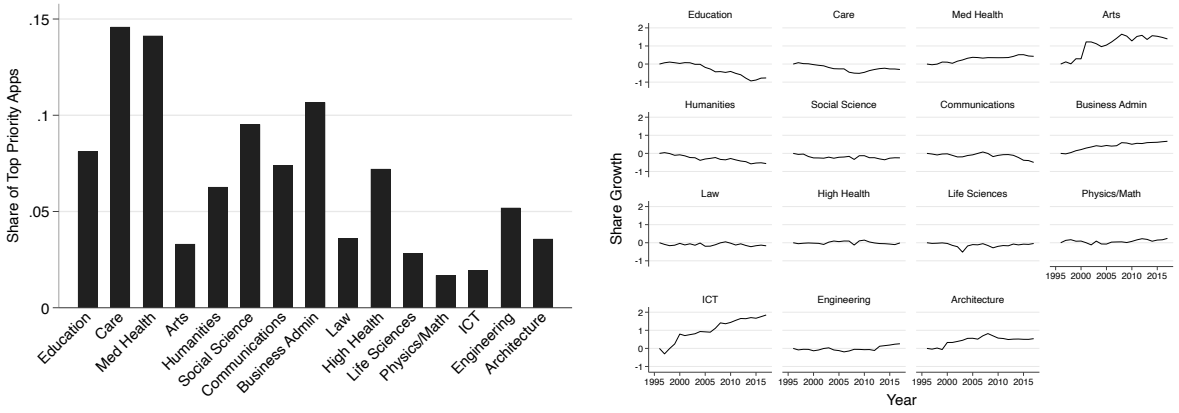
[6] This is in line with the fact that students may have poor information about alternatives, see (Wiswall and Zafar, 2015; Kapor et al., 2020).

[7] See https://www.ug.dk/

Figure 1: Summary Statistics on First Priority

(a) Unconditional Distribution

(b) Growth Over Time



by KOT.

Table 1 provides insight into students' preferences by examining the relationship between first-choice fields (columns) and second-choice fields (rows). Two patterns stand out. First, the data show strong diagonal dominance across fields, indicating that most students select a second-choice program within the same field as their first choice[8]. Second, preferences vary significantly by discipline. In STEM fields such as life sciences, physics/mathematics, and engineering, there is notable clustering, with a large share of students selecting second or third choices within these closely related fields. In contrast, preferences in Education and Health appear more narrowly focused, with students showing a lower likelihood of considering programs outside their initial field of interest. This suggests a strong tendency toward specialization in these disciplines.

[8]In our subsequent analysis, we do not distinguish between students choosing a similar program within the same institution (e.g., Physics versus Math at Aarhus) or applying to the same program at different institutions. We do not find a clear pattern that suggests students engage in one kind of behavior or another: 37% of the time students apply across programs within the same institution, and the other 63% of the time they apply for similar programs across institutions.

	First Priority														
Second Priority	Education	Care	Med Health	Arts	Humanities	Social Science	Communications	Business Admin	Law	High Health	Life Sciences	Physics/Math	ICT	Engineering	Architecture
Education	75.92			4.54	4.19										
Care	8.38	88.43	5.03												
Med Health	3.51	4.02	85.57							7.56	4.80				
Arts				46.98											7.54
Humanities				13.27	51.48	13.89	16.81		5.72						
Social Science				9.40	15.40	49.09	11.60	9.46	16.07	4.10		7.62	4.69		
Communications				9.98	15.48	7.46	42.61	10.59	5.84				3.79		
Business Admin						9.32	14.11	66.77	22.30			4.62	6.98	4.13	
Law						4.74			39.98						
High Health										67.36	11.37	3.57			
Life Sciences										9.25	50.37	15.56	4.60	7.53	
Physics/Math											10.61	39.84	4.85	6.39	
ICT												4.17	46.64	9.18	
Engineering											11.37	13.35	19.15	53.37	14.73
Architecture				5.85										10.11	55.75

Table 1: $P(Priority_{r+1}|Priority_r)$

Notes: The columns display students' first-choice fields and the rows display students' second-choice fields. The definition of fields is based on the Danish variant of the ISCED classification fields and is discussed in Section 2.2. Cells are censored below 3.5% for readability.

3 Motivating Model Ingredients

In this section, we present three stylized facts about the relationship between fields of study and career choices. Our analysis focuses on students and workers with tertiary education. The purpose of this section is to motivate our model, which incorporates both a career choice component and a field choice component. The facts presented below highlight that occupational reallocation is influenced by two margins: how incumbents reallocate themselves and how new students adjust their human capital investment decisions.

Fact 1: Different fields of study are associated with specialization in different occupations.

We first demonstrate that fields of study prepare students for a narrow band of occupations, but that there is substantial heterogeneity in specificity. To measure how specialized a field is, we calculate the Herfindahl-Hirschman Index (HHI) across occupations for each field:

$$HHI_m = \sum_{o \in \mathcal{O}} s_{o|m}^2,$$

where $s_{o|m}$ is the share of workers from education field m working in occupation o. We calculate shares pooling across the whole economy. Figure 2 contains a bar graph for each field of study, as well as the HHI calculated by pooling all workers (in grey).

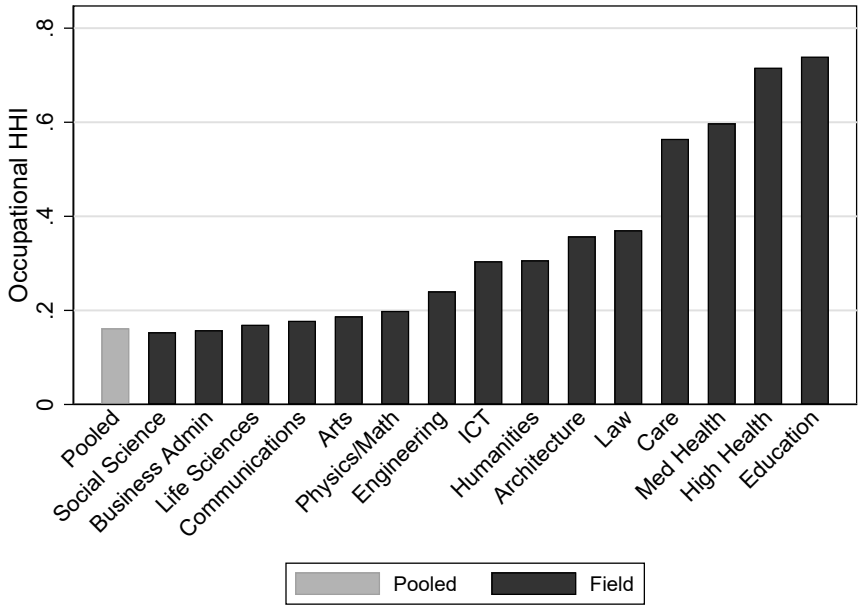
We highlight two takeaways from this graph. First, fields of study are generally highly specialized. The pooled HHI—unconditional on field of study—is about .2, while the average HHI for any given field is nearly double this. Second, there is substantial heterogeneity across fields, which suggests that there could be significant heterogeneity in how workers with different backgrounds weather shocks. For example, Social Science has a lower HHI than the pooled HHI, while health and education fields are nearly fully specialized.

Fact 2: Occupational reallocation is driven equally by new cohorts and incumbents.

We now establish that a substantial fraction of occupational reallocation occurs through the entry of new cohorts. Changes in occupational shares can occur in two ways: incumbents can switch occupations (or exit the labor force), or new cohorts can engage in a different set of occupations.

In order to quantify these channels we adopt the approach of Porzio et al. (2022), who studied

Figure 2: Field of Study-Occupation HHIs



Notes: The black bars represent the HHI for different fields of study. The definition of fields is based on the Danish variant of the ISCED classification fields and is discussed in Section 2.2. The grey bar represents the HHI calculated by pooling all workers.

reallocation of workers in and out of agriculture in emerging economies. Let $s_{o|c,t}$ be the share of workers in cohort c working in occupation o at time t. We project this share on cohort and time fixed effects for each occupation:

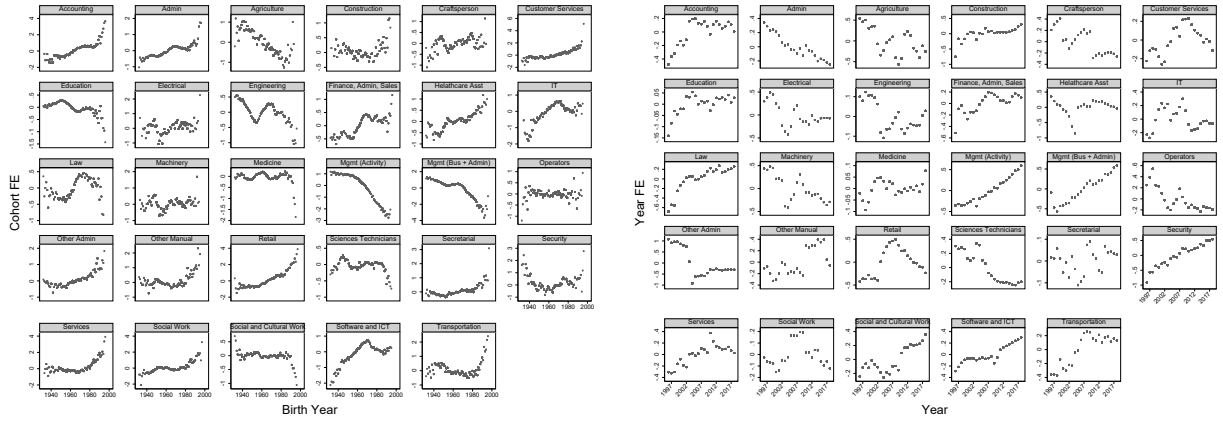
$$logs_{o|c,t} = \varphi_{oc} + \psi_{ot} + \varepsilon_{oct}. \tag{1}$$

To interpret this equation, consider the case that cohort effects are irrelevant: workers are freely mobile (or at least equally mobile) across generations. In this case, the conditional share of workers in each occupation would be similar, so that $s_{o|c,t} \approx s_{o|c',t} \forall c, c'$. In this case, reallocation would reflect common demand or supply shocks, and be absorbed into the year fixed effects. On the other hand, if workers are completely immobile, then all reallocation would occur through new cohorts, and only those fixed effects would matter.

Figure 3 plots both cohort and year fixed effects for each occupation. The first panel plots cohort effects, with birth year of the cohort on the x-axis, and the fixed effect on the y-axis. Both sets of fixed effects clearly matter. However, there are some clear differences, where for some occupations the cohort effect is close to irrelevant, while for others the time effect is small. This motivates a model that captures both margins of adjustment for the economy.

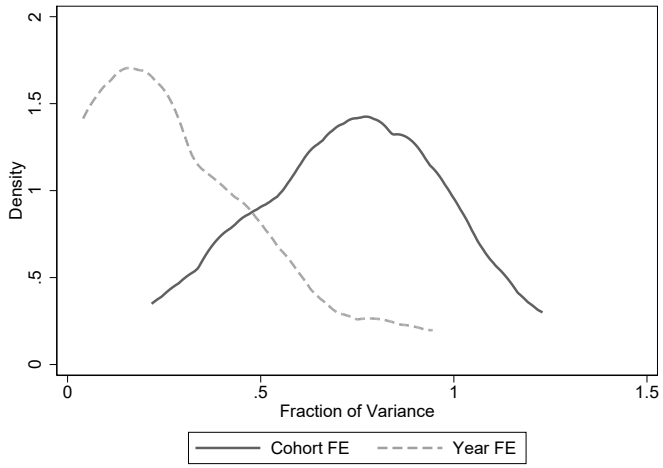
Figure 3: Reallocation Decomposition

(a) Cohort Fixed Effects (b) Year Fixed Effects



Notes: The figure plots the cohort and year fixed effects from (1). The null hypothesis is a flat line bunched at zero.

Figure 4: Variance Explained by Cohort and Year Effects



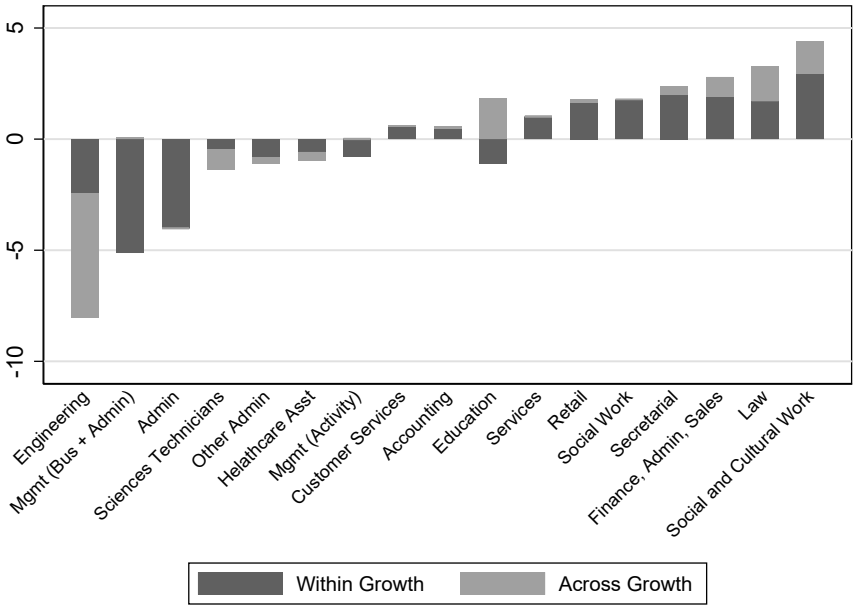
To further quantify the importance and heterogeneity of the cohort effect, we calculate the ratio of variance in fixed effects to the variance of the fitted log shares: $\sigma^2_{\varphi_{oj}} / \sigma^2_{\widehat{\log s_{o|c,t}}}, j \in \{c,t\}$. This number can be larger than one due to the covariance of cohort and year fixed effects. Figure 4 plots the density of this ratio for both year and cohort fixed effects. The cohort effects are to the right of the year fixed effects, suggesting that, if anything, they account for a larger share of variance—with the caveat that the covariance term (which tends to be negligible) is absent here. We finally turn to establishing that new cohorts reallocate by switching their field of study.

Fact 3: New cohorts reallocate through their choice of field of study.

Finally, in our third fact, we document the extent to which reallocation across cohorts occurs through changes in the composition of fields of study. To do this, we do a Kitagawa-Oaxaca-Blinder decomposition of changes in initial occupations for new cohorts into within- and across-field of study. Specifically we consider the following decomposition:

$$\Delta s_{ot} = \underbrace{\sum_m \Delta s_{o|m} \times s_{m,t-1}}_{\text{Within}} + \underbrace{\sum_m \Delta s_m \times s_{o|m,t-1}}_{\text{Across}} + \underbrace{\sum_m \Delta s_m \times \Delta s_{o|m}}_{\text{Growth Covariance}}. \tag{2}$$

Figure 5: Decomposing the New Cohort Entry Margin



Notes: The decomposition of occupational reallocation follows (2). Occupations are sorted based on total growth over 1998 to 2018.

Figure 5 plots this decomposition for all occupations that experienced at least a 5% change in share.[9] The light gray bars measures changes through field choice. For nearly all occupations, a substantial fraction of changes in initial occupation shares occurs through changes in field choices. There are some notable deviations—for example, management and administration. These occupations may be less matched to a particular field.

Together, our facts paint a clear picture of how the occupational structure of the economy

[9]The covariance term tends to be very small, so it is omitted. The full decomposition for all occupations can be found in appendix figure 9.

changes: incumbents and new cohorts both play a large role in shaping occupation reallocation. Moreover, new cohorts shift their initial occupation in response to shocks, but also shift their field of study over time. With this in mind, we now turn to our model of labor supply and field choice, and its estimation.

4 A Model of Labor Supply and Field Choice

In this section we describe students' education and labor supply decisions, with labor demand left to Section 7. Our labor supply model combines three main ingredients: first, students make forward-looking choices over which field to pursue; second, workers with different educational backgrounds make forward-looking occupational choices; third, we incorporate latent heterogeneity in preferences over fields, as well as latent comparative advantage across occupations.

4.1 Environment

We consider a dynamic, small, and open economy. Time, indexed by t, is discrete. There is a measure L of workers, differentiated by their field of study $m \in \mathcal{F}$ and comparative advantage type $k \in \{1, ..., K\}$. Workers supply one unit of labor inelasticially to an occupation $o \in \mathcal{O}$. They exit the labor force at an exogenous rate δ, and are replaced by a new cohort. New cohorts are differentiated by their preferences, l, their test score, s, and other covariates, X.

Figure 6 provides a schematic illustration of the labor supply model. High school graduates start their lives in the *field choice block*, in which they pick their field of study based on anticipated returns to different fields and a preference shock—returns and preferences can depend on their initial state. Students graduate into the *career choice block* with their chosen field of study and comparative advantages across different occupations. New graduates make a forward-looking decision on their initial occupation. In each period afterwards, incumbent workers can choose to switch occupations, subject to switching costs. In the sequel, we describe the model backwards, starting from the incumbents, moving onto the new cohorts, and finally the students.

4.1.1 Incumbent Workers

Incumbent workers who spent the last period in occupation o choose an occupation o' at the *beginning* of period t. In order to switch occupations they face moving costs $C_{oo'km}$ and receive a stochastic occupation preference shock, ε_o. Workers of type (k, m) supply human capital $H_o(k, m)$

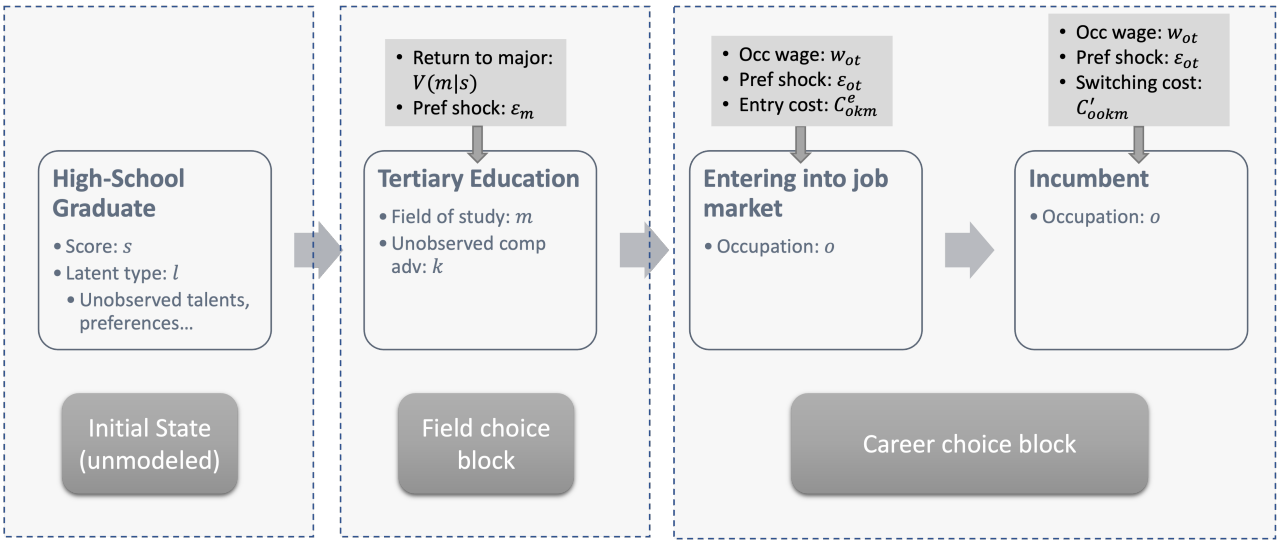


Figure 6: Illustration of the Labor Supply Model

to occupation o. In order to parameterize H we first project each occupation onto Z tasks. Denote by $\zeta_{z,o}$ the loading on task z in occupation o, where a larger value of ζ means that the task is more important for output in that occupation. With this notation, each occupation is attached to a vector in the positive orthant of $\mathbb{R}^Z$, $\zeta_o \in \mathbb{R}^Z_{++}$. Each type of worker has a different ability at each task, $\alpha_{k,z}$. Worker's log output in an occupation is the dot product of α and ζ. We additionally allow for a field effect, α_{om}, that scales up and down the human capital supply of workers in each occupation, and a type effect α_k that reflects the absolute advantage of type-k workers within a field. Putting this together we have that,

$$\log H_o(k,m) = \alpha_{om} + \alpha_k + \sum_{z=1}^{Z} \alpha_{k,z}\zeta_{z,o}. \tag{3}$$

A few comments are in order. First, workers have comparative advantage across different occupations through differences in their type k. As an example, suppose there were two occupations, two tasks, and two types. Suppose that for the first occupation, $\zeta = (1,0)$, and for the other occupation, $\zeta = (0,1)$. Similarly, suppose that $\alpha_1 = (1,0)$ and $\alpha_2 = (0,1)$. Then, type 1 workers would only produce output in occupation 1, while type 2 workers would only produce output in occupation 2. Nevertheless, given this horizontal differentiation there is no clear way to argue that one type has any absolute advantage. Second, through the α_{om} terms, workers in different

fields of study can have different comparative advantages, even conditional on type k. For example, workers in STEM backgrounds may have large values of α_{om} in technical occupations, such as engineering. Later, we discuss how during the education phase, the conditional distribution of types, k, is a function of m.

Returns to an occupation are determined competitively, with skill price w_{ot}. Putting this together, the worker's Bellman equation in period t is given by,

$$v(k, m, o, \varepsilon) = \max_{o'} w_{o't} H_{o'}(k, m) - C_{oo'km} - \varepsilon_{o't} + \beta E_t V(k, m, o'), \tag{4}$$

where $V \equiv \int v(\cdot, \varepsilon) dG(\varepsilon)$, is the value function v integrated over shocks, and β is the discount rate, $\tilde{\beta}$, multiplied by the probability of exit, δ.

Henceforth we assume that preference shocks, ε_o, are conditionally independent of other variables, and are distributed Gumbel with mean 0 and scale parameter ν. Under these assumptions, V takes the following form:

$$V_t(k, m, o) = \nu \log \left(\sum_{o'} \exp \left(\frac{-C_{oo'km} + w_{o't} H_{o'}(k, m) + \beta E_t V_{t+1}(k, m, o')}{\nu} \right) \right). \tag{5}$$

Moreover, the policy function for workers aggregates to the following transition probabilities across occupations and states:

$$\lambda_t(o' | o, k, m) = \frac{\exp \left(\frac{-C_{oo'km} + w_{o't} H_{o'}(k,m) + \beta E_t V_{t+1}(k,m,o')}{\nu} \right)}{\sum_{o''} \exp \left(\frac{-C_{oo''km} + w_{o''t} H_{o''}(k,m) + \beta E_t V_{t+1}(k,m,o'')}{\nu} \right)}. \tag{6}$$

Equation (6) illustrates the interpretation of ν as governing the elasticity of occupational switching with respect to fundamentals. Suppose that $\nu \to 0$. In this case, the transition probabilities collapse to either 0 or 1, taking on the value of 1 only if the numerator is the highest utility occupation. Hence, for low values of ν, earnings differentials determine worker transitions, and gross worker flows are close to net worker flows. At the other extreme, if $\nu \to \infty$, then (6) collapses to $1/|\mathcal{O}|$—i.e., workers randomly allocate themselves uniformly across occupations in each period. And so, for large values of ν, earnings differentials are irrelevant.

4.1.2 New Entrants

Like incumbents, new entrants enter with a (k, m) pair. They make a multinomial choice over initial occupation, according to the utility function,

$$v_t^e(k,m) = \max_o -C_{o,km}^e - \epsilon_o + w_{ot} H_o(k,m) + \beta E_t V_{t+1}(k,m,o) \tag{7}$$

We assume ϵ is Gumbel with the same scale parameter as for incumbent workers. This yields the following choice probability for new cohorts:

$$\lambda_t^e(o|k,m) = \frac{\exp\left(\frac{-C_{o,km}^e + w_{ot} H_o(k,m) + \beta E_t V_{t+1}(k,m,o)}{\nu}\right)}{\sum_{o'} \exp\left(\frac{-C_{o',km}^e + w_{o't} H_{o'}(k,m) + \beta E_t V_{t+1}(k,m,o')}{\nu}\right)}. \tag{8}$$

Moreover, one can integrate to solve for the value of being type (k, m) at a time t:

$$V_t^e(k,m) = \nu \log \sum_{o'} \exp\left(\frac{-C_{o,km}^e + w_{o't} H_{o'}(k,m) + \beta E_t V_{t+1}(k,m,o')}{\nu}\right).$$

4.2 Field Choice

In this subsection we first describe the returns to a field, and then discuss how students construct their ranked choice lists. Students enter their field choice decision with a test score, s, and a preference type, $l \in 1, ..., L$. They also have covariates, X, such as their high school background. In this subsection, we suppress these characteristics for ease of discussion, but everything is conditional on these characteristics.

Returns to Field

When students choose their field of study, they consider two effects on labor market returns. First, there is the common productivity effect, α_{om} in equation (3). For example, all economics degrees have some mean return in managerial occupations that may be different than the return to communications degrees. Second, each field of study induces a conditional distribution over $k \in K$, that may also depend on other characteristics: $\pi(k|m,s)$. This stochastic term allows for students within the same field of study to have different talents on the market. Given this conditional

distribution, the value of field m for a student with score s is:

$$V_t(m|s) = \sum_k V^e_{t+5}(k,m)\pi(k|m,s),$$

where $V^e_{t+5}(k,m)$ is the lifetime payoff to the student upon graduation with a labor market type k and a field m.

In addition to their test scores, students may have different preferences over fields of study. Student preferences are drawn from the set $l \in L$. We suppose that for a type-l student, there is a non-pecuniary, time-invariant, utility to a field denoted by ϑ_{lm}. Moreover, students can differ in the importance they assign to future earnings, which is governed by the parameter θ_l. Regardless of their type, students also receive a conditionally independent Gumbel(0,1) preference shock over fields, ε^m_m.

Finally, different fields of study can be associated to a varying number of programs. For example, within "Math/Physics" are all mathematics programs and physics programs, at all institutions. As the focus of our paper is not on institutional quality, nor on within-field variation, we simply treat programs within a field of study as symmetric. Instead, we assume that students have a love-of-variety over the number of programs, governed by v.

Combining all of these terms, student i of type l with test score s has the following utility for each field:

$$U_{imsl} = \vartheta_{lm} + \theta_l V_t(m|s) + v\log(N_m) + \varepsilon^m_{im}, \tag{9}$$

where ε^m is the field choice shock.

Ranked List Determination

Upon receiving their shocks, students fill out a ranked list of fields of length R. Students are assigned to a field via the deferred acceptance algorithm. We assume that students reveal their preferences over fields truthfully *over their consideration sets*.[10] Students' consideration sets can depend on their test score (for example, students with low scores will not apply to programs which they have no probability being admitted to) and their high school background (for example, whether they studied math or not). We denote consideration sets by $\mathcal{C} \subset \mathcal{F}$.

For ease of notation, let $\Theta_{lmt}(s)$ be the value of field m to type l at time t as a function of test

[10]Given that the list is truncated list, students may be incentivized to include a "safety" choice. At our level of aggregation, few students are ever actually bound on the list, so we ignore this constraint.

score s. I.e., the terms in (9) except for the shock. Let m_r be how an individual ranks their fields, so that m_1 means ranked first, and so on. Then, under the Gumbel assumption, one can show that the probability of observing a list $(m_1, m_2, ...,)$ is given by,

$$P(m_1, m_2, ..., m_R; s, \mathcal{C}) = \prod_{r=1}^{R} \frac{\exp(\Theta_{lm_r t}(s))}{\sum_{m \in \mathcal{C} \setminus m_{-r}} \exp(\Theta_{lmt}(s))}.$$

This expression is a concatenated probability of the usual logit choice probabilities, with the choice set adjusted at each stage.

Despite the seeming complexity of our demand system, it is quite tractable and encompasses standard models as special cases. First, if there is no unobservable heterogeneity or occupational choice component in the labor market, then this model collapses to one where workers estimate the net present value of earnings by taking the net present value of cross-sectional average earnings in each field. Second, if $\beta = 0$, then this collapses to a model where workers only care about their initial earnings on the job market.

5 Estimation

We present our estimation strategy in the same order as the model: first we estimate the career choice block of the model, then turn to the field choice block of the model.

5.1 Labor Supply of Incumbent Workers

Our estimation of the labor supply parameters for incumbent workers largely follows Traiberman (2019), building on Arcidiacono and Miller (2011) and Artuç et al. (2010). We implement a two stage estimator: first we use the EM algorithm to estimate the parameters of human capital supply, the distribution of unobserved heterogeneity, and the transition probabilities across occupations; second, we estimate the remaining structural parameters, C and ν, by inverting workers' choice probabilities (Hotz and Miller, 1993).

Turning to the first stage, we assume that log observed wages of individual i at t are the model's wages plus a Gaussian noise term.[11] Let $f(w_{it}|k, o, m, t; \alpha)$ be the Gaussian PDF of log wages conditional on type, field, occupation, and year, with parameters α. Let q_k be the probabil-

[11] In Traiberman (2019), this noise term is treated as ex post iid heterogeneity. This can be included in the model, as it shifts the value of an occupation. Nevertheless, since this ex post shock has no effect on selection, we ignore it in our present setting.

ity of being type k. In this case, the likelihood contribution of person i is given by,

$$\tilde{\mathcal{L}}_i = \sum_{k=1}^{K} q_k \times \lambda^e(o_1, m|k) f(w_{i1}|k, o, m, 1; \alpha) \times \left[\prod_{t=2}^{T} f(w_{it}|k, o, m, t; \alpha) \times \lambda_t(o_t|o_{t-1}, m, k) \right].$$

The full likelihood, given that q_k is unobserved, is the product over $\tilde{\mathcal{L}}_i$. Optimizing this function using typical optimization routines is difficult for two reasons: (1) the likelihood is not log-linear; and (2) λ_t is the outcome of the solution of the structural model, requiring knowing future continuation values and their expectations.

A standard solution to the first issue is to use the Expectation-Maximization (EM) algorithm. This necessitates defining an auxiliary function, the mean of the likelihood conditional on the distribution of k (as if it were known). The second issue is more specific to our setting. However, as pointed out by Arcidiacono and Miller (2011), we can treat λ as a direct parameter to be estimated, and then map this back to structural parameters in a second stage. Putting this all together yields the following function, which we continue to call the log likelihood in a mild abuse of terminology:

$$\hat{\mathcal{L}} = \sum_{i=1}^{N} \sum_{k=1}^{K} \left[\sum_{t=1}^{T} \log f(w_{it}|k, o_{it}, m_i, t; \alpha) + \log \hat{\lambda}_t(o_{it}|o_{i,t-1}, m_i, k) \right] q_{ik}, \tag{10}$$

where q_{ik} is the probability that individual i is type k, and the hats atop the λ terms imply that they are not outputs of the model, but estimated directly from the data.

One final issue in maximizing (10) is how to deal with initial condition, $\lambda_t(o_1)$, since workers initial occupation is non-random. We follow Wooldridge (2005) and estimate a function, $\lambda(o_{i0}|m_i, k_i, t_{0i})$—where t_{0i} is a cohort fixed effect—and then estimate separately $q(k|m, t_0, s)$, where t_0 is again cohort fixed effects and s are students' GPA (which we observe for a subset of workers). We parameterize $\lambda(o_{i0})$ with a logit link function.[12]

Appendix D contains a full description of how we implement the EM algorithm, but we now briefly describe the idea. Given an initial guess of q_k and α, one can proceed in the following loop:

[12] Formally,

$$q(k|m, t_0, s) = \frac{\exp(\delta_{kmt_0} + \beta_k s)}{\sum_{k'} \exp(\delta_{k'mt_0} + \beta_{k'} s)}.$$

If s were not continuous we could use a fully non-parametric bin estimator. Indeed, for cohorts where we do not observe test scores, this reduces to a bin estimator. For cohorts with test scores, we also include a dummy whenever the score is missing (which we treat as an exogenous event).

1. Update the conditional probabilities for each type:

$$q'_{ik} = \frac{\left[\sum_{t=1}^{T} \log f(w_{it}|k, o_{it}, m_i, t; \alpha) + \log \hat{\lambda}_t(o_{it}|o_{i,t-1}, m_i, k)\right] q_k}{\sum_{k'=1}^{K} \left[\sum_{t=1}^{T} \log f(w_{it}|k', o_{it}, m_i, t; \alpha) + \log \hat{\lambda}_t(o_{it}|o_{i,t-1}, m_i, k')\right] q_{k'}},$$

   which is Bayes' rule.

2. Update model parameters. We update q_k by taking the mean of q_{ik} across individuals. Wage parameters can be estimated using a weighted linear regression of $\log w_{it}$ on occupation-year fixed effects, occupation-field fixed effects, and tasks, where the weights are q_{ik}. $\hat{\lambda}$ can be updated through a weighted bin estimator of transition rates:

$$\hat{\lambda}(o'|o, k, m, t) = \frac{\sum_{i \in (o', o, k, m, t)} q_{ik}}{\sum_{i' \in (o, k, m)} q_{i'k}}.$$

This process is guaranteed to monotonically increase towards the maximizer of the likelihood function.

In the second stage of the estimation procedure, we use the Hotz-Miller inversion to recover the switching cost parameters and the scale of the Gumbel distribution. To see how this is done, notice that after estimating the first stage, $w_{ot}H_o(k, m)$ is observed for every possible (o, k, m), as are the switching probabilities, $\lambda_t(o'|o, k, m)$.[13] In Appendix D.2.2, we show how to use (6) to arrive at the following regression equation:

$$\log\left[\frac{\lambda_t(o'|o, k, m)}{\lambda_t(o|o, k, m)}\right] + \beta \log\left[\frac{\lambda_{t+1}(o'|o', k, m)}{\lambda_{t+1}(o'|o, k, m)}\right] = -\frac{C_{oo'km}(1-\beta)}{\nu} + \frac{1}{\nu}\left(w_{o't}H_o(k, m) - w_{ot}H_o(k, m)\right) + \xi_{oo'kmt+1}, \quad (11)$$

where $\xi_{oo'kmt+1}$ is a time $t+1$ expectational error. Compared to equation (6), the current and next period's continuation values are differenced out, which addresses the remaining selection issue. Under rational expectations, the error is uncorrelated with time t variables, so that equation (11) can be estimated by ordinary least squares.

[13] Even though these are generated regressors, we will avoid the cumbersome use of hat notation. Instead, it should be understood that these are estimated.

5.2 Labor Supply of New Cohorts

The only parameter that remains to be estimated for new cohorts are the entry costs C_e. We can exploit a very similar strategy as above to back out these parameters. We normalize, without loss of generality, the entry cost into the occupation 1 ("Management") to zero. In Appendix D.2.2 we derive the following estimating equation:

$$C^e_{o,km} = \Big(w_{ot}H(o,k,m) - w_{1t}H(1,k,m)\Big) + \beta C_{1,o,km} - \nu \times \left\{\log \frac{\lambda^e_t(o|k,m)}{\lambda^e_t(1|k,m)} + \beta \log \frac{\lambda_{t+1}(o|o,k,m)}{\lambda_{t+1}(o|1,k,m)}\right\}. \quad (12)$$

We take the time average of this expression as our estimate of C_e.

Before we turn to the estimation of the field choice block, we stress that at this point our estimation strategy has avoided explicitly computing V_e in levels. Instead, we have used the inversion of choice probabilities to construct regression equations that difference out these terms. Nevertheless, the net present value of entering an occupation is an input into the field choice problem. However, we now describe a recursive formulation of the utility function that allows us to avoid needing to observe V^e_t.

5.3 Field Choice

First, we demonstrate how to represent payoffs recursively. This will allow us to exploit the panel dimension of our applications data. We begin by deriving all of our results under the assumption that students' latent type, l, is observed. Afterward, we describe how to use the EM algorithm to obtain the distribution of types.

The first step is to estimate the education production function through the regression:

$$q_{ik} = \beta_m + f(s;\beta^s_m) + \epsilon_{ikm},$$

where q_{ik} is the probability that individual i is type k, m is their field, and s is their test score, and f is a function of test scores.[14] We then construct the payoff to a field m at time t for type l individuals as,

$$U_{tlm}(s) = \vartheta_{lm} + v\log(N_{mT}) + \theta_l\left[\sum_k \pi(k|s,m)V^e_{t+5}(k,m)\right]. \quad (13)$$

As mentioned at the end of section 5.2, V^e is not directly estimated. Technically, one only needs to know the differences in V^e across fields, but even these differences cannot be directly estimated

[14] In the previous section we had estimated a similar function for initial conditions. The difference here is that we bin students into quintiles to economize on computation time.

as there is no simple representation of these differences that does not still require knowledge of some terminal value. To resolve this issue, we do not try to estimate the structural parameters for the last period in our data, instead we directly estimate the composite term $\tilde{U}_{tlm}(s) \equiv (U_{Tlm}(s) - v\log(N_{mT}))$.

Estimating $\tilde{U}_{tlm}(s)$ provides an estimate of differences in V^e within it. It turns out that we can use this to estimate the structural parameters using earlier periods of data. Specifically, in Appendix D.2.2, we derive the following representation for the value of field m and time t for a type (l, s) student:

$$U_{\bar{T}-\tau,lm}(s) = \vartheta_{lm}(1-\beta^{\bar{T}-\tau}) + v\log(N_{m,T-\tau}) + \theta_l \underbrace{\left[\sum_{\tau'=\tau}^{\bar{T}-1} \beta^{\tau'-\tau} V^F_{\bar{T}-\tau'+5}(m,s)\right]}_{XV^F_{\tau,\bar{T}}(m,s)} + \beta^{\bar{T}-\tau}\tilde{U}_{\bar{T},lm}(s). \quad (14)$$

where,

$$V^F_{t+5}(m,s) = \sum_k \pi(k|m,s)\left\{-(1-\beta)C^e_{okm} + w_{o,t+5}H(1,k,m) - v\log\lambda^e_{t+5}(o|k,m) - \beta v\log\left(\frac{\lambda_{t+6}(o|o,k,m)}{\lambda^e_{t+6}(o|k,m)}\right)\right\},$$

is an entry value term that *only* depends on observable data. Notice that if $\tau = \bar{T}$, this collapses to the final period utility. Otherwise, this is a discounted sum of observed payoffs, using the discounted final period utility as a control for the continuation value. With (14) in hand, we turn

to deriving the likelihood function. We continue conditional on l, and discuss estimating the latent type distribution at the end. To construct the likelihood conditional on type, let i index individuals, let $\mathcal{C}_i$ be their choice set (which is a function of calendar time, GPA, and high school background), and let $(m_1, m_2, ..., m_R)$ be their ranked choice list. Suppose first that we observe individual i in the final period of the education data, $T-6$. For their *first* ranked choice, conditional on being of type l and having score in decile s_i, the likelihood contribution will be:

$$L_{i,T-6,1} = \frac{N^v_{m_{r_i}} \exp\left(U_{m_r l_i T-6}(s_i)\right)}{\sum_{m'\in\mathcal{C}_i} N^v_{m'} \exp\left(\tilde{U}_{m' l_i T-6}(s_i)\right)}.$$

After a choice has been made, the number of options needs to be adjusted. The notation can now be cumbersome, so let $N_{i,t,r,m}$ be the number of options person i at time t has at position r in field m given their past choices. For example, if there were 40 nursing programs, and an individual

filled out a nursing program as their first priority, then there are 39 programs still available for the second priority. With this notation, the likelihood contribution for the full ranked list is given by,

$$L_{i,T-6}(m_1, m_2, ..., m_R, s_i) = \prod_{r=1}^{R} \frac{N^v_{i,T-6,r,m_{r_i}} \exp\left(\tilde{U}_{m_r l_i T-6}(s_i)\right)}{\sum_{m' \in \mathcal{C}_i} N^v_{i,T-6,,r,m'} \exp\left(\tilde{U}_{m' l_i T-6}(s_i)\right)}.$$

While the utility term, $\tilde{U}$, appears in the remainder of the likelihood, for computational convenience, we will only update this term using last period data. Ultimately, we will use the EM algorithm once again, since l is unobserved. To this end, letting q_{il} be the probability that i is of type l, we can write the log likelihood contribution from the final period as,

$$LL_{T-6} = \sum_l \sum_i \sum_r q_{il} \left(v \log N_{i,r,m_{r_i}} + \tilde{U}_{m_r l_i T-6}(s_i) - \log \left(\sum_{m' \in \mathcal{C}_i} N^v_{i,r,m'} \exp\left(\tilde{U}_{m' l_i T-6}(s_i)\right) \right) \right).$$

Given a current guess of $\tilde{U}_{\bar{T},lm}(s)$, we can construct the individual likelihood contribution for individual i at time t if they are at type l using the multinomial logit probabilities. Taking the logarithm and summing across all periods and individuals leads to the following objective:

$$\begin{aligned} LL_t = \sum_l \sum_i \sum_r q_{itl} \Bigg(& v \log N_{irm_{r_i}} + \vartheta_{lm}(1-\beta^{T-6-t}) + \theta_l X V^F_{t,T-6}(m,s) + \beta^{T-6-t} U_{T-6,lm}(s) - \\ & \log \left(\sum_{m' \in \mathcal{C}_i} N^v_{irm'} \exp\left(\vartheta_{lm}(1-\beta^{T-6-t}) + \theta_l X V^F_{t,T-6}(m,s) + \beta^{T-6-t} U_{T-6,lm}(s) \right) \right) \Bigg) \end{aligned} . \quad (15)$$

We maximize (15) by iterating on the first order condition, as in Berry et al. (1995). Details, including the exact expression from the FOCs, and their connection to various data moments, can be found in Appendix D.2.2.

Finally we turn to estimation when l is unknown. In this case, we use the EM algorithm estimate the distribution of types. To update the conditional probability that individual i of type l, conditional on a guess of parameters, we use Bayes' rule:

$$q^{(g+1)}_{itl} = \frac{\exp(LL^{(g)}_{itl}) q^{(g)}_l}{\sum_{l'} \exp(LL^{(g)}_{itl'}) q^{(g)}_{l'}}.$$

We update the model parameters given q_{il} by iterating on the likelihood's first order conditions, as described above. We repeat this process, increasing the likelihood in each step, until the parameters converge within some tolerance.

5.4 Identification

We discuss the intuition for identification by breaking down the discussion into the same steps as the estimation. First, identification of wage parameters comes from the panel structure of the wage data, and have a similar identification strategy to fixed effects. If there were no occupations, this would be tantamount to clustering workers by their long-run average quantile in the wage distribution.[15] With occupations, switchers determine how to bunch workers according to comparative advantage. When a worker switches occupation, if they are in a similar position of the income distribution, persistently, then they appear to have absolute advantage in both occupations; on the other hand, if their wage is lower (relative to peers *in the new occupation*) after the switch, they are deemed to have comparative advantage in the initial occupation. These situations are illustrated in Appendix Figure 11 for the case of two types. We caveat that this intuition is incomplete, as we also control for occupational histories, so that workers who cluster in similar occupations are also grouped together.

Second, the identification of switching costs parameters—C, C_e, and ν— follows straightforwardly from the regression specification and comes from the response of occupational switching given changes in wage differentials. The response of *net* flows to wage *differentials* pins down the switching parameter, while the overall level of flows—*gross* flows—pins down the level of switching costs. Finally, differential switching patterns conditional on ν pin down how C varies across occupation pairs.

Finally, we discuss the intuition for the field choice parameters. While all parameters are jointly identified via likelihood maximization—making any discussion an approximation—the most intuitive discussion separated each piece of the EM algorithm. That is to say, we discuss identification conditional on l, and then identification of the type distribution conditional on the other parameters. First, given a distribution of types, the identification of θ_l comes from changes in field choice given changes in returns to the observed component of that field, V^F. Similarly, the fixed effects, ν_{lm}, reflect long-term sorting patterns of students, as well as how students of the same type make decisions given different choice sets. Second, identification of the distribution of types comes from information on the ranked list of choices. Conditional on a set of parameters, students with similar lists are clustered together. Clarity on this point may be easiest to see with an example where there were no clustering. Suppose there were only type of student. In this case,

[15] This is similar in spirit to Bonhomme et al. (2019), and indeed the EM algorithm is often how k-means clustering is implemented in practice.

since the preference shocks are conditionally independent, one's second choice of field would be independent of their first choice, $P(m_{r_2}|m_{r_1}) = P(m_{r_2})$. If there is dependence between first and second choices, students are clustered in a way that ensures conditional independence conditional on type.[16]

6 Estimation Results

6.1 Labor Supply

We separate the discussion of the labor market block of the model into two sets of parameters: those governing human capital, H, and those governing switching costs. Recall that human capital depends on both one's field of study, and their latent type k. In our estimation, we have used 4 types. Turning to human capital, we first discuss the specificity of fields. We focus on a measure that captures the idea that some fields are strongly attached to a few occupations, while other fields send workers to many occupations. To this end, we focus on the parameter α_{om}, which captures the comparative advantage of workers with field m in occupation o, holding their latent type constant.

Table 2: Specificity of Different Fields

	Max	Skew	β
High Health	0.75	1.93	2.21
Architecture	0.47	1.58	1.68
Education	0.43	0.94	1.53
Care	0.41	-0.14	1.66
Law	0.37	-0.03	1.72
Communication	0.33	0.82	1.19
Business Admin	0.30	-0.04	1.55
Physics/Math	0.28	-0.19	1.27
Life Sciences	0.27	0.33	1.20
ICT	0.27	0.32	1.29
Engineering	0.26	0.16	1.42
Medium Health	0.23	-0.35	1.55
Humanities	0.23	0.31	1.13
Arts	0.23	0.34	1.38
Social Science	0.20	-0.40	1.46

In our estimation, we cannot recover α_{om} in levels, we can only recover differences. This is

[16] See Greene and Hensher (2010) or Conlon et al. (2024) for more on the identification of unobservable preference parameters from ordered choice data.

intuitive—we can only recover comparative advantage, which requires both a base occupation, and base field. We construct a measure of specificity based on the Revealed Comparative Advantage (RCA) measure of Balassa (1965). First we construct RCA:

$$RCA_{om} = \frac{\exp\left(\alpha_{om}/\nu\right) / \sum_{o'} \exp\left(\alpha_{o'm}/\nu\right)}{\frac{1}{M-1}\sum_{m'}\left[\exp\left(\alpha_{om'}/\nu\right) / \sum_{o'} \exp\left(\alpha_{o'm'}/\nu\right)\right]} - 1. \tag{16}$$

This measure is the probability that a worker with field m would enter occupation o if only α_{om} determined the returns to an occupation relative to the mean of that probability for all other fields. We subtract 1 for ease of interpretation. When this number is negative, it implies that workers are less likely than everyone else to enter an occupation, and if it is positive, they are more likely. Moreover, the magnitude is informative about the degree of sorting relative to everyone else. Appendix Figure 10 plots RCA across occupations for each field. While there are many ways to summarize this graph, we use the max as our definition of specificity:

$$\text{Specificity}_m \triangleq \max_o RCA_{om}. \tag{17}$$

Table 2 displays our measure for all fields. The second and third columns contain two additional measures: the skewness of RCA_{om}, and the slope of RCA_{om} on its rank for each field. The different measures tell similar stories. Health, Education, and Care are relatively specialized fields. STEM fields are in the middle, while Humanistic and social science fields are much more general. An interesting finding is that for a handful of fields, the skew is negative. This implies that rather than being very well matched to a specific occupation, there might be fields that match very *poorly* to a few occupations, but are otherwise quite general skills.

Turning to the parameters governing switching, table 3 displays estimates for this block of the model. The first row is the value of $1/\nu$ we obtain, scaled so that average wages are equal to 1. The second row is the average value of C/ν across all types and occupations. This second row is a simple average over all occupations and is not weighted by actual switches, nor adjusted for moving shocks. The columns are the results of three different assumptions: the first column presents estimates with no controls for field, or latent types, in the second column we introduce controls for field of study, and the third column presents results for the full model.

Table 3: Labor Supply Parameters

Parameter	(1)	(2)	(3)
$\bar{w}/\nu$	0.30	0.38	1.85
	(.122)	(.041)	(.238)
C/ν	5.77	4.60	4.01
Field Controls		✓	✓
Latent Types			✓

Notes: The first row presents the estimates for $1/\nu$ in (5), scaled by the mean wage. The second row presents the estimates for C/ν in (5). The first column reports the estimates using Artuç et al. (2010), which does not consider heterogeneity by field or heterogeneity in unobserved comparative advantage. The second column introduces worker heterogeneity by field. The third column includes both.

To interpret $1/\nu$, we can calculate the switching semi-elasticity with respect to a temporary shock to wages in occupation o:

$$\frac{d(1-\pi_{oo})}{d\log w_o} = -\frac{w_o}{\nu}\pi_{oo}(1-\pi_{oo}). \tag{18}$$

We focus on the semi-elasticity since π_{oo} is a probability. This number is the percentage point increase in exit from occupation o given a temporary shock to wages. This semi-elasticity scales linearly in $1/\nu$, so we can conclude that controlling for latent type and field increases workers' responsiveness to shocks by a factor of four. Moreover, moving costs decrease by 30%. This result is similar to Dix-Carneiro (2014) and Traiberman (2019), who point out that ignoring heterogeneity equates the earnings of workers on the margin of switching with inframarginal workers. This pushes down ν, and consequently raises the measure of costs.

More novel to our setting is that because workers of different fields differ both in their exit rates and earnings, we can calculate the semi-elasticity in (18) for different fields of study. To do so, for each worker, we calculate their semi-elasticity given their current earnings, current occupation, and the probability they are type k, we then take the mean over all workers and years in the economy within each field. We refer to this as the mean semi-elasticity. $\mathcal{E}_m$.[17]

Table 4 shows the value of $\mathcal{E}_m$ for each field of study. These values are correlated with our previous measure of specificity, but there are clear differences. For example, business administration

[17] Formally, let

$$\mathcal{E}_{imt} = \frac{1}{\nu}\sum_k q_{ik}\hat{w}_{imkt} \times \hat{\lambda}_{o(i)o(i)mkt} \times (1-\hat{\lambda}_{o(i)o(i)mkt}).$$

Then we can define the mean semi-elasticity as,

$$\mathcal{E}_m = \frac{1}{NT}\sum_{i\in m}\sum_t \mathcal{E}_{imt}.$$

Table 4: Elasticity values by Education Field

Field of Study	$\mathcal{E}_m$
Business Admin	0.270
Social Science	0.244
Engineering	0.228
ICT	0.215
Physics/Math	0.193
Law	0.185
Life Sciences	0.180
Architecture/Construction	0.170
Arts	0.134
Communications	0.131
Humanities	0.132
High Health	0.081
Care	0.060
Education	0.050
Med Health	0.046

is highly elastic, but in the middle of the pack using our specificity measure. The differences arise for several reasons, including the fact that $\mathcal{E}_m$ depends not just on m but also on the correlation between m with latent type, k, and other parameters of the model. One interesting source of differences arises from the fact that w_o appears in 18 in levels, while in (17), implicitly, only differences matter. For example, if two fields of study had the same distribution of λ across occupations, and w_o was shifted up uniformly for one field, then that field would be more sensitive to shocks. Hence, even though business administration degrees are more specialized to specific occupations than, say, arts degrees, they are nevertheless more sensitive to shocks because their earnings are higher.

A striking feature of Table 4 is the level of heterogeneity in semi-elasticity across fields of study. For example, in the most responsive field—business administration—a 1% decline in wages raises the exit rate by nearly 6 times as in the least responsive field—medium health. The economy-wide average semi-elasticity is .13. Thus, a 1% temporary decline in earnings in an occupation raises the exit rate by .13 percentage points. This number may seem small, but this is because a 1% temporary shock is itself small. Thus, to put this number in perspective, consider a permanent shock, approximated by $w_o/(1-\beta)$.[18] In this case, a 1% permanent decline in earnings raises the exit rate by 1.7 percentage points.

[18]Formally, the value function changes by the change in wages, as well as through ripple effects that arise from changes in switching probabilities. We ignore these latter terms for simplicity.

6.2 Field Choice

In this subsection, we focus on three aspects of field choice estimation: the sensitivity of field choice with respect to net present value, heterogeneity in preferences by type, and substitution patterns between fields following a shock to net present values. The full set of parameters can be found in Appendix (XX). In our estimation, we set $L = 4$, which was computationally fast while allowing for a good deal of heterogeneity.

Table 5: Field Choice Parameters

Parameter	Type 1	Type 2	Type 3	Type 4
θ_l	0.07	0.47	0.03	0.07
q_l	0.16	0.21	0.34	0.29
$\mathcal{E}^f$	0.17	1.42	0.08	0.20

Notes: The table reports the results of field choice estimation by latent type l. The first row reports the estimates of θ_l in (13), the second row reports the share of each type in the population, the third row reports the semi-elasticity of students' first choice field with respect to a shock.

Turning to the first set of results, Table 5 contains our estimates of θ for all types, as well as their distribution. In addition, we measure the mean semi-elasticity of students' first choice field with respect to a shock. Specifically, for each student i we calculate,

$$\mathcal{E}^f_{il} \equiv \frac{d\pi_{m(i)l}}{d\log V^e} = \pi_{im(i)l}(1-\pi_{im(i)l})\theta_l V^e(m, s(i)),$$

where $m(i)$ and $s(i)$ are the field choice and score of student i. At this point, we need to measure V^e in levels, albeit it is measured in differences. To do so, we make a simple assumption: that the entry cost into retail is one quarter of a year's earnings, as one quarter is the average time it takes to find a retail job.[19] Our measure of sensitivity is the average of this across students, weighted by their probability of being a specific type:

$$\mathcal{E}^f_l = \frac{1}{\sum_i q_{il}} \sum_i q_{il} \times \mathcal{E}^f_{il}.$$

This measure has the interpretation as being the percentage point decrease in the fraction of students' choosing a field given a 1% decline in the NPV of that field.

[19] Our results are not particularly sensitive to this normalization. For example, we have also used an approximation to V^e based simply on the net present value on the labor market—$\sum_o \pi_{om} \frac{w_o + \nu \log(\pi_{oo})}{1-\beta}$—and find similar numbers. This is because those occupations with high costs of entry are precisely those that workers enter very rarely (e.g., management). We have also used $w_o/(1-\beta)$ to consider a shock to income, holding occupational choices fixed, and this also does not qualitatively change results. Appendix E contains details on how we approximate V_e, including how we take averages over time.

Table 5 displays two key results. First, students are relatively sensitive to shocks. Taking the type-weighted average of the last row—which is .41— suggests that a 10% decrease in the net present value of a field leads to a 4 percentage point decrease in the choice probability for that field. While small, this number is substantially larger than many estimates in the literature. For example, Beffy et al. (2012) estimates that a 10% increase in returns to STEM fields only increases entry by a quarter of a percentage point. Wiswall and Zafar (2015) estimate that the elasticity of the log odds ratio of switching majors with respect to a shock to be 1.6, while the equivalent number in our estimation would be 2.5. There are many possible explanations for our findings. The Danish system may simply better inform students of the returns to different fields. We also consider a larger set of fields, and as we will see below, a good amount of reallocation is within broad categories, such as STEM.[20]

Our second key finding is that there is substantial heterogeneity across students in their sensitivity. For example, the most sensitive students—those identified d as Type 2—are 18 times more sensitive to shocks than the least sensitive—those identified as Type 3. This type of heterogeneity can be very important for the effects of counterfactuals—shocks that primarily affect relatively flexible types will have quantitatively and qualitatively different economic consequences than those affecting relatively inflexible types.

Aside from heterogeneity in sensitivity, we find that there are significant differences in preferences by type. Table 6 shows the distribution of first priority by student type, with the last column being the unconditional distribution in the population. The four types align relatively well with common wisdom, and with the priority matrix plotted in section 2.3. Type 1 strongly prefers education, nursing, and care. The second type is a "health care" type, with a strong preference for high health (doctors, dentistry, etc.) or medium health (nursing, technicians, etc.). The third type is a STEM type. Interestingly, there is still a substantial fraction of this type that chooses Business or Care. There are two reasons for this. First, STEM is not popular overall, as can be seen in the last column, so that most students, even within this type, will opt for fields like Business. Second, not all students of the same type have the same access to choice sets, either because of their test scores or their high school background. As we will see below, substitution between fields paints

[20]Interestingly, we do not find that the differences are driven by our use of ranked choice data, which is a different source of variation than in many other papers. If we only estimate one type, and ignore the ranked lists, we estimate similar mean semi-elasticities—but very different substitution patterns. On the other hand, we do find that controlling for choice set is very important, and assuming that all students have access to all fields of study lowers our estimates considerably. At the extreme of ignoring all observable heterogeneity, and simply inverting aggregate choice probabilities, our estimates are close to 0.

Table 6: Distribution of First Priority by Student Type

	Type				Population
	1	2	3	4	
Education	60.6	0.7	0.7	1.9	10.4
Medium Health	7.7	54.4	1.3	0.7	13.2
Care	18.6	1.9	22.1	0.8	11.1
Arts	0.5	0.0	0.9	6.3	2.2
Humanities	2.8	0.7	1.0	22.2	7.4
Communication	2.1	1.2	5.7	18.1	7.8
Social Science	3.1	12.1	2.5	21.7	10.2
Business Admin	0.7	3.8	17.3	15.5	11.4
Law	0.3	3.7	4.4	5.7	4.0
Life Sciences	0.4	3.0	8.1	0.4	3.6
Physics/Math	0.2	0.6	4.5	0.7	1.9
ICT	0.0	0.0	3.2	1.4	1.5
Engineering	0.5	0.5	13.1	1.4	5.1
Architecture	0.2	0.2	6.2	2.6	2.9
High Health	2.3	17.2	8.9	0.8	7.2

a slightly different picture of student preferences. Finally, Type 4 is a humanities/social science type.

	Initial Field														
Substitute Field	Education	Medium Health	Care	Arts	Humanities	Communication	Social Science	Business Admin	Law	Life Sciences	Physics/Math	ICT	Engineering	Architecture	High Health
Education		5.26	37.80	2.82	3.78	3.18	1.61	1.49	0.69	0.71	0.55	1.74	0.85	1.20	1.19
Medium Health	17.60		15.26	1.11	1.19	1.67	4.89	1.34	1.80	31.18	3.33	0.98	2.65	1.31	60.38
Care	51.72	14.83		1.53	1.31	2.91	0.93	3.16	0.82	3.14	4.00	2.65	8.49	6.06	1.50
Arts	1.10	0.21	1.87		8.34	6.80	2.08	2.70	1.08	0.20	0.51	5.10	0.87	2.59	0.06
Humanities	9.04	1.95	3.23	28.21		29.78	12.18	10.15	5.54	0.79	1.99	3.41	1.27	6.25	1.28
Communication	6.01	2.24	6.53	18.37	26.23		10.67	12.44	6.02	1.09	2.92	3.96	3.12	4.46	1.33
Social Science	5.46	6.64	3.07	24.05	31.98	24.23		37.18	51.90	2.75	10.82	12.39	3.64	12.63	12.23
Business Admin	2.22	1.72	7.65	12.65	14.71	18.90	23.42		21.08	4.02	12.54	16.08	13.86	13.47	2.26
Law	0.66	1.63	3.49	5.99	6.52	5.97	22.62	15.94		2.09	4.30	6.64	4.73	6.62	2.60
Life Sciences	0.67	7.26	6.38	0.32	0.38	0.72	0.43	1.76	0.67		13.61	6.39	16.20	7.52	11.09
Physics/Math	0.25	0.77	1.99	0.45	0.68	0.59	2.32	2.54	1.52	4.35		4.56	9.03	4.78	2.44
ICT	0.06	0.01	0.97	1.39	0.69	0.81	0.22	1.92	0.26	1.08	2.98		7.95	7.64	0.12
Engineering	0.60	1.10	4.28	0.75	0.83	0.66	0.38	2.90	0.51	9.72	17.93	20.30		22.22	2.99
Architecture	0.37	0.46	3.06	1.34	2.04	1.19	0.91	2.54	0.73	2.68	5.50	12.88	15.72		0.54
High Health	4.23	55.92	4.41	1.02	1.31	2.60	17.33	3.95	7.38	36.19	19.03	2.92	11.62	3.24	

Table 7: Substitution Matrix $\left(\frac{d\pi_{m'}/d\log V_m}{d(1-\pi_m)/d\log V_m}\right)$

Finally, before turning to counterfactuals, we look at aggregate substitution patterns between fields of study. To do so, for each field m we calculate the shares of where students go following a shock to that field's NPV. For example, if there were a 1% shock to the returns to being in education, we see what fraction go to Medium Health, to Care, etc. In a standard multinomial choice model without heterogeneity, it is well known that these substitution patterns are independent of any parameters, and proportional to initial conditional probabilities: $\frac{d\pi_{m'}/d\log V_m}{d(1-\pi_m)/d\log V_m} = -\pi_{m'}/(1-\pi_m)$. However, when agents are heterogeneous this need not be the case. Instead,

$$\frac{d\pi_{m'}/d\log V_m}{d(1-\pi_m)/d\log V_m} = -\frac{\sum_i \sum_l q_{il}\pi_{im'}\pi_{im}\theta_l V_{im}^e}{\sum_i \sum_l q_{il}(1-\pi_{im})\pi_{im}\theta_l V_{im}^e},$$

where the i indexing refers to the variables at student i's characteristics. Aside from heterogeneity in test scores and choice sets, students all differ due to their latent preferences. Table 7 presents the substitution matrix, with the shocked field in the columns, and the other fields in the rows (hence, columns sum to 100%).

In line with the previous discussion, the substitution patterns reveal clear "clustering" of preferences. Two striking examples are the Education column and the Engineering column. In education, nearly all substitution is towards social work and health fields, with a small share of students moving into other fields. This is in line with most students opting for education are Type 1 students, who also have strong preferences for Care and Medium Health. However, these kinds of patterns could not be matched without our model of preferences. Following a decline in the NPV of engineering, 25% of students that switch out move into Physics, Math, or Life Sciences, even though these are only 5.5% of total students. Thus, the substitution patterns are realistic in terms of picking up certain clusters.[21] In Appendix Table 12 we present the substitution matrix without estimating any type heterogeneity, and using only students' first priority to recover preferences. However, we still allow for observable heterogeneity in test scores and in choice sets (arising from different backgrounds). In this case, we find unrealistic substitution patterns. Most strikingly, in all fields, nearly 10% of students that switch fields following a shock switch into Education. This contrasts dramatically with Table 7 where for most fields, this number is less than 2%, often less than 1%. This finding in particular highlights the importance of carefully modeling preference heterogeneity, and shows the value in our method.

[21] Relatedly, notice that shocks to STEM fields (Physics, ICT, Engineering), do not lead to a large exodus to Care, despite the fact that these were clustered together by preference type. This reflects the choice set differences between those students who could opt into STEM fields in the first place.

The substitution patterns we demonstrate above have important implications for our counterfactuals—while students respond to income shocks, they still appear to largely cluster in a narrow set of fields. Hence, even sizable shocks may lead to limited reallocation. This will depend on the exact shocks under consideration and their correlation with students' preference patterns. In the next and final section of the paper we use counterfactual simulations to shed light on our model's implications for reallocation and welfare in response to realistic shocks.

7 Counterfactuals

This section first outlines the remaining ingredients of the model, and then describes the two sets of counterfactual experiments. The first counterfactual evaluates the impact of easing capacity constraints of tertiary education. The second one evaluates the impact of simulated labor market shocks under different policy scenarios.

7.1 Closing the Model

We briefly describe the labor demand side of the model. More details are available in Appendix F. In what follows, $s \in \mathcal{S}$ index industries, and D and F refer to Denmark and Foreign, respectively.

There is a final non-traded good from which agents consumed. The final good aggregate industrial outputs, as well as domestic and foreign varieties in a nested-CES fashion. In the top nest, the final good aggregator, Y_t, is given by

$$Y_t = \left(\sum_s \mu_s^{\frac{1}{\sigma}} Y_{st}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{19}$$

where Y_{st} are industrial outputs and μ_s is an exogenous demand shifter for industry s. Industrial outputs are themselves non-traded aggregates of domestic and foreign varieties given by:

$$Y_{st} = \left(\left(Y_{st}^D\right)^{\frac{\eta-1}{\eta}} + \left(Y_{st}^F\right)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \tag{20}$$

where Y_{st}^D is local absorption of domestic (i.e., Danish) output and Y_{st}^F are imported varieties for an aggregate foreign output, with P_{st}^D and P_{st}^F being their appropriately defined price indices. The elasticity of substitution on industrial outputs, σ, and on domestic versus foreign varieties, η, obey the standard regularity condition $\eta > \sigma > 0$.

In addition to consumer demand at home, there is demand for domestic output abroad. We assume that export demand in industry s is exogenously determined by the following demand curve:

$$X_{st} = A_s^F \left(P_{st}^D\right)^{-\eta}, \tag{21}$$

where A_s^F is a demand shifter.

Domestic production in each industry s is carried out by a large number of perfectly competitive production units. Each unit produces domestic output Q_{st} by aggregating production and non-production services. Accordingly, we categorize occupations into production and non-production groups, denoted by $\mathcal{O}^P$ and $\mathcal{O}^N$, respectively. Production takes place according to the following CES function:

$$Q_{st} = A_{st}\left[\left(A_s^N\right)^{\frac{1}{\alpha}}\left(L_{st}^N\right)^{\frac{\alpha-1}{\alpha}} + \left(\left(A_s^P\right)^{\frac{1}{\rho}}\left(L_{st}^P\right)^{\frac{\rho-1}{\rho}} + \left(A_s^M\right)^{\frac{1}{\rho}}\left(Q_{st}^M\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}\frac{\alpha-1}{\alpha}}\right]^{\frac{\alpha}{\alpha-1}}, \tag{22}$$

where L_{st}^N is the aggregate output by non-production occupation services, L_{st}^P is the aggregate output by non-production occupation services, Q_{st}^M is the aggregated imported material inputs, A_{st}, A_s^N, A_s^P, and A_s^M are the sector-specific TFP, and productivities of non-production, production services and imported inputs, respectively, $\alpha > 0$ is the elasticity of substitution between L_{st}^N and the the production aggregate within each industry, and $\rho > 0$ is the elasticity of substitution between L_{st}^N and Q_{st}^M within each industry.

We assume that production and non-production inputs are themselves Cobb-Douglas aggregates over occupations, for $f = \{N, P\}$

$$L_{st}^f = \prod_{o \in \mathcal{O}^f}\left(\frac{L_{sot}^f}{\beta_{so}}\right)^{\beta_{so}}, \tag{23}$$

where $\sum_{o\in\mathcal{O}^f}\beta_{so} = 1$, and L_{sot} is the efficiency units (human capital) in each occupation within each industry. Since workers are indifferent across sectors within an occupation, total human capital within an occupation is given by $L_{ot} = \sum_s L_{sot}$. Finally, consistent with the labor supply model, workers spend their entire income in the period it is earned. The aggregate expenditure is

therefore equivalent to the aggregate income:

$$E_t = \sum_o L_{ot} W_{ot}.$$

To run the counterfactual analysis, we specify the full set of demand-side parameters by calibrating directly from aggregate labor shares or borrowing from literature. We provide more details in Appendix F.

7.2 Capacity Constraints in General Equilibrium

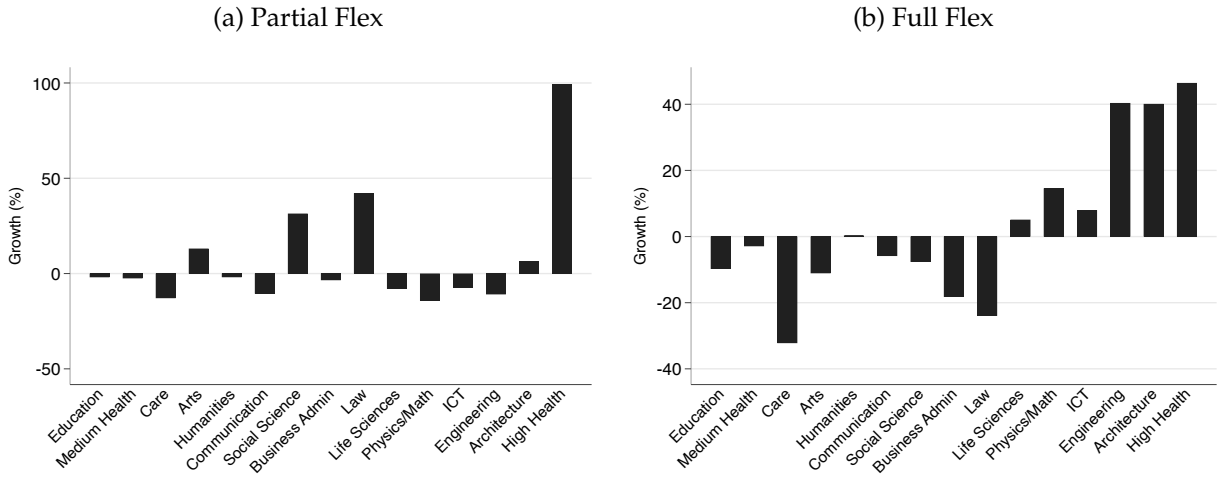
In our first counterfactual we compare the effects of Danish capacity constraints and tracking on worker allocations and GDP. Specifically, we consider two changes. First, we remove capacity constraints but maintain high school tracking. That is to say, we only allow students to apply to programs for which the highest scoring students in their high school track were eligible. Second, we remove capacity constraints and tracking, by allowing students to freely enter into any field of study. We refer to these as "Partial Flex" and "Full Flex," respectively. Both scenarios map to different policy-relevant counterfactuals. For example, Denmark has expanded upper secondary supplemental coursework to offer students more flexibility after secondary school (*Gymnasiale suppleringskurser*).

At the outset, there are two caveats to keep in mind. First, we do not take into account the costs of increasing capacity. Second, we assume that students who start programs are able to finish them. The second issue is partially tempered by the fact that students' outcomes can depend on their high school performance. Thus, students may understand that they may be low-earners in a particular field, and thus still go against. That said, our model may still overestimate students' abilities to graduate in certain majors. To this end, in Appendix (XX), we offer one further scenario—we do maintain cutoffs and tracking specifically for medicine, a field that attracts many students but our model may be ill-suited for.

Turning to the results, Figure 7 plots field reallocation under both scenarios. In both cases, there is substantial reallocation of majors. In the partial flexibility case, 9% of individuals switch their major across steady states, and this number is 21% in the full flexibility case. This suggests that constraints on students are strong and binding given their preferences.

Appendix Table 13 shows the matrix of conditional switching shares—the share of switchers entering row field m' from column field m. The switching patterns are in line with the patterns

Figure 7: Field Reallocation w/o Capacity Constraints



Notes:

in students' ranked lists, and reflect the preference heterogeneity estimated in the model. Humanistic and social science reallocate among themselves other, often into higher paying fields like law. Similarly, there is reallocation among STEM fields. These changes, which imply that students tend to reallocate across relatively similar fields, rather than across very different fields, may explain the small change in national income. We return to this point when we discuss occupational reallocation. First, we turn to the economic consequences of relaxing constraints.

In designing systems with capacity constraints, policy makers may face a trade-off between respecting the preferences of students, and the allocation of students that maximize national objectives. We do not take a stand on the source of this trade-off, but there are a few potential issues. First, it could be that even after providing students with substantial information, students are still poorly informed about potential outcomes. Thus, strong preferences may still reflect a lack of information, so there may be paternalistic motives. Second, policy makers may wish to maximize fiscal capacity in order to subsidize the education system. While we do not take a stand on the particulars of the social welfare function, we can use our model to ask two questions: what happens to national income, and what happens to income inequality, under different institutional arrangements?

Table 8 displays the results from this exercise. The first row displays the percent change in real income, and the second displays the percent change in the standard deviation of income. Perhaps surprisingly, we do not find that capacity constraints and tracking increase national income. In fact, we find that output would be 1.7% higher under the fully flexible system. Maintaining

Table 8: GE Consequences of Lifting Capacity Constraints

	Partial Flex	Full Flex
%Δ GDP	0.03	1.16
%Δ s.d. log Income	0.90	0.43
Occupational Reallocation	3.47	7.50

tracking, but easing capacity constraints would still increase national income by .3%. While small, neither number suggests that the rigidity of the education system generates higher income.

To understand why national income responds little to the reallocation in majors, we look at amount of occupational reallocation in the economy. Our measure of occupational reallocation is the mean of the absolute value of the change in shares across occupations:

$$Reallocation = 100 \times \frac{1}{|\mathcal{O}|} \sum_{o \in \mathcal{O}} |\log(s_{o,cf}) - \log(s_{o,ss})|,$$

where $s_{o,j}$ is the share of college educated workers in occupation o under steady state ($j = ss$) and the counterfactual ($j = cf$). Occupational reallocation is an order of magnitude smaller than field reallocation.

In order to assess distributional consequences, we simulate and follow a large cohort of individuals. More precisely, we simulate the field choice and labor market outcomes of 120,000 individuals drawn from the underlying distribution of student observables. For each individual we then compute a measure of long run earnings:

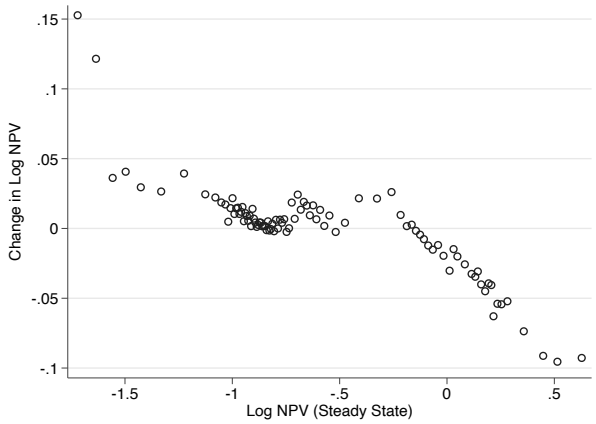
$$NPV_i = (1 - \beta) \times \sum_{t=0}^{50} \beta^t w_{it}.$$

There is a small increase in the standard deviation of incomes. For example, in the partial flexibility case, the standard deviation increases by around 1 tenth of a percent. To make sense of these forces, note that, holding earnings fixed, the effects of relaxing capacity constraints depends on the correlation between capacity constraints and earnings. If capacity constraints tend to bind on lower earning fields, then allowing marginal students to enter these fields could lead to more inequality. This is especially true since students below the cutoffs have lower test scores, and so the earnings losses in low paying fields may be exacerbated. On the other hand, if capacity constraints allow lower scoring students into higher earnings fields, then this could raise their earnings (and consequently put competitive pressure on higher scoring students, further lowering

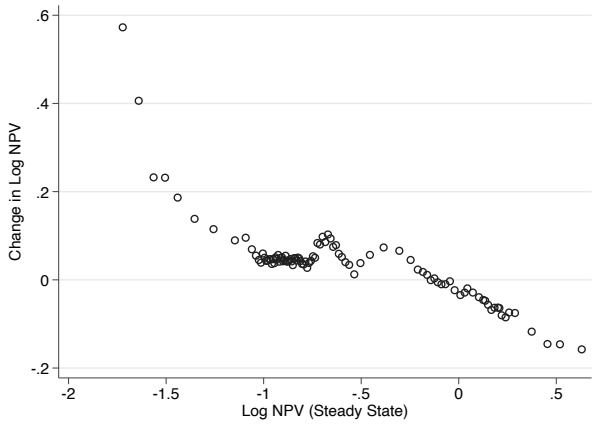
their earnings). The steady state correlation between the probability of being accepted to a program and earnings is low—suggesting little to no systematic pattern, explaining the small change in inequality. In addition to this logic, there are general equilibrium forces at play, where students reallocating across fields, and subsequently occupation, change the earnings of those who do not change their field.

Figure 8: Removing Capacity Constraints and Inequality

(a) Partial Flex



(b) Full Flex



Notes:

The focus on standard deviation, while convenient, can potentially obscure more interesting distributional consequences. To gain a more granular understanding of how the distribution of income changes, in Figure 8 we show a bin scatter of the change in the long run income of a simulated individual in each student against their long run income in the capacity constrained steady state.

We see that in the center of the distribution—which is the majority of the weight—there is no relationship, or a small positive relationship, between initial earnings and the change in earnings. This is in line with the small change in standard deviation. However, we see that in the tails of the distribution there is strong convergence—the lowest earning workers see their earnings rise the most, and the highest earning workers see their earnings decline.

Our results suggest that relaxing the rigidity in the education system may improve welfare and output, with negligible impacts on inequality. These results are tempered by the lack of information on the costs of relaxing these constraints, but they are indicative that offering students more flexibility may not be particularly detrimental. This discussion focuses on steady state, and less on how different students fare in response to shocks. In our second counterfactual, we turn to

understanding the resilience of different fields of study.

7.3 Field Choice and Labor Market Shocks

In this section, we examine the effects of two labor market shocks on welfare, occupational and field-of-study reallocation, income dispersion, and intergenerational inequality under different policy scenarios. First, we analyze a foreign price shock that lowers the cost of imported goods and offshored inputs, shifting the relative demand for domestic labor across occupations. Second, we consider a technology shock driven by AI, which changes skill prices in certain occupations, altering employment patterns and the incentives for skill acquisition.

7.3.1 International Trade Shocks

n this exercise, the analysis begins from the steady-state equilibrium projected forward from 2018. We calibrate an average 20% reduction in prices, roughly matching the observed decline in foreign prices in Denmark over the past two decades. These shocks generate heterogeneous labor market effects across occupations: reductions in the cost of imported goods affect all sectors, while declines in offshoring costs disproportionately impact occupations that are more substitutable with offshored inputs.

In our model, labor market adjustment to shocks occurs not only through wage changes and occupational re-sorting but also through shifts in students' field-of-study choices, which shape the long-term supply of skills. To assess the role of educational flexibility in this process, we compare four policy scenarios that vary the extent to which students can adjust their field choices in response to labor market shifts.

- No Field Reallocation: A counterfactual scenario in which students' field choices remain fixed, isolating the effects of occupational reallocation among incumbent workers.
- Current System: Reflecting existing institutional constraints, where students can adjust their field choices but are limited by capacity restrictions.
- Partial Flex: A scenario in which capacity constraints are removed, but students remain subject to high school tracking, limiting entry to programs based on prior academic eligibility.
- Full Flex: A scenario in which both capacity constraints and tracking restrictions are lifted, allowing students to freely select any field of study.

To assess the impact of labor market flexibility on welfare and inequality, we compare outcomes across the four policy scenarios. Our analysis evaluates both long-run and short-run effects, focusing on key economic indicators such as welfare changes, income differences, field reallocation, and intergenerational inequality. In the long run, we compare steady-state outcomes across different policy environments, using the baseline steady state as the reference point. For the short run, we examine the transition dynamics over the first 50 years following the labor market shock, focusing on incumbent workers rather than new entrants. Intergenerational inequality is assessed by comparing the lifetime income of incumbents in transition to that of individuals in the new steady state under each counterfactual scenario.

Welfare and Income Effects

Table 9 presents the percentage change in welfare and lifetime income across the four policy scenarios. In the long run, welfare gains increase with greater field choice flexibility, rising from 0.58% in the *No Field Reallocation* scenario to 3.71% under *Full Flex*.[22] Similarly, lifetime income gains follow a similar trend, with the *Full Flex* system yielding the highest increase (27.0%).[23] These results highlight the role of educational flexibility in enhancing long-term economic efficiency by allowing students to better align their skill choices with evolving labor market demands.

Field and Occupational Reallocation

A key mechanism driving these welfare and income effects is the extent of field reallocation. By construction, the *No Field Reallocation* scenario results in zero field mobility, while the Current System allows for limited adjustments (2.57% of individuals switch fields). The *Partial Flex* and *Full Flex* systems lead to significantly higher reallocation rates (18.3% and 26.5%, respectively), reflecting the extent to which relaxing capacity constraints enables students to optimize their field

[22]Specifically, we compute the average percentage differences in welfare as follows:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{U_i^{CF}-U_i^{Base}}{U_i^{Base}}, \tag{24}$$

where U_i^{CF} and U_i^{Base} denote individual welfare under the counterfactual and baseline scenarios, respectively.

[23]The percentage change in lifetime income is calculated as:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{NPV_i^{CF}-NPV_i^{Base}}{NPV_i^{Base}}, \tag{25}$$

where NPV_i^{CF} and NPV_i^{Base} denote the net present value of lifetime income for individual i under the counterfactual and baseline scenarios, respectively.

choices. These shifts in field allocation translate into larger lifetime income gains, demonstrating how greater flexibility in education decisions improves long-run labor market outcomes.

Short-Run vs. Long-Run Adjustments

For the short run, we examine the transition dynamics over the first 50 years following the labor market shock, focusing on incumbent workers rather than new entrants. Unlike the long-run results, welfare gains remain constant across all policy scenarios in the short run (0.46%), suggesting that the initial adjustment primarily occurs through wage and occupational changes rather than educational reallocation. Similarly, income differences across policies are relatively small in the short run, with all scenarios yielding gains of approximately 9.86%–9.87%. These results emphasize the gradual nature of field reallocation and its increasing importance over time as new cohorts enter the labor market.

Intergenerational Inequality

We also evaluate intergenerational inequality, comparing the lifetime income of incumbents in transition to that of individuals in the new steady state under each counterfactual scenario. The results indicate that greater field flexibility slightly increases intergenerational inequality, rising from 0.94 in the *No Field Reallocation* scenario to 0.99 under *Full Flex*. This suggests that while increased educational mobility enhances overall welfare and income levels, it also amplifies differences across cohorts, as younger generations benefit more from adapting to evolving labor market conditions.

Together, these findings highlight the critical role of field-of-study flexibility in shaping economic outcomes. While short-run gains primarily occur through wage and occupational adjustments, long-run effects are increasingly driven by educational reallocation, underscoring the importance of policies that ease capacity constraints in tertiary education.

Table 9: Results across Policy Scenarios

	No Field Reallocation (1)	Current System (2)	Partially Flex System (3)	Fully Flex System (4)
		Long Run		
%△ Welfare	0.58	0.75	1.42	3.71
%△ Income	19.3	20.4	24.4	27.0
%△ s.d. log Income	0.16	0.16	0.17	0.17
Field Reallocation	0%	2.57%	18.3%	26.5%
		Short Run		
%△ Welfare	0.46	0.46	0.46	0.46
%△ Income	9.86	9.86	9.87	9.86
%△ s.d. log Income	-0.34	-0.34	-0.34	-0.34
		Long Run vs. Short Run		
Inter-generational Inequality	0.94	0.94	0.98	0.99

7.3.2 AI Shock

to be completed...

8 Conclusion

In this paper, we model how students adjust their education decisions in response to globalization. Understanding how responsive students are to shocks is a key input to policymakers who wish to use education policy as a lever to address labor market shocks such as trade, immigration, or automation. Our model captures selection on unobservable comparative advantage and preferences, while remaining tractable enough for estimation on standard administrative datasets.

We find that fields of study tend to be quite specialized, but there is substantial heterogeneity with some fields imparting quite general skills. This heterogeneity is not easily characterized by coarse measures such as STEM and Humanities, militating for our disaggregated approach. We also find that students are more sensitive to shocks to future income than many past studies, but that these preferences are very heterogeneous across students. Some students are almost completely inelastic, while others are more willing to switch. We also show the power of our method by showing that modeling latent heterogeneity, exploiting our data on students' ranked preferences, captures realistic substitution patterns across fields. Ignoring this information yields similar sensitivity to shocks, but very different substitution patterns.

In counterfactuals, we find that modeling field of study leads to quantitatively different findings than in models ignoring this channel: the gains from offshoring are up 12% larger, and there is twice as much reallocation in the economy. Perhaps more interesting, our model shows that policy makers can shape the labor market impacts of shocks through educational institutions. Specifically, the "American System" leads to less occupational reallocation than the current Danish system, but leads to 10% larger gains from offshoring, and less income inequality in the economy.

Our research leaves substantial room for future work. For example, in our model we treat one's high school background as fixed. It is plausible, indeed likely, that these decisions are also endogenous. We have also abstracted from institution quality, and other factors that might matter for students' decisions, such as their parents' backgrounds. Finally, the largest determinant of students' choices are embedded in the time-invariant component of preferences—essentially a residual. Opening up this black box—whether it reflects information frictions, secondary school inputs, or "deeper" preferences—could shed further light on what kinds of policy interventions in educational institutions will have impacts on labor market outcomes.

References

Adda, Jerome and Christian Dustmann, "Sources of wage growth," *Journal of Political Economy*, 2023, *131* (2), 456–503.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 2018, *86* (2), 391–444.

Altonji, J.G., P. Arcidiacono, and A. Maurel, "Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 5, Elsevier, 2016, pp. 305–396.

Altonji, Joseph G., Erica Blom, and Costas Meghir, "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 2012, *4* (1), 185–223.

_ , _ , and _ , "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 2012, *4* (1), 185–223.

Andrews, Rodney J, Scott A Imberman, Michael F Lovenheim, and Kevin M Stange, "The returns to college major choice: Average and distributional effects, career trajectories, and earnings variability," Technical Report, National Bureau of Economic Research 2022.

Arcidiacono, Peter, "Ability sorting and the returns to college major," *Journal of econometrics*, 2004, *121* (1-2), 343–375.

_ and Robert A. Miller, "CCP Estimation of Dynamic Models with Unobserved Heterogeneity," *Econometrica*, November 2011, *79* (6), 1823–1868.

_ , Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College Attrition and the Dynamics of Information Revelation," Working Paper 22325, National Bureau of Economic Research June 2016.

_ , V. Joseph Hotz, Arnaud Maurel, and Teresa Romano, "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 2020, *128* (12), 4475–4522.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, June 2010, *100* (3), 1008–45.

Atalay, Enghin, "How Important Are Sectoral Shocks?," *American Economic Journal: Macroeconomics*, October 2017, *9* (4), 254–80.

Balassa, Bela, "Trade Liberalisation and "Revealed" Comparative Advantage," *The Manchester School*, 1965, *33* (2), 99–123.

Balinski, Michel and Tayfun Sonmez, "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 1999, *84* (1), 73–94.

Beffy, Magali, Denis Fougere, and Arnaud Maurel, "Choosing the field of study in postsecondary education: Do expected earnings matter?," *Review of Economics and Statistics*, 2012, *94* (1), 334–347.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–90.

Bleemer, Zachary and Aashish Mehta, "Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major," *American Economic Journal: Applied Economics*, April 2022, *14* (2), 1–22.

Bonhomme, Stephane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, May 2006, *121* (2), 541–585.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Conlon, Chris, Julie Mortimer, and Paul Sarkis, "Estimating Preferences and Substitution Patterns from Second-Choice Data Alone," Technical Report, New York University 2024. Accessed: 2024-08-14.

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885.

Doraszelski, Ulrich and Jordi Jaumandreu, "Measuring the Bias of Technological Change," *Journal of Political Economy*, 2018, *126* (3), 1027–1084.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economic Review*, April 2019, *109* (4), 1486–1529.

Feenstra, Robert C., "New Product Varieties and the Measurement of International Prices," *The American Economic Review*, 1994, *84* (1), 157–177.

Greene, William H and David A Hensher, *Modeling ordered choices: A primer*, Cambridge University Press, 2010.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman, "Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile," Working Paper 19241, National Bureau of Economic Research July 2013.

Hotz, V. Joseph and Robert A. Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, July 1993, *60* (3), 497–529.

Humlum, Anders, "Robot Adoption and Labor Mafrket Dynamics," Working Paper, University of Chicago Booth School of Business 2021.

Kapor, Adam J., Christopher A. Neilson, and Seth D. Zimmerman, "Heterogeneous Beliefs and School Choice Mechanisms," *American Economic Review*, May 2020, *110* (5), 1274–1315.

Kapor, Adam, Mohit Karnani, and Christopher Neilson, "Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms," *Journal of Political Economy*, 2024, *132* (7), 2346–2395.

Keane, Michael and Kenneth Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Khanna, Gaurav and Nicolas Morales, "The IT Boom and Other Unintended Consequences of Chasing the American Dream - Working Paper 460," Working Papers 460, Center for Global Development August 2017.

Kinsler, Josh and Ronni Pavan, "The Specificity of General Human Capital: Evidence from College Major Choice," *Journal of Labor Economics*, 2015, *33* (4), 933–972.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Larroucau, Tomas and Ignacio Rios, "Dynamic College Admissions," 2022. Working paper.

Leighton, Margaret and Jamin D. Speer, "Labor market returns to college major specificity," *European Economic Review*, 2020, *128*, 103489.

Lise, Jeremy and Fabien Postel-Vinay, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, August 2020, *110* (8), 2328–76.

Llull, Joan, "Immigration, wages, and education: A labour market equilibrium structural model," *The Review of Economic Studies*, 2018, *85* (3), 1852–1896.

Luflade, Margaux, "The Value of Information in Centralized School Choice Systems," *Mimeo*, 2019.

Monras, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo, "The Human Side of Structural Transformation," *American Economic Review*, August 2022, *112* (8), 2774–2814.

Robles, Verónica C. Frisancho and Kala Krishna, "Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch," NBER Working Papers 17727, National Bureau of Economic Research, Inc January 2012.

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Traiberman, Sharon, "Occupations and Import Competition: Evidence from Denmark," *American Economic Review*, December 2019, *109* (12), 4260–4301.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 2015, *82* (2), 791–824.

Wooldridge, Jeffrey M, "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of applied econometrics*, 2005, *20* (1), 39–54.

A Additional Tables

Table 10: Field of Study Classification

Field	Field Name	ISCED-F Code	Notes
1	Education	11	
2	Arts	21	
3	Humanities	22, 23	
4	Social Science	31	
5	Communication	32	
6	Business Admin	41	
7	Law	42	
8	Life Sciences	51, 52	
9	Physics/Math	53, 54	
10	ICT	61	
11	Engineering	71, 72	Combined only for highest level of education
12	Manufacturing	72	
13	Architecture	73	
14	Agriculture	80	
15	Medium Health	91	Separated between university and professional bachelor's
16	Care	92	
17	Personal Services	101, 102	
18	High Health	91	Separated between university and professional bachelor's
19	Security	103	
20	Transport	104	

Table 11: Occupation Classification

Occupation	Occupation Name	Disco-08 2 Digit Codes	Notes
1	Management (Business + Admin)	10, 11, 12	
2	Management (Activity)	13, 14	
3	Engineering	21	
4	Medicine	22	
5	Education	23	
6	Finance, Admin, and Sales	24	
7	Software and ICT	25	
8	Social and Cultural Work	26	
9	Law	26	Split out of 3 digit code 261
10	Science Technicians	31	
11	Healthcare Assistance	32	
12	Admin	33, 34	
13	IT	35	
14	Secretarial	41	
15	Customer Services	42	
16	Accounting	43	
17	Other Admin	44	
18	Services	51, 94	
19	Retail	52	
20	Social Work	53	
21	Security	54	
22	Agriculture	60, 92	
23	Machinery	72	
24	Electrical	74	
25	Craftsperson	71, 73, 75	
26	Operators	81, 82	
27	Transportation	83	Includes 9621
28	Other Manual Work	91, 96	Excludes 9621
29	Construction	93	

	Initial Field														
Substitute Field	Education	Medium Health	Care	Arts	Humanities	Communication	Social Science	Business Admin	Law	Life Sciences	Physics/Math	ICT	Engineering	Architecture	High Health
Education		15.15	19.35	10.47	11.97	13.24	9.57	11.49	8.91	8.60	7.13	8.07	9.03	8.83	7.50
Medium Health	18.53		24.78	15.69	13.98	15.37	11.65	15.44	11.71	10.71	9.11	11.79	10.89	11.37	9.49
Care	18.53	19.59		11.51	10.91	13.17	7.55	12.88	7.47	6.85	5.57	6.03	6.01	6.12	5.40
Arts	1.88	2.68	2.51		2.14	2.16	2.19	2.06	2.15	1.57	1.59	2.13	1.43	1.64	1.98
Humanities	9.73	9.52	9.50	8.93		10.71	10.23	7.31	8.33	7.13	7.57	3.47	5.24	4.85	8.90
Communication	9.96	9.21	9.33	8.65	11.30		9.24	8.99	8.35	5.28	5.23	3.04	4.09	3.74	6.57
Social Science	9.59	9.16	7.53	10.33	12.97	12.12		15.99	15.59	11.27	13.47	7.95	10.22	9.29	16.65
Business Admin	11.03	12.21	11.21	11.45	11.27	12.76	13.79		14.23	10.80	11.35	9.26	10.22	9.82	12.12
Law	3.53	3.82	3.17	5.06	4.51	4.37	6.28	6.13		4.41	4.38	5.17	5.27	4.98	5.64
Life Sciences	3.15	3.29	2.65	2.90	3.41	2.45	4.37	2.62	3.48		6.06	7.00	9.21	6.97	5.99
Physics/Math	1.25	1.35	0.94	1.14	1.59	1.01	2.76	1.91	2.02	3.59		2.69	3.90	2.99	4.24
ICT	0.67	1.09	0.70	1.16	0.70	0.58	0.99	1.19	1.03	1.90	1.56		3.45	3.30	1.53
Engineering	3.56	4.35	2.22	3.13	3.65	2.02	5.42	4.21	4.43	12.50	10.87	16.22		16.59	9.46
Architecture	2.17	2.63	1.73	2.14	2.10	1.54	2.91	2.59	2.49	5.57	4.81	8.80	9.84		4.52
High Health	6.41	5.96	4.37	7.44	9.50	8.51	13.05	7.18	9.82	9.82	11.30	8.37	11.19	9.51	

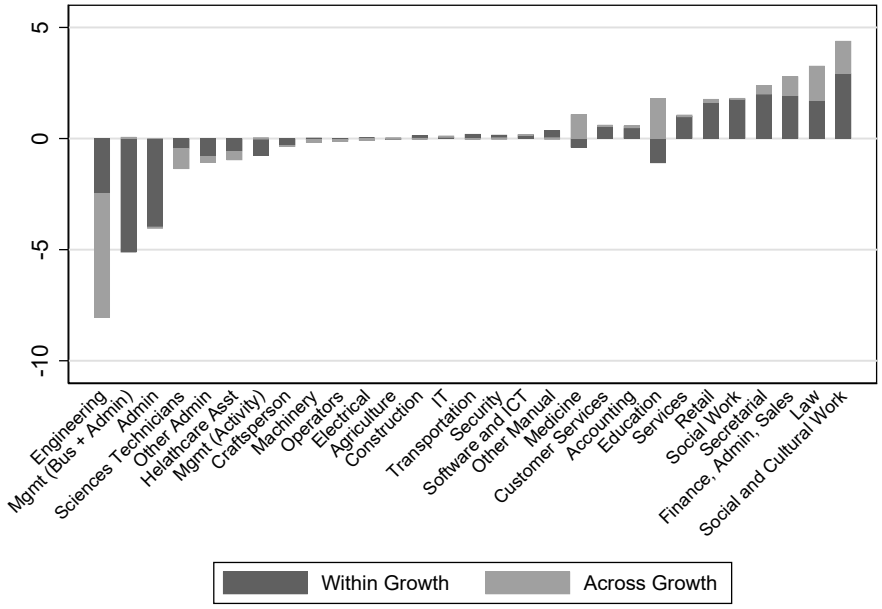
Table 12: Substitution Matrix $\left(\frac{d\pi_{m'}/d\log V_m}{d(1-\pi_m)/d\log V_m}\right)$

Table 13: Reallocation Table

	Initial Field														
Counterfactual Field	Education	Medium Health	Care	Arts	Humanities	Communication	Social Science	Business Admin	Law	Life Sciences	Physics/Math	ICT	Engineering	Architecture	High Health
Education			12.51				0.45	0.17		1.56		0.43			
Medium Health	12.73		4.01			0.75			8.18	0.17		1.10			
Care	2.06							0.09	0.68	0.30	0.35				
Arts	11.24	0.50	3.48		5.12	4.11	4.50	3.94	10.49	3.18	3.08	4.31	3.64	4.29	11.11
Humanities	7.30	1.00	1.64	1.62		8.22	1.35	5.44	10.19	4.64	6.32	5.35	7.39	3.43	
Communication	3.18	0.13	5.22	0.54	0.34		0.45	0.69	1.85	2.64	3.25	2.53	3.52	1.29	
Social Science	26.03	10.90	3.69	35.14	42.49	33.25		32.06	51.23	18.27	26.15	27.19	23.38	35.19	
Business Admin	2.62	0.38	10.18	2.70	2.90	4.73	1.80		8.33	5.82	7.35	5.50	8.78	4.72	38.89
Law	14.79	2.26	6.49	52.43	43.34	33.50	83.33	39.70		19.00	21.88	23.77	18.94	36.05	27.78
Life Sciences	2.25	3.63	4.96	0.54	0.51	1.00	0.45	1.04	4.01		5.30	4.75	6.52	3.00	
Physics/Math	0.19		1.21			0.12	1.35	0.58	1.23	0.64		0.30	0.69		
ICT	0.56	0.25	5.17	2.16	1.02	1.74	0.90	1.39	2.78	2.36	2.91		2.71	0.43	
Engineering	1.31		1.16	0.54	0.34	0.50	0.90		0.93	1.64	2.91	1.19			5.56
Architecture	7.49	0.63	4.96	1.08	1.37	3.11	3.15	2.43	4.32	4.64	5.64	5.05	6.64		16.67
High Health	8.24	80.33	35.30	3.24	2.56	8.97	1.80	12.73	4.63	28.45	14.19	19.76	14.78	11.16	

B Additional Figures

Figure 9: Decomposing the New Cohort Entry Margin



Notes: The decomposition of occupational reallocation follows (2). Occupations are sorted based on total growth over 1998 to 2018.

C Additional Counterfactuals

C.1 Maintaining Medicine Cutoffs

Figure 10: RCA by Field

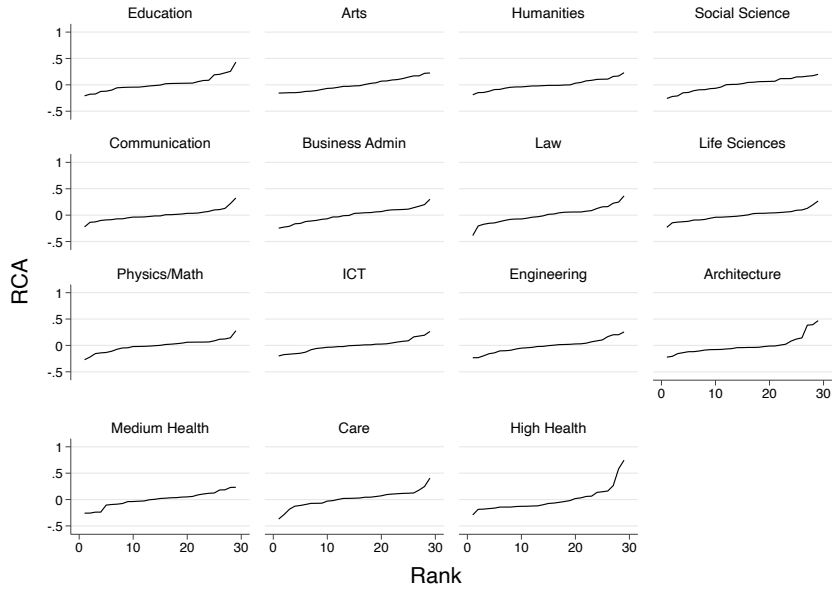
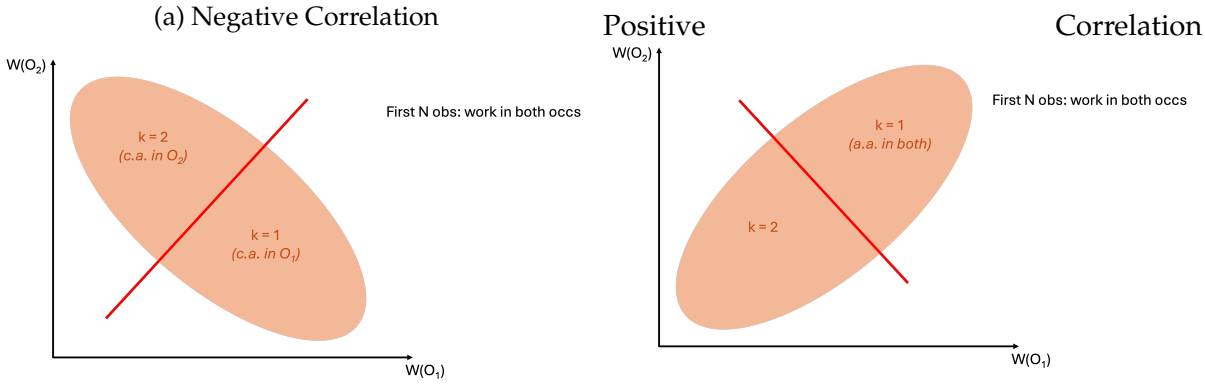


Figure 11: Intuition for the EM Algorithm



D Estimation Details

D.1 Generating Data for Monte Carlo

This section assumes that one has already defined the model and solved for the parameters of the simulation. Subroutines for this procedure are defined elsewhere. Here we outline the loop to generate simulated data. We will index each simulated individual by i and calendar time by t.

1. Simulate skill prices, w_{ot} and solve for the value functions and policy functions in the labor market.

2. Given a cohort of size N_c and T periods, pre-allocate the following data matrices:

   - DLM: an $N_c(T+1)T/2 \times 4$ (this will be excessive given death and graduation time) matrix where each row is (i, t, o_{it}, w_{it})
   - DF: $N_c T \times 2$ vector that contains vector of i's field and test score
   - DK: $N_c T \times 1$ vector that contains vector of i's unobserved labor market type
   - DL: $N_c T \times 1$ vector that contains vector of i's unobserved ex ante type
   - DR: $N_c T \times R$ matrix that contains each N

3. For each i simulate a career path:

   (a) Generate a test score s_i for i

   (b) Generate a type for i from the conditional CDF of the type distribution given s_i:

   $$l_i = 1 + \sum_l \mathbf{1}(e > Q_l(s_i)),$$

   where e is a uniform random variable.

   (c) For $r = 1, ..., R$ calculate the ranked list of choices for individual i:

   - First calculate the conditional CDF of choices for field:

   $$\tilde{Q}_{ir}^F(m) = \frac{\sum_{m'=1\setminus m_{-r}}^{m} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l \sum_k q(k|s,m) V_{t+5}(k,m)\right)}{\sum_{m'=1\setminus m_{-r}}^{M} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l \sum_k q(k|s,m) V_{t+5}(k,m)\right)}$$

   - Second generate a draw from this distribution:

   $$m_{r_i} = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_{ir}^F(m)),$$

   for a uniform random variable.

   (d) Assign the agent the highest ranked m in their choice set conditional on s. If no choice meets the criterion remove them from the sample.

   (e) Determine the agents type:

   $$k_i = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_i^K(m, s)),$$

where Q_k is type CDF for agent i with field m and score s, and e is a uniform random variable.

(f) Assign an initial occupation beginning in t+5 based on the initial occupation policy function, λ^e_{t+5}:

$$o_{it} = 1 + \sum_{o=1}^{|\mathcal{O}|} \mathbf{1}(e > \sum_{o'=1}^{o} \lambda^e_t(o, m_i, k_i)),$$

where e is a uniform random variable.

(g) Starting the agent at $t + 5$, simulate a labor market history

i. Determine if the agent exits by testing a uniform random variable against δ. If retire, continue to next i. Else...

ii. Draw next occupation and wage:

$$\begin{aligned} o_{it} =& 1 + \sum_{o'=1}^{|\mathcal{O}|} \mathbf{1}(e > \sum_{o''=1}^{o''} \lambda_t(o'|o_{it-1}, k, m)) \\ w_{it} =& w_{o_{it}t} \times H(o_{it}, k, m) \end{aligned}$$

D.2 Labor Block Details

D.2.1 EM Algorithm Implementation Details

Inputs into the EM Algorithm are the following:

File Name	Description	Size
iDex	Key for *worker* for each i and t in panel	$NT \times 1$
tDex	Key for *calendar year* for each i and t in panel	$NT \times 1$
ptDex	Worker's observation within panel for each i and t	$NT \times 1$
field	Each worker's field	$N \times 1$
occupation	Each worker's occupation at time t	$NT \times 1$
logW	Each worker's log wage at time t	$NT \times 1$
zeta	Occupational characteristics	$\lvert\mathcal{O}\rvert \times Z$

These are the files read in by the Algorithm, however they are an incomplete list of what is needed for the Algorithm to run.

Step 1. Initialize the EM algorithm.

- Set types by k-means clustering on wages, ignoring occupations (as one possibility)
- Assign initial probabilities by,

$$q_{ik} = \begin{cases} 0.95 & \text{if clustered into } k \\ 0.05 & \text{if clustered into } k \end{cases}$$

- Construct the following sparse matrices *once*:
  - **XOT**: $NT \times |\mathcal{O}|$ matrix of occ-year dummies
  - **XOM**: $NT \times (M|\mathcal{O}| - M)$ matrix of occ-field dummies
  - For each k construct a matrix **XKZ** that is $NT \times (K-1) \times Z$ with all zeros for $k = 1$ (this is a normalization) and for $k \geq 2$:

$$\mathbf{XKZ}(k)(:, Z * (k-2) + Z) = \zeta_{o_{it},z}$$

  - Finally, construct the design matrix for each type:

$$\mathbf{XTEMP}(k) = [\mathbf{XOT}, \mathbf{XOM}, \mathbf{XKZ}(k)]$$

Note: In matlab this can be stored as a cell or multi-dimensional in array. In other languages, a multidimensional array. The notation here is Matlab's cell. This matrix does *not* change with iterations—only the weight matrix is updated.

- Estimate initial parameters for wages.
  (a) Preallocate the covariance matrix, **XtX**, to a $|\mathcal{O}| \times T + (M-1) \times |\mathcal{O}| + (K-1) \times Z$ square matrix of zeros.
  (b) Preallocate the covaraince matrix between Y and X, **XtY**, to a $|\mathcal{O}| \times T + (M-1) \times |\mathcal{O}| + (K-1) \times Z$ length vector.
  (c) For $k = 1, ..., K$...
    i. Generate the weight matrix,

$$W = diag(q_{ik})$$

    ii. Update the RHS covariance matrix:

$$\mathbf{XtX} = \mathbf{XtX} + \mathbf{XTEMP}(k)'W\mathbf{XTEMP}(k)/(NT)$$

    iii. Update the LHS covariance matrix

$$\mathbf{XtY} = \mathbf{XtY} + \mathbf{XTEMP}(k)'W[\log w]/(NT)$$

  (d) Estimate the Mincer regression parameters,

$$\beta^{Mincer} = (XtX)^{-1}XtY$$

  (e) Estimate residual variance
    – For each k,

$$r_{itk} = \log w_{it} - XTEMP(k)_{it} \times \beta^{Mincer}$$

    – Estimate variance:

$$\hat{\sigma}^2 = \frac{1}{NT}\sum_k \sum_i \sum_t r_{itk}^2 q_{ik}$$

- Estimate transitions:

$$\hat{\lambda}(o'|o,t,m,k) = \frac{\sum_{i:o_{it-1}=o \cap o_{it}=o' \cap i \in m} q_{ik}}{\sum_{i:o_{it-1}=o \cap i \in m} q_{ik}}$$

- Estimate initial conditions:

$$\hat{\lambda}^0(o|t,m,k) = \frac{\sum_{i:o_{it}=o \cap i \in m} q_{ik} \mathbf{1}(p_i^1 = t)}{\sum_{i:i \in m} q_{ik} \mathbf{1}(p_i^1 = t)},$$

  where p_i^1 is the first year the worker appears in the data.

Step 2. Initial value of EM objective, $J^{(0)} = 0$. **Begin E Step:**

Step 3. For iteration g, update the EM objective, J^g. For each k calculate the log likelihood, LL_k

$$LL_k = \sum_{i=1}^{N} q_{ik} \left[\sum_{t=1}^{T} \left(-r_{itk}^2/(2\sigma^2) - \log(\sigma^2)/2 + \log \hat{\lambda}(o_{it}|o_{it-1}, m_i, k) \right) + \log(\hat{\lambda}^0(o_{io}|t_0, m_i, k) \right]$$

And define,

$$LL = \sum_k LL_k$$

Finally add cdf term:

$$J = \sum_k LL_k + \sum q_k \log(q_k),$$

where $q_k = \frac{1}{N} \sum_i q_{ik}$.

Step 4. Update weights,

$$q_{ik} = \frac{\exp(LL_{ik}) q_k}{\sum_{k'} \exp(LL_{ik'}) q_{k'}}$$

Begin M Step:

Step 5. Update unconditional weights,

$$q_k = \frac{1}{N} \sum_i q_{ik}$$

Step 6. Update Mincer Parameters

(a) For $k = 1, ..., K$...

i. Generate the weight matrix,

$$W = diag(q_{ik})$$

ii. Update the RHS covariance matrix:

$$\mathbf{XtX} = \mathbf{XtX} + \mathbf{XTEMP}(k)'W\mathbf{XTEMP}(k)/(NT)$$

iii. Update the LHS covariance matrix

$$\mathbf{XtY} = \mathbf{XtY} + \mathbf{XTEMP}(k)'W[\log w]/(NT)$$

(b) Estimate the Mincer regression parameters,

$$\beta^{Mincer} = (XtX)^{-1}XtY$$

(c) Update residual variance

- For each k,

$$r_{itk} = \log w_{it} - XTEMP(k)_{it} \times \beta^{Mincer}$$

- Estimate variance:

$$\hat{\sigma}^2 = \frac{1}{NT}\sum_k \sum_i \sum_t r_{itk}^2 q_{ik}$$

Step 7. Update transitions:

$$\hat{\lambda}(o'|o,t,m,k) = \frac{\sum_{i:o_{it-1}=o \cap o_{it}=o' \cap i \in m} q_{ik}}{\sum_{i:o_{it-1}=o \cap i \in m} q_{ik}}$$

Step 8. Update initial conditions:

$$\hat{\lambda}^0(o|t,m,k) = \frac{\sum_{i:o_{it}=o \cap i \in m} q_{ik}\mathbf{1}(p_i^1 = t)}{\sum_{i:i \in m} q_{ik}\mathbf{1}(p_i^1 = t)},$$

Step 9. Step 8: If $|J^g - J^{g-1}| < tol$, where g is iteration, stop. Else, go to Step 2.

Outputs to be recovered from β^{Mincer}:

File Name	Description	Size
$\hat{\mu}_{ot}$	Skill Prices	$\|\mathcal{O}\| \times T$
$\hat{\alpha}_{om}$	Occ-Field FE	$\|\mathcal{O}\| \times M$
$\hat{\alpha}_{kz}$	Type-Task Productivity	$K \times Z$

Additional outputs to save:

File Name	Description	Size
$\hat{\lambda}_{oo}$	Transition matrix estimate	$\lvert\mathcal{O}\rvert \times \lvert\mathcal{O}\rvert \times T \times M \times K$
$\hat{Q}_K$	Matrix of type weights	$NT \times K$

D.2.2 Stage 2 of Labor Supply

To derive (11),

$$\log\left[\frac{\lambda_t(o'|o,k,m)}{\lambda_t(o|o,k,m)}\right] = -\frac{1}{\nu}\Big(-C_{oo'km} + w_{o't}H_{o'}(k,m) - w_{ot}H_o(k,m) + \beta\big(V_{t+1}(k,m,o') - V_{t+1}(k,m,o)\big)\Big),$$

and

$$\begin{aligned}\log\left[\frac{\lambda_{t+1}(o'|o',k,m)}{\lambda_{t+1}(o'|o,k,m)}\right] =& -\frac{1}{\nu}\Bigg(C_{oo'km} + w_{o't+1}H_{o'}(k,m) - w_{o't+1}H_{o'}(k,m)\\ &+\beta\big(V_{t+2}(k,m,o') - V_{t+2}(k,m,o')\big) + \big(V_{t+1}(k,m,o') - V_{t+1}(k,m,o)\big)\Bigg)\\ =& -\frac{C_{oo'km} + \big(V_{t+1}(k,m,o) - V_{t+1}(k,m,o')\big)}{\nu}\end{aligned}$$

Combining, we get,

$$\begin{aligned}\log\left[\frac{\lambda_t(o'|o,k,m)}{\lambda_t(o|o,k,m)}\right] + \beta\log\left[\frac{\lambda_{t+1}(o'|o't,k,m)}{\lambda_{t+1}(o'|o,k,m)}\right] =& -\frac{C_{oo'km}(1-\beta)}{\nu} +\\ &\frac{1}{\nu}\left(w_{o't}H_o(k,m) - w_{ot}H_o(k,m)\right) + \xi_{oo'kmt+1}\end{aligned},$$

In principle, to ensure positivity of the costs we apply an exponential transform. Specifically we estimate the following regression via unweighted NLLS:

$$\begin{aligned}\log\left(\frac{\hat{\lambda}_{oo'tmk}}{\hat{\lambda}_{ootmk}}\right) + \beta\log\left(\frac{\hat{\lambda}_{o'o't+1,mk}}{\hat{\lambda}_{oo't+1,mk}}\right) =& \exp\left(\gamma_{km} + \sum_z \gamma_z|\zeta_{o'z} - \zeta_{oz}|\right)\times(1-\beta)\tilde{\nu} -\\ &\tilde{\nu}\left(e^{\hat{\mu}_{o't}}\hat{H}(o',m,k) - e^{\hat{\mu}_{ot}}\hat{H}(o,m,k)\right)\end{aligned},$$

where $\tilde{\nu} = \nu^{-1}$ and $\hat{H}(o,m,k) = \exp(\alpha_{om} + \sum_z \alpha_{kz}\zeta_{zo})$.

D.2.3 Estimating Entry Costs

To derive (12),

$$\log\left[\frac{\lambda^e_t(o|k,m)}{\lambda^e_t(1|k,m)}\right] = -\frac{1}{\nu}\left(C^e_{okm} + w_{ot}H_o(k,m) - w_{1t}H_1(k,m) + \beta\Big(V_{t+1}(k,m,o) - V_{t+1}(k,m,1)\Big)\right)$$

Combining with $\log\left[\frac{\lambda_{t+1}(o'|o',k,m)}{\lambda_{t+1}(o'|o,k,m)}\right]$ and after some algebra, we obtain (12). In practice we take the mean across all occupations.

D.3 Details on Field Choice

D.3.1 Deriving the Recursive Representation

First we derive the transformation for the entry costs regression by expressing the entry value function in terms of future entry values and

$$\begin{aligned}
V_t^e(k,m) =& \nu \log\left(\sum_o \exp\left(\frac{-C^e_{okm} + w_{ot}H(o,k,m) + \beta V_{t+1}(o,k,m)}{\nu}\right)\right) \\
=& -C^e_{okm} + w_{ot}H(o,k,m) + \beta V_{t+1}(o,k,m) - \nu \log \lambda_t^e(o|k,m) \\
=& -C^e_{okm} + w_{ot}H(o,k,m) + \beta\left\{w_{o,t+1}H(o,k,m) + \beta V_{t+2}(o,k,m)\right\} - \nu \log \lambda_t^e(o|k,m) - \\
& \beta\nu \log \lambda_{t+1}(o|o,k,m) \\
=& -(1-\beta)C^e_{okm} + w_{ot}H(o,k,m) + \\
& \beta\left\{-C^e_{okm} + w_{o,t+1}H(o,k,m) + \beta V_{t+2}(o,k,m) - \nu \log \lambda^e_{t+1}(o|k,m)\right\} - \\
& \nu \log \lambda_t^e(o|k,m) - \beta\nu \log \frac{\lambda_{t+1}(o|o,k,m)}{\lambda^e_{t+1}(o|k,m)} \\
=& -(1-\beta)C^e_{okm} + w_{ot}H(o,k,m) + \beta V^e_{t+1}(k,m) - \nu \log \lambda_t^e(o,k,m) - \beta\nu \log \frac{\lambda_{t+1}(o|o,k,m)}{\lambda^e_{t+1}(o|k,m)}.
\end{aligned}$$

Group observable or estimated terms—wages, costs, and transition probabilties—into $V_t^F(k,m,s)$, where dependence on m and s is suppressed for convenience. Thus we have,

$$V_t^e(k,m) = V_t^F(k,m,s) + \beta V^e_{t+1}(k,m).$$

Averaging over k we have,

$$V_t^e(m,s) = \left[\sum_k \pi(k|m,s)V_t^F(k,m,s)\right] + \beta V_t^e(m,s)$$

Hence,

$$\begin{aligned}\tilde{U}_t(l,m,s) =& \vartheta_{lm} + \theta_l V_t^e(m,s) \\ =& (1-\beta)\vartheta_{lm} + \theta_l \left[\sum_k \pi(k|m,s) V_t^F(k,m,s)\right] + \beta\left[\vartheta_{lm} + \theta_l V_{t+1}^e(m,s)\right] \\ =& (1-\beta)\vartheta_{lm} + \theta_l \left[\sum_k \pi(k|m,s) V_t^F(k,m,s)\right] + \beta\tilde{U}_{t+1}(l,m,s).\end{aligned}$$

Iterating forward yields the expression for arbitrary t, and we fix $\tilde{U}_T$ as the terminal value for all periods.

D.3.2 Likelihood Conditional on l

The FOCs are given by,

$$\frac{\partial LL}{\partial \vartheta_{lm}} = \sum_t (1-\beta^{T-6-t}) \sum_i q_{itl} \sum_r \left[\mathbf{1}(m_{r_i}=m) - \frac{N_{irm}^v \exp(U_{ml,t}(s_i))}{\sum_{m'\in\mathcal{C}_i} N_{irm'}^v \exp(U_{m'l,t}(s_i))}\right] \tag{26}$$

$$\frac{\partial LL}{\partial \theta_l} = \sum_t \sum_i q_{itl} \sum_r \left[XV_{t,T-6}^F(m_{r_i},s_i) - \frac{\sum_{m\in\mathcal{C}_i} XV_{t,T-6}^F(m,s_i) \times N_{irm}^v \exp(U_{ml,t}(s_i))}{\sum_{m\in\mathcal{C}_i} N_{irm'}^v \exp(U_{ml,t}(s_i))}\right] \tag{27}$$

$$\frac{\partial LL}{\partial v} = \sum_t \sum_l \sum_i q_{itl} \sum_r \left[\log N_{irm} - \frac{\sum_{m\in\mathcal{C}_i} \log N_{irm} \times N_{irm}^v \exp(U_{ml,t}(s_i))}{\sum_{m\in\mathcal{C}_i} N_{irm'}^v \exp(U_{ml,t}(s_i))}\right]. \tag{28}$$

These first order conditions are intuitive.[24] The first says that ϑ_{lm} ensures that the time-averaged choice probabilities for each field by type l match the empirical shares. The second and third says that the sensitivity of individuals with respect to labor market fundamentals and the number of programs ensures that the model-implied average labor market outcome equals and log number of programs equals their empirical means. The discounting also has an intuitive appeal from a identification standpoint. The term $(1-\beta^{T-t-6})$ is larger the further away t is from the terminal period. Hence, the likelihood places more weight on earlier periods, where there is a longer history of observed discounted wage differentials, than on later periods. On a final note, the residuals in the $U_{lmt}(s)$ choice utilities will reflect measurement error and the expectational error that arise from using observed labor market outcomes in lieu of expectations.

[24] In the implementation, to keep parameters bounded from 0, we parameterize them as $\theta = \exp(\vartheta)$ and search over ϑ. The first order conditions only change in that they are multiplied by the level of the current guess.

D.3.3 Computational Algorithm Details

D.4 Parameter Adjustments for Levels 1 + 2

After estimating the model for levels 1 and 2, we have estimates of μ_{ot}—in principle skill prices—for each group. Here we denote these by μ_{otg}, where g is for skill group. While it is the case that the average level of wages for each worker should be correctly matched because we normalize the *population* mean wages to 1, for each group we can only estimate a *relative* μ. That is to say, *something* must be normalized in the model. To see why, briefly, consider observing wages over time for different occupations and workers of different types (and ignore other heterogeneity), but condition on group. In this case, one has,

$$w_{otgk} = \mu_{otg} + \beta_{okg}. \tag{29}$$

This is essentially the decomposition we perform. This system is rank deficient. To see why is easy—one can add a constant to all the β and subtract this from all the μ and solve the system. To this end, we have imposed a normalization in estimation that sets $\beta_{11_g} = 0 \forall g$. In reality, however, we do not want to do this because it is *not* the case that $\beta_{111} = \beta_{113}$. We need to renormalize. However, there is a further issue, which is that μ_{ot} are supposed to be [relative] skill prices, and we want these to match. I.e., we would like for $\mu_{ot1} = \mu_{ot2}$ for all groups.

We solve these issues with a simple renormalization:

$$\begin{aligned} w_{otgk} =& \mu_{otg} + \beta_{okg} \\ =& \mu_{ot3} + \underbrace{\left[\beta_{okg} + \mu_{otg} - \mu_{ot3}\right]}_{\tilde{\beta}_{okg}}. \end{aligned}$$

Another way to put this is that we want things measured in the same units. This normalization ensures that.

Unfortunately, this leads to time-varying β's. In order to avoid this, we only do the normalization for the year that we initialize from for counterfactuals. And so in practice we set,

$$\tilde{\alpha}_{omg} = \alpha_{omg} + \mu_{o0g} - \mu_{o03}$$

for all skill groups.

D.5 Bootstrapping

We calculate standard errors using the Bayesian block bootstrap, drawing weights for different stages. For the first stage of the labor block, the EM algorithm, by arguments in Arcidiacono and Miller (2011), is $\sqrt{N}$ consistent, so long as conditional on year, individuals remaining errors are independent. While Arcidiacono and Miller (2011) provide analytical covariance matrices, we use a bootstrap procedure so that in the second stage, where we bootstrap instead over year, we can capture first stage uncertainty. Specifically in the first stage we employ a Bayesian block bootstrap. For each bootstrap iteration, we draw from a Dirichlet distribution of length N with $\alpha_j = 1 \forall j$. We multiply this by N. This will be our vector of weights, w_i. Notice that this samples uniformly over individuals regardless of their total length in the sample, so that the total size of the sample (which we often abusively write as NT can change). Given the weights, we re-estimate the first stage where the likelihood is now weighted by w_i.

In the first stage, we estimate $\hat{\pi}$, which are the *true* transition rates. However, these transitions rates are based on *expectations* about the future. Thus, they encode information about $E_t V_{t+1}$. These expectations are not separately identified from other parameters, but as described in the main text, we plug in *future* transition rates and exploit finite dependence to estimate structural parameters in a linear regression. In *this* regression, that expectational errors may be correlated across units—now the occupation—within a year. Ignoring this clustering and re-estimating this stage using the bootstrapped values from the first stage would only capture uncertainty stemming from sampling error, but miss the within-year correlation. To this end, in this stage we also cluster by year. We now block bootstrap in the time dimension. Specifically, note that we may write the regression as depending on three periods $(t-1, t, t+1)$. We draw weights from a $T-2$ Dirichlet for the parameters in this estimation.

We employ the same Bayesian bootstrapping procedure (with newly drawn weights) for the Field Choice estimation. As a final caveat, we employ a minimum burn-in for the EM algorithm. Because of this, mechanically, the procedure runs for longer for the bootstrap occasionally (when the initial guess is, e.g., very good). Because of this, we recenter estimates at the parameter estimates. For a given parameter, $\hat{x}$, we set $\hat{x}BS^{Recentered} = \hat{x}BS - \overline{\hat{x}^{BS}} + \hat{x}^0$, where $\hat{x}^0$ is the initial estimate. In practice this has very little impact.

E Results Appendix

Approximating V^e

Given the Gumbel shocks, we derive above the following expression:

$$V_t^e(k,m) = -(1-\beta)C_{okm}^e + w_{ot}H(o,k,m) + \beta V_{t+1}^e(k,m) - \nu \log \lambda_t^e(o,k,m) - \beta\nu \log \frac{\lambda_{t+1}(o|o,k,m)}{\lambda_{t+1}^e(o|k,m)},$$

where λ are choice probabilities, and o is any choice of base occupation. Here we begin our approximation. Assuming that $V_t^e \approx V_{t+1}^e$ (i.e, that we are in a stationary environment),

$$V_t^e(k,m) \approx \frac{-(1-\beta)C_{okm}^e + w_{ot}H(o,k,m) - \nu \log \lambda_t^e(o,k,m) - \beta\nu \log \frac{\lambda_{t+1}(o|o,k,m)}{\lambda_{t+1}^e(o|k,m)}}{1-\beta}.$$

The issue with the approximation is that we cannot estimate C_{okm}^e in levels, only in differences. To this end, we need to take stand on this value. We assume that for retail, $C_{okm}^e \approx .25$, which is one quarter of the average annual income in the economy (which has been normalized to 1). Given this normalization, we can calculate the above for every value of o and every period t in the economy, and take the average. And so, calling the right hand side, $\tilde{V}_{ot}^e(k,m)$, where the $\tilde{V}$ refers to the approximated value evaluated at o, we defne,

$$V^{e,approx}(k,m) = \frac{1}{|\mathcal{O}| \times T} \sum_o \sum_t \tilde{V}_{ot}^e(k,m).$$

We then have,

$$V^{e,approx}(m,s) = \sum_k \pi(k|m,s) V^{e,approx}(k,m).$$

F Labor Demand Model

F.1 Model details

Given wages and prices of imported inputs, the profit maximizing choice of input quantity and the zero profit condition yield demand for occupation o in industry s:

$$H^P_{sot} = A_{st} \times \frac{\beta^P_{so}}{W_{ot}} \times A^P_s \times Q_{st} \times \left(P^D_{st}\right)^{\alpha} \times \left(P^P_{st}\right)^{\rho-\alpha} \times \left(W^P_{st}\right)^{-\rho}, \quad \forall o \in \mathcal{O}^P; \tag{30}$$

$$H^N_{sot} = A_{st} \times \frac{\beta^N_{so}}{W_{ot}} \times A^N_s \times Q_{st} \times \left(P^D_{st}\right)^{\alpha} \times \left(W^N_{st}\right)^{-\alpha}, \quad \forall o \in \mathcal{O}^N; \tag{31}$$

and demand for imported goods is

$$Q^M_{st} = A_{st} \times A^M_s \times Q_{st} \times \left(P^D_{st}\right)^{\alpha} \times \left(P^P_{st}\right)^{\rho-\alpha} \times \left(P^M_{st}\right)^{-\rho}, \tag{32}$$

where

$$Q_{st} = A^F_s \left(P^D_{st}\right)^{-\eta} + \mu_s E_t \times \left(\frac{P_{st}}{P_t}\right)^{1-\sigma} \times \frac{\left(P^D_{st}\right)^{-\eta}}{\left(P^D_{st}\right)^{1-\eta} + \left(P^F_{st}\right)^{1-\eta}}; \tag{33}$$

$$W^f_{st} = \prod_{o \in \mathcal{O}^f} w^{\beta_{so}}_{ot}, \quad f = \{P, N\}; \tag{34}$$

$$P^P_{st} = \left[A^P_s \left(W^P_{st}\right)^{1-\rho} + A^M_s \left(P^M_{st}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}; \tag{35}$$

$$P^D_{st} = \left[A^N_s \left(W^N_{st}\right)^{1-\alpha} + \left(P^P_{st}\right)^{1-\alpha}\right]^{\frac{1}{1-\alpha}} \tag{36}$$

$$P_{st} = A^{-1}_{st} \left[\left(P^D_{st}\right)^{1-\eta} + \left(P^F_{st}\right)^{1-\eta}\right]^{\frac{1}{1-\eta}} \tag{37}$$

$$P_t = \left[\sum_s \mu_s \left(P_{st}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{38}$$

$$E_t = \sum_o w_{ot} \left(\sum_s L_{sot}\right) \tag{39}$$

Finally, occupation wages, w_{ot}, adjust to clear the labor market, for $k = \{P, N\}$

$$H_{ot} = \sum_s H^k_{sot} \tag{40}$$

F.2 Demand Side Parameters Calibration

Table 14 summarizes the full set of parameters to be calibrated in order to close the model. All production coefficients are estimated as the time-averaged expenditure shares.

Table 14: Labor Demand Parameters and Variables

Parameters	Meaning	Value/Data Source
β_{so}^P, β_{so}^N	Labor inputs coefficients	IDA
P_{st}^M	Offshored input prices	UHDI, Broda and Weinstein (2006)
P_{st}^F	Imported good prices	UHDI, Broda and Weinstein (2006)
w_{ot}	Occupation wage	IDA
ρ, α	EOS in production	Doraszelski and Jaumandreu (2018)
σ	EOS in final good	Atalay (2017)
η	EOS btwn domestic and foreign	Simonovska and Waugh (2014)

Total wage bills in occupation-industry pairs identify the labor input coefficients. Danish customs data provides provide and quantity level information on goods sold to Denmark by each trading partner, which can be fed into the procedure of Broda and Weinstein (2006) to calculate the price indices for offshored inputs and imported varieties.[25] We set the trade elasticity as 4, following Simonovska and Waugh (2014). We obtain raw wage data from the Integrated Database for Labour Market Research (IDA) at Statistics Denmark. Occupation wages are computed by first running a mincerian regression and then divide the residuals by the computed human capital using (3).

G Hat Algebra

G.1 Trade shocks

Consider shocks $\{P_{st}^{F,0}\} \to \{P_{st}^{F,1}\}$, $\{P_{st}^{M,0}\} \to \{P_{st}^{M,1}\}$. We use 0 superscripts to denote the initial steady state, and 1 superscripts to denote the final steady state.

Start from the initial steady state: $\{w_o^0\}$. We assume q_l and $\pi(k|m,s) = \frac{\exp(\delta_{km}+\beta_k s)}{\sum_{k'}\exp(\delta_{k'm}+\beta_{k'}s)}$ are constant. Denote relative changes in variable x by $\hat{x} = \frac{x^1}{x^0}$.

Step 1. Guess $\{w_o^1\}$

Step 2. Compute changes in prices

$$\hat{W}_s^f = \prod_{o\in\mathcal{O}^f} \hat{W}_o^{\beta_{so}}; \quad f = P, N$$

[25] We follow the approach in Feenstra (1994) to correct for changes in product varieties across time.

$$\hat{P}_s^P = \left(S_{st}^P \left(\hat{W}_s^P \right)^{1-\rho} + (1 - S_{st}^P) \left(\hat{P}_s^M \right)^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

$$\hat{P}_s^D = \frac{1}{\hat{A}_s} \left[S_{st}^N \left(\hat{W}_s^N \right)^{1-\alpha} + (1 - S_{st}^N) \left(\hat{P}_s^P \right)^{\frac{1-\alpha}{1-\rho}} \right]^{\frac{1}{1-\alpha}}$$

$$\hat{P}_s = \left[(1 - S_{st}^I) \left(\hat{P}_s^D \right)^{1-\eta} + S_{st}^I \left(\hat{P}_s^F \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}$$

$$\hat{P} = \left(\sum_s S_{st} \left(\hat{P}_s \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$P = \hat{P} P_0$$

Step 3. Compute Bellman Equations by iterating on the following until convergence:

$$V^1(k, m, o) = \nu \log \left(\sum_{o'} \exp \left(\frac{-C_{oo'km} + w_{o'}^1 H_{o'}(k, m)/P + \beta V^1(k, m, o')}{\nu} \right) \right)$$

Step 4. Using $V^1(k, m, o)$, solve for the inclusive value of being type (k, m):

$$V^{e,1}(k, m) = \nu \log \sum_{o'} \exp \left(\frac{-C_{o,km}^e + w_{o'}^1 H_{o'}(k, m)/P + \beta V^1(k, m, o')}{\nu} \right)$$

Step 5. Compute the NPV for each field m

$$V^1(m|s) = \sum_k V^{e,1}(k, m) \pi(k|m, s).$$

Step 6. Compute $\{L_{km}^1\}$ by Monte Carlo simulation

(a) **Outside the loop** Simulate N students. For each student i draw a test score s_i, gender f_i, high school background, x_i, and type, l_i from the joint distribution $Q(l, s, f, x)$.

(a) For each individual i, generate a test score s_i for i

(b) Generate a type for i from the conditional CDF of the type distribution given s_i:

$$l_i = 1 + \sum_l \mathbf{1}(e > Q_l(s_i)),$$

where e is a uniform random variable.

(c) For $r = 1, ..., R$ calculate the ranked list of choices for individual i:

- Calculate the conditional CDF of choices for field:

$$\tilde{Q}_{ir}^{F}(m) = \frac{\sum_{m'=1\setminus m_{-r}}^{m} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V^1(m|s)\right)}{\sum_{m'=1\setminus m_{-r}}^{M} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V^1(m|s)\right)}$$

- Generate R draws from this distribution:

$$m_{r_i} = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_{ir}^{F}(m)),$$

for a uniform random variable.

(d) Assign the agent sequentially to a field in their choice set based on acceptance probabilities (conditional on s). If the agent is not admitted to the first $R-1$ choices, then they are automatically allocated to the R-th choice.

(e) Determine the agents type:

$$k_i = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_{i}^{K}(m, s)),$$

where Q_k is type CDF for agent i with field m and score s, and e is a uniform random variable.

(f) Calculate the share of individuals belong to type (k, m), scale up to $\bar{L}$.

Step 7. Using $V^1(k, m, o)$ solve for λ_{km}^1 and $\lambda^{e,1}(o|k, m)$

$$\lambda^1(o'|o, k, m) = \frac{\exp\left(\frac{-C_{oo'km} + w_{o't} H_{o'}(k,m)/P + \beta V^1(k,m,o')}{\nu}\right)}{\sum_{o''} \exp\left(\frac{-C_{oo''km} + w_{o''t} H_{o''}(k,m) + \beta V^1(k,m,o'')}{\nu}\right)}$$

$$\lambda^{e,1}(o|k, m) = \frac{\exp\left(\frac{-C_{o,km} + w_{ot} H_{o}(k,m)/P + \beta V^1(k,m,o)}{\nu}\right)}{\sum_{o'} \exp\left(\frac{-C_{o,km} + w_{o't} H_{o'}(k,m) + \beta V^1(k,m,o')}{\nu}\right)}$$

Step 8. Compute $\hat{H}_o^{supply}$

(a) Consider the flow equation:

$$L_{omk} = (1-\delta)\sum_{o'} \lambda^1(o|o',k,m)L_{o'mk} + \delta\lambda^{e,1}(o|k,m)L^e_{mk}$$

This implies,

$$L_{\cdot,mk} = \delta\left(I_{NOCC} - (1-\delta)\lambda^1(\cdot|o',k,m)'\right)^{-1} \times \left(\lambda^{e,1}(\cdot|k,m)L^e_{mk}\right)$$

(b) Construct H^1_{omk}:

$$H_{omk} = L_{omk} \times H(o,m,k)$$

(c) Compute changes in human capital supply across occupations

$$\hat{H}^{supply}_o = \sum_{k,m} H^1_{omk}/H^0_o$$

Step 9. Calculate aggregate expenditure (and income)

$$\hat{E} = \sum_o S_{ot}\hat{H}_o\hat{W}_o$$

Step 10. Calculate changes in domestic output

$$\hat{Q}_s = S^x_{st}\left(\hat{P}^D_s\right)^{-\eta} + (1-S^x_{st})\hat{E}\hat{P}^{\eta-\sigma}_s\hat{P}^{\sigma-1}\left(\hat{P}^D_s\right)^{-\eta}$$

Step 11. Calculate a new guess for wages, by enforcing labor market clearing and output market clearing. Choose $\hat{W}_o$ such that

$$\begin{aligned}\hat{H}_o &= \sum_s \frac{H_{sot}}{H_{ot}}\hat{H}_{so} \\ &= \sum_s \frac{H_{sot}}{H_{ot}}\frac{1}{\hat{W}_o}\hat{Q}_s\left(\hat{P}^D_s\right)^{\alpha}\left(\hat{W}^N_s\right)^{-\alpha},\end{aligned}$$

for all $o \in \mathcal{O}^N$; and

$$\begin{aligned}\hat{H}_o &= \sum_s \frac{H_{sot}}{H_{ot}} \hat{H}_{so} \\ &= \sum_s \frac{H_{sot}}{H_{ot}} \frac{1}{\hat{W}_o} \hat{Q}_s \left(\hat{P}_s^D\right)^{\alpha} \left(\hat{P}_s^P\right)^{\rho-\alpha} \left(\hat{W}_s^P\right)^{-\rho},\end{aligned}$$

for all $o \in \mathcal{O}^P$.

Step 12. Update $\hat{W}_o$ and iterate from Step 1 again.

After the algorithm completes, one can update shares and quantities:

$$\begin{aligned}
S_{s1}^P =& A^P (W_{s1}^P)^{1-\rho} / (P_1^P)^{1-\rho} = S_{s0}^P \times \left(\frac{\hat{W}_s^P}{\hat{P}^P}\right)^{1-\rho} \\
S_{s1}^N =& A^P (W_{s1}^N)^{1-\alpha} / (P_1^D)^{1-\alpha} = S_{s0}^N \times \left(\frac{\hat{W}_s^N}{\hat{P}^D}\right)^{1-\alpha} \\
S_{s1}^I =& 1 - A (P_{s1}^D)^{1-\eta} / (P_1^D)^{1-\eta} = 1 - (1 - S_{s0}^I) \times \left(\frac{\hat{P}_s^D}{\hat{P}_s}\right)^{1-\eta} \\
S_{s1}^S =& \mu_s P_{s1}^{1-\sigma} / P_0^{1-\sigma} = S_{s0}^S \times \left(\frac{\hat{P}_s}{\hat{P}}\right)^{1-\sigma} \\
S_{s1}^x =& \frac{A_s^F (P_{s1}^D)^{1-\eta}}{Q_{s1} P_{s1}^D} = S_{s0}^x \times \frac{(\hat{P}_s^D)^{-\eta}}{\hat{Q}_s} \\
S_{o1} =& \frac{w_{o1} H_{o1}}{E_1} = \frac{\hat{w}_o \hat{H}_o}{\hat{E}_h} \\
\Pi_{os1}^P =& \frac{(A_s^P)^{1-\rho} \beta_{so} Q_{s1} (P_{s1}^D)^{\alpha} (P_{s1}^P)^{\rho-\alpha} (W_{s1}^P)^{1-\rho} / W_{o1}}{H_{o1}} = \frac{\hat{\beta}_{os} \hat{Q}_s (\hat{P}_s^D)^{\alpha} (\hat{P}_s^P)^{\rho-\alpha} (\hat{W}_s^P)^{-\rho} / \hat{W}_o}{\hat{H}_o} \Pi_{os0}^P \\
\Pi_{os1}^N =& \frac{(A_s^N)^{1-\alpha} \beta_{so} Q_{s1} (P_{s1}^D)^{\alpha} (W_{s1}^N)^{-\alpha} / W_{o1}}{H_{o1}} = \frac{\hat{\beta}_{os} \hat{Q}_s (\hat{P}_s^D)^{\alpha} (\hat{W}_s^N)^{-\alpha} / \hat{W}_o}{\hat{H}_o} \Pi_{os0}^N
\end{aligned}$$

G.2 AI Shock

G.2.1 Defining the AI Shock

Suppose there is an external measure that predicts declines in *relative* labor demand. We do not model aggregate reductions in labor demand. Specifically we assume that there is a measure z_o such that within each sector we assume that at *original prices*,

$$\frac{w_{so}H'_{so}}{w_sH'_s} \propto z_o.$$

Let $\tilde{z}_{so}$ be a sector-occupation specific shock that ensures the wage share is left unchanged:

$$\tilde{z}_{so} = z_o \times \frac{1}{\frac{w_s^N L_s^N}{w_S L_S}\left(\sum_{o\in N}\beta_{so}z_o\right) + \frac{w_s^P L_s^P}{w_S L_S}\left(\sum_{o\in P}\beta_{so}z_o\right)}$$

We want to solve for $\tilde{\beta}$ and $\tilde{A}$ that matches this shock. We do this in two stages, first we look at non-production workers, then production workers. But the strategy is ultimately identical.

For non-production workers, the share of the wage bill is given by,

$$\frac{w_oL_{so}}{w_sL_s} = \beta_{so}^N A_{sN}\left(\frac{w_s^N}{P_s^D}\right)^{1-\alpha}.$$

So we need to solve the following:

$$\tilde{\beta}_{so}^N \tilde{A}_{sN}\left(\frac{w_s^N}{P_s^D}\right)^{1-\alpha} = \tilde{z}_{so}\beta_{so}^N A_{sN}\left(\frac{w_s^N}{P_s^D}\right)^{1-\alpha},$$

subject to the restriction that $\sum_{o\in N}\tilde{\beta}_{so}^N = 1$. Dividing out the price term and summing we have,

$$\tilde{A}_{sN} = A_{sN}\sum_o \tilde{z}_{so}\beta_o^N.$$

Hence,

$$\hat{A}_{sN} = \sum_o \tilde{z}_{so}\beta_{so}^N.$$

We also have,

$$\tilde{\beta}_{so}^N = \frac{\tilde{z}_{so}\beta_{so}^N}{\sum_{o'}\tilde{z}_{so'}\beta_{so'}^N}.$$

For production workers, there is a similar common price term that ultimately cancels out, so we

have the same expressions:

$$\hat{A}_{sP} = \sum_o \tilde{z}_{so} \beta^P_{so}$$
$$\tilde{\beta}^P_{so} = \frac{\tilde{z}_{so} \beta^P_{so}}{\sum_{o'} \tilde{z}_{so'} \beta^P_{so'}}$$

G.2.2 Modified Algorithm

Consider shocks $\{P^{F,0}_{st}\} \to \{P^{F,1}_{st}\}$, $\{P^{M,0}_{st}\} \to \{P^{M,1}_{st}\}$ as well as $\beta_{so0} \to \beta_{so1}$. We use 0 superscripts to denote the initial steady state, and 1 superscripts to denote the final steady state.

Start from the initial steady state: $\{w^0_o\}$. We assume q_l and $\pi(k|m,s) = \frac{\exp(\delta_{km} + \beta_k s)}{\sum_{k'} \exp(\delta_{k'm} + \beta_{k'} s)}$ are constant. Denote relative changes in variable x by $\hat{x} = \frac{x^1}{x^0}$.

Step 1. Guess $\{w^1_o\}$

Step 2. Compute changes in prices

$$\hat{W}^f_s = \frac{\prod_{o \in \mathcal{O}^f} W^{\beta_{so1}}_{o1}}{\prod_{o \in \mathcal{O}^f} W^{\beta_{so0}}_{o0}}; \quad f = P, N$$

$$\hat{P}^P_s = \left(S^P_{st} \hat{A}_{sP} \left(\hat{W}^P_s \right)^{1-\rho} + (1 - S^P_{st}) \left(\hat{P}^M_s \right)^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

$$\hat{P}^D_s = \frac{1}{\hat{A}_s} \left[S^N_{st} \hat{A}_{sN} \left(\hat{W}^N_s \right)^{1-\alpha} + (1 - S^N_{st}) \left(\hat{P}^P_s \right)^{\frac{1-\alpha}{1-\rho}} \right]^{\frac{1}{1-\alpha}}$$

$$\hat{P}_s = \left[(1 - S^I_{st}) \left(\hat{P}^D_s \right)^{1-\eta} + S^I_{st} \left(\hat{P}^F_s \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}$$

$$\hat{P} = \left(\sum_s S_{st} \left(\hat{P}_s \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$P = \hat{P} P_0$$

Step 3. Compute Bellman Equations by iterating on the following until convergence:

$$V^1(k,m,o) = \nu \log \left(\sum_{o'} \exp \left(\frac{-C_{oo'km} + w^1_{o'} H_{o'}(k,m)/P + \beta V^1(k,m,o')}{\nu} \right) \right)$$

Step 4. Using $V^1(k,m,o)$, solve for the inclusive value of being type (k,m):

$$V^{e,1}(k,m) = \nu \log \sum_{o'} \exp\left(\frac{-C^e_{o,km} + w^1_{o'} H_{o'}(k,m)/P + \beta V^1(k,m,o')}{\nu}\right)$$

Step 5. Compute the NPV for each field m

$$V^1(m|s) = \sum_k V^{e,1}(k,m)\pi(k|m,s).$$

Step 6. Compute $\{L^1_{km}\}$ by Monte Carlo simulation

(a) For each individual i, generate a test score s_i for i

(b) Generate a type for i from the conditional CDF of the type distribution given s_i:

$$l_i = 1 + \sum_l \mathbf{1}(e > Q_l(s_i)),$$

where e is a uniform random variable.

(c) For $r = 1, ..., R$ calculate the ranked list of choices for individual i:

- Calculate the conditional CDF of choices for field:

$$\tilde{Q}^F_{ir}(m) = \frac{\sum_{m'=1\setminus m_{-r}}^{m} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V^1(m|s)\right)}{\sum_{m'=1\setminus m_{-r}}^{M} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V^1(m|s)\right)}$$

- Generate R draws from this distribution:

$$m_{r_i} = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}^F_{ir}(m)),$$

for a uniform random variable.

(d) Assign the agent sequentially to a field in their choice set based on acceptance probabilities (conditional on s). If the agent is not admitted to the first $R-1$ choices, then they are automatically allocated to the R-th choice.

(e) Determine the agents type:

$$k_i = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}^K_i(m,s)),$$

where Q_k is type CDF for agent i with field m and score s, and e is a uniform random variable.

(f) Calculate the share of individuals belong to type (k, m), scale up to $\bar{L}$.

Step 7. Using $V^1(k, m, o)$ solve for λ^1_{km} and $\lambda^{e,1}(o|k, m)$

$$\lambda^1(o'|o, k, m) = \frac{\exp\left(\frac{-C_{oo'km} + w_{o't} H_{o'}(k,m)/P + \beta V^1(k,m,o')}{\nu}\right)}{\sum_{o''} \exp\left(\frac{-C_{oo''km} + w_{o''t} H_{o''}(k,m) + \beta V^1(k,m,o'')}{\nu}\right)}$$

$$\lambda^{e,1}(o|k, m) = \frac{\exp\left(\frac{-C_{o,km} + w_{ot} H_o(k,m)/P + \beta V^1(k,m,o)}{\nu}\right)}{\sum_{o'} \exp\left(\frac{-C_{o,km} + w_{o't} H_{o'}(k,m) + \beta V^1(k,m,o')}{\nu}\right)}$$

Step 8. Compute $\hat{H}_o^{supply}$

(a) Consider the flow equation:

$$L_{omk} = (1 - \delta) \sum_{o'} \lambda^1(o|o', k, m) L_{o'mk} + \delta \lambda^{e,1}(o|k, m) L^e_{mk}$$

This implies,

$$L_{\cdot,mk} = \delta \left(I_{NOCC} - (1 - \delta)\lambda^1(\cdot|o', k, m)'\right)^{-1} \times \left(\lambda^{e,1}(\cdot|k, m) L^e_{mk}\right)$$

(b) Construct H^1_{omk}:

$$H_{omk} = L_{omk} \times H(o, m, k)$$

(c) Compute changes in human capital supply across occupations

$$\hat{H}_o^{supply} = \sum_{k,m} H^1_{omk} / H^0_o$$

Step 9. Calculate aggregate expenditure (and income)

$$\hat{E} = \sum_o S_{ot} \hat{H}_o \hat{W}_o$$

Step 10. Calculate changes in domestic output

$$\hat{Q}_s = S_{st}^x \left(\hat{P}_s^D\right)^{-\eta} + (1 - S_{st}^x)\hat{E}\hat{P}_s^{\eta-\sigma}\hat{P}^{\sigma-1}\left(\hat{P}_s^D\right)^{-\eta}$$

Step 11. Calculate a new guess for wages, by enforcing labor market clearing and output market clearing. Choose $\hat{W}_o$ such that

$$\begin{aligned}\hat{H}_o &= \sum_s \frac{H_{sot}}{H_{ot}}\hat{H}_{so} \\ &= \sum_s \frac{H_{sot}}{H_{ot}}\frac{1}{\hat{W}_o}\hat{Q}_s\left(\hat{P}_s^D\right)^{\alpha}\left(\hat{W}_s^N\right)^{-\alpha}\hat{\beta}_{so}^N\hat{A}_{sN},\end{aligned}$$

for all $o \in \mathcal{O}^N$; and

$$\begin{aligned}\hat{H}_o &= \sum_s \frac{H_{sot}}{H_{ot}}\hat{H}_{so} \\ &= \sum_s \frac{H_{sot}}{H_{ot}}\frac{1}{\hat{W}_o}\hat{Q}_s\left(\hat{P}_s^D\right)^{\alpha}\left(\hat{P}_s^P\right)^{\rho-\alpha}\left(\hat{W}_s^P\right)^{-\rho}\hat{\beta}_{so}^P\hat{A}_{sP},\end{aligned}$$

for all $o \in \mathcal{O}^P$.

Step 12. Update $\hat{W}_o$ and iterate from Step 1 again.

After the algorithm completes, one can update shares and quantities

$$
\begin{aligned}
S^P_{s1} =& A^P (W^P_{s1})^{1-\rho} / (P^P_1)^{1-\rho} = S^P_{s0} \times \left(\frac{\hat{W}^P_s}{\hat{P}^P} \right)^{1-\rho} \\
S^N_{s1} =& A^P (W^N_{s1})^{1-\alpha} / (P^D_1)^{1-\alpha} = S^N_{s0} \times \left(\frac{\hat{W}^N_s}{\hat{P}^D} \right)^{1-\alpha} \\
S^I_{s1} =& 1 - A (P^D_{s1})^{1-\eta} / (P^D_1)^{1-\eta} = 1 - (1 - S^I_{s0}) \times \left(\frac{\hat{P}^D_s}{\hat{P}_s} \right)^{1-\eta} \\
S^S_{s1} =& \mu_s P^{1-\sigma}_{s1} / P^{1-\sigma}_0 = S^S_{s0} \times \left(\frac{\hat{P}_s}{\hat{P}} \right)^{1-\sigma} \\
S^x_{s1} =& \frac{A^F_s (P^D_{s1})^{1-\eta}}{Q_{s1} P^D_{s1}} = S^x_{s0} \times \frac{(\hat{P}^D_s)^{-\eta}}{\hat{Q}_s} \\
S_{o1} =& \frac{w_{o1} H_{o1}}{E_1} = \frac{\hat{w}_o \hat{H}_o}{\hat{E}_h} \\
\Pi^P_{os1} =& \frac{(A^P_s)^{1-\rho} \beta_{so} Q_{s1} (P^D_{s1})^{\alpha} (P^P_{s1})^{\rho-\alpha} (W^P_{s1})^{1-\rho} / W_{o1}}{H_{o1}} = \frac{\hat{\beta}_{os} \hat{Q}_s (\hat{P}^D_s)^{\alpha} (\hat{P}^P_s)^{\rho-\alpha} (\hat{W}^P_s)^{-\rho} / \hat{W}_o}{\hat{H}_o} \Pi^P_{os0} \\
\Pi^N_{os1} =& \frac{(A^N_s)^{1-\alpha} \beta_{so} Q_{s1} (P^D_{s1})^{\alpha} (W^N_{s1})^{-\alpha} / W_{o1}}{H_{o1}} = \frac{\hat{\beta}_{os} \hat{Q}_s (\hat{P}^D_s)^{\alpha} (\hat{W}^N_s)^{-\alpha} / \hat{W}_o}{\hat{H}_o} \Pi^N_{os0}
\end{aligned}
$$

G.3 Transition Dynamics

Step 1. Calculate final steady state

Step 2. Guess a path of wages $\{w_{ot}\}_{o,t=1,T-1}$

Step 3. Solve for the path of prices using the hat algebra:

$$\hat{W}_{st}^{f} = \frac{\prod_{o \in \mathcal{O}^f} W_{ot}^{\beta_{sot}}}{\prod_{o \in \mathcal{O}^f} W_{o0}^{\beta_{so0}}}; \quad f = P, N$$

$$\hat{P}_{st}^{P} = \left(S_{s0}^{P} \hat{A}_{stP} \left(\hat{W}_{st}^{P} \right)^{1-\rho} + (1 - S_{s0}^{P}) \left(\hat{P}_{st}^{M} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

$$\hat{P}_{st}^{D} = \frac{1}{\hat{A}_{st}} \left[S_{s0}^{N} \hat{A}_{stN} \left(\hat{W}_{st}^{N} \right)^{1-\alpha} + (1 - S_{s0}^{N}) \left(\hat{P}_{st}^{P} \right)^{\frac{1-\alpha}{1-\rho}} \right]^{\frac{1}{1-\alpha}}$$

$$\hat{P}_{st} = \left[(1 - S_{s0}^{I}) \left(\hat{P}_{st}^{D} \right)^{1-\eta} + S_{s0}^{I} \left(\hat{P}_{st}^{F} \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}$$

$$\hat{P}_{t} = \left(\sum_{s} S_{s0} \left(\hat{P}_{st} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$P_t = \hat{P}_t P_0$$

Step 4. Solve the path of Bellman equations:

$$V_t(k, m, o) = \nu \log \left(\sum_{o'} \exp \left(\frac{-C_{oo'km} + w_{o't} H_{o'}(k, m)/P_t + \beta V_{t+1}(k, m, o')}{\nu} \right) \right)$$

Step 5. Solve for the path of inclusive values of being type (k, m):

$$V_t^e(k, m) = \nu \log \sum_{o'} \exp \left(\frac{-C_{o,km}^e + w_{o't} 1 H_{o'}(k, m)/P_t + \beta V_{t+1}(k, m, o')}{\nu} \right)$$

Step 6. Compute the NPV for each field m

$$V_t(m|s)_t = \sum_{k} V_{t+5}^e(k, m) \pi(k|m, s).$$

Notice that for $t = T-4, ..., T-1$, one will use the steady state value.

Step 7. Compute $\{L_{kmt}\}$ by Monte Carlo simulation

(a) For each individual i and period t, generate a test score s_i for i. We do not index by t because they will be copies each period.

(b) Generate a type for i from the conditional CDF of the type distribution given s_i:

$$l_i = 1 + \sum_l \mathbf{1}(e > Q_l(s_i)),$$

where e is a uniform random variable.

(c) For $r = 1, ..., R$ calculate the ranked list of choices for individual i:

- Calculate the conditional CDF of choices for field:

$$\tilde{Q}_{ir}^F(m) = \frac{\sum_{m'=1\setminus m_{-r}}^{m} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V_t(m|s)\right)}{\sum_{m'=1\setminus m_{-r}}^{M} \exp\left(\vartheta_{mt} + \vartheta_{lm} + \theta_l V_t(m|s)\right)}$$

  Note: For cohorts $t = 1, ..., 5$ allocate students according to the steady state *values of $V_t(m|s)$. For remaining students, allocate according to $V_t(m|s)$.*

- Generate R draws from this distribution:

$$m_{r_i} = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_{ir}^F(m)),$$

  for a uniform random variable.

(d) Assign the agent sequentially to a field in their choice set based on acceptance probabilities (conditional on s). If the agent is not admitted to the first $R-1$ choices, then they are automatically allocated to the R-th choice.

(e) Determine the agents type:

$$k_i = 1 + \sum_{m \in m_{-r}} \mathbf{1}(e > \tilde{Q}_i^K(m, s)),$$

where Q_k is type CDF for agent i with field m and score s, and e is a uniform random variable.

(f) Calculate the share of individuals belong to type (k, m), scale up to $\bar{L}$.

Step 8. Using $V_t(k,m,o)$ solve for λ_{kmt} and $\lambda_t^e(o|k,m)$

$$\lambda_t(o'|o,k,m) = \frac{\exp\left(\frac{-C_{oo'km}+w_{o't}H_{o'}(k,m)/P_t+\beta V_{t+1}(k,m,o')}{\nu}\right)}{\sum_{o''}\exp\left(\frac{-C_{oo''km}+w_{o''t}H_{o''}(k,m)/P_t+\beta V_{t+1}(k,m,o'')}{\nu}\right)}$$

$$\lambda_t^e(o|k,m) = \frac{\exp\left(\frac{-C_{o,km}+w_{ot}H_o(k,m)/P_t+\beta V_{t+1}(k,m,o)}{\nu}\right)}{\sum_{o'}\exp\left(\frac{-C_{o,km}+w_{o't}H_{o'}(k,m)/P_t+\beta V_{t+1}(k,m,o')}{\nu}\right)}$$

Step 9. Compute $\hat{H}_o^{supply}$

(a) Use the flow equation:

$$L_{omkt} = (1-\delta)\sum_{o'}\lambda_t(o|o',k,m)L_{o'mkt-1} + \delta\lambda_t^e(o|k,m)L_{mkt}^e$$

(b) Construct H_{omkt}:

$$H_{omkt} = L_{omkt} \times H(o,m,k)$$

(c) Compute changes in human capital supply across occupations

$$\hat{H}_{ot}^{supply} = \sum_{k,m} H_{omkt}/H_{o0}$$

Step 10. Calculate aggregate expenditure (and income)

$$\hat{E}_t = \sum_o S_{o0}\hat{H}_{ot}\hat{W}_{ot}$$

Step 11. Calculate changes in domestic output

$$\hat{Q}_{st} = S_{s0}^x\left(\hat{P}_{st}^D\right)^{-\eta} + (1-S_{s0}^x)\hat{E}\hat{P}_{st}^{\eta-\sigma}\hat{P}_t^{\sigma-1}\left(\hat{P}_{st}^D\right)^{-\eta}$$

Step 12. Calculate a new guess for wages, by enforcing labor market clearing and output market clearing. Choose $\hat{W}_o$ such that

$$\begin{aligned}\hat{H}_{ot} &= \sum_s \frac{H_{so0}}{H_{o0}}\hat{H}_{so} \\ &= \sum_s \frac{H_{so0}}{H_{o0}}\frac{1}{\hat{W}_{ot}}\hat{Q}_{st}\left(\hat{P}_{st}^D\right)^{\alpha}\left(\hat{W}_{st}^N\right)^{-\alpha}\hat{\beta}_{so}^N\hat{A}_{sN},\end{aligned}$$

for all $o \in \mathcal{O}^N$; and

$$\begin{aligned}\hat{H}_{ot} &= \sum_s \frac{H_{so0}}{H_{o0}} \hat{H}_{sot} \\ &= \sum_s \frac{H_{so0}}{H_{o0}} \frac{1}{\hat{W}_{ot}} \hat{Q}_{st} \left(\hat{P}^D_{st}\right)^{\alpha} \left(\hat{P}^P_{st}\right)^{\rho-\alpha} \left(\hat{W}^P_{st}\right)^{-\rho} \hat{\beta}^P_{sot} \hat{A}_{sP},\end{aligned}$$

for all $o \in \mathcal{O}^P$.

Step 13. Update $\hat{W}_{ot}$ and iterate from Step 2 again.