Camila Comunello

Essays in Labour and Education Economics

Thesis submitted for the degree of

Doctor of Philosophy in Economics

University of Essex

September 6

2024

AUTHOR'S STATEMENT

I, Camila Comunello, declare that the work presented in this thesis is my own. Where information derives from other sources, this has been indicated and referenced. I declare that the work in this thesis was carried out in accordance with the requirements of the University's Regulations and that it has not been submitted for any other academic award.

Chapter 1 and 2 of this thesis are solo-autored. Chapter 3 is joint work with Carlos Carillo-Tudela, Alex Clymo, Anette Jackle, David Zentler-Munro and Ludo Visschers.

Errors are my own.

Signed by: Camila Comunello
Date: 6th of September, 2024

DATA AND FUNDING STATEMENT

There are a number of data sources used in this thesis. Chapter 1 and 2 utilise data from the identified Annual Census of Higher Education (ACoHE), the National High School Exam (ENEM) and employer-employee records from the Annual Register of Social Information (RAIS).

Access to all data sets was provided by the Protected Data Access Service (SEDAP) at the Anisio Teixeira National Institute of Research in Education (INEP). The analysis was carried out in a secure lab in Brazil and the data outputs presented here have been deidentified and cleared by the competing authorities. I gratefully acknowledge the exquisite support of the staff at SEDAP: Augusto, Marco, Maruska e Solymar, muito obrigada.

Chapter 3 uses the publicly available UK Labour Force Survey (LFS) provided by Office for National Statistics via the UK Data Service.

The completion of this thesis was only possible with the financial support from the Economic and Social Research Council through the South East Network for Social Sciences Doctoral Scholarship.

ACKNOWLEDGEMENTS

"*Every man I meet is my master in some point, and in that I learn of him*".

Ralph Waldo Emerson - Letters and Social Aims, 1875

There is great uncertainty in embarking on a research project of any kind and I was lucky to count with the encouragement and advice of great scholars. In special, I would like to thank my supervisors, Marco Francesconi, Elif Kubilay and Michel Serafinelli for their guidance and unwavering support. I am very grateful to my co-authors, Carlos Carrillo-Tudela, Alex Clymo, David Zentler-Munro and Ludo Visschers for the trust laid upon me - from day one. This thesis also benefited from the expert advice of Bob Miller, Cristina Gualdani, Joan Lull, João Thereze, Lisa Spantig, Mikhail Freer and Ursula Mello. To all, a heartfelt thank you. I am honoured to have learned and continue to learn from you.

While it seems appropriate to name some of the academic mentors who contributed to my development over the last few years, there are less noticeable agents who were fundamental to this work. To my friends, co-workers, school and university mates, students - dropouts and graduates, I hope this work can capture some of the complexity of the choices you made and trusted me to confide in.

A final thank you to my family, whose faith and support made this possible.

SUMMARY

This thesis is comprised of three stand-alone papers. The first paper, "Learning from What: The Informational Value of Grades and Wages," explores how imperfect information about personal schooling ability and labor market productivity impacts college enrollment and career transitions. Using Brazilian data, I estimate a dynamic model of college and work decisions, where individuals update their beliefs about their abilities based on noisy signals from grades and wages. I find that a significant portion of the variability in wages and grades is due to gradually revealed ability components. Counterfactual simulations show that perfect information about abilities would increase college graduation rates by 6.2 percentage points and that the option to return to education after a period of experimentation in the labour markets plays an important role in providing information that fills this gap.

The second paper, "The Role of Financial Aid for Low-Income Low-Achievers," examines the impact of financial aid on low-income students. By analyzing eligibility discontinuities for financial aid, I estimate the effects of University loans and/or grants on students' outcomes along their exam score distribution. Findings indicate that low-interest loans are more effective than full or partial grants in ensuring college completion for students from similar low-socioeconomic backgrounds. This effectiveness is attributed to loans acting as a commitment device, easing financial constraints while introducing a dropout penalty.

The final paper, co-authored with researchers from the University of Essex and the University of Edinburgh, investigates the UK labour market during the COVID-19 pandemic. Using novel job search data from the UK Longitudinal Household Survey, we document how individuals adjust their job search in response to changing employment patterns. Workers shifted their search towards expanding occupations and industries, but non-employed workers remain more attached to their previous roles, and those with lower education are more likely to target declining occupations. Workers from declining occupations make fewer transitions to expanding fields, indicating that those on the labour market's margins struggle more to transition away from declining jobs.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 LEARNING FROM WHAT: THE INFORMATIONAL VALUE OF GRADES AND WAGES

This paper investigates the informational value of grades and wages in shaping individuals' career trajectories and educational attainment. Using a unique dataset that integrates administrative records from Brazil, I analyze how imperfect self-knowledge regarding ability influences career and educational choices over a Bayesian learning framework. Results suggest that while education provides important signals through grades, wages emerge as a more accurate measure of an individual's abilities, Blackwell-dominating grades in informational value. Simulations reveal that a learning model results in delayed but optimized sorting on ability, increasing the college graduation rate by 6.2 p.p. thus enhancing wage outcomes for non-college goers and reducing inequality when compared to a full-information model. This highlights the fact that enrolling into higher education after a period of work experience may be a cost-efficient way to reduce the uncertainty involved in higher education decisions.

1.1 Introduction

Of all decisions made by individuals over the life cycle, career decisions could be the most decisive. Career decisions can determine people's level of income and wealth, happiness and satisfaction, fertility and health, and inter-generational propagation of wealth (Keane and Wolpin, 1997; Francesconi, 2002; Conti et al., 2010; Bertrand et al., 2015; Carneiro et al., 2021). Imperfect self-knowledge is one of the many sources of uncertainty that govern individual expectations about the future and consequent decisions[1]. With imperfect information about their own characteristics, such as tastes or *abilities*, it is not unusual that individuals experiment with different activities and/or occupations early in their lives before they settle on a career path[2].

In fact, many individuals follow a gap in their studies from high school graduation to when they join higher education. Late-bloomers, or students who join higher education after the age of 21, amount to approximately 44% of tertiary enrollments in the US, 38% in the UK, and 49% in Brazil [3]. Even when post-secondary enrolment

[1]According to Benabou and Tirole (2003), the three sources of *humanity* that produce uncertainty in decision-making: imperfect self—knowledge (treated in this paper), imperfect willpower, and imperfect recall.

[2]The time spent at University is generally treated as a time of experimentation in common knowledge as well as in economic theory (Manski, 1989).

[3]Based on 2017 data from: US - Beginning Postsecondary Students Longitudinal Study (BPS)

happens after high school graduation, high University attrition rates are common worldwide[4]. A less-documented fact is that late-bloomers' dropout rates are, on average, 8.9 percentage points lower than that of early entrants, which would suggest that some of the uncertainty about ability involved in the enrolment decision may be solved over time[5]. Not accounting for the process of information acquisition over time, and how valuable different ability signals can be, may lead to biased estimates of the returns to education (Heckman et al., 2008).

This paper attempts to measure the value of information about one's own ability acquired over different experience paths on career outcomes. First, I measure the impact of imperfect information from the extent to which the learning model compares to a model of full-information decision-making. As expected, eliminating information frictions increases sorting on ability (Arcidiacono, 2004; Arcidiacono et al., 2024). Learning follows a Bayesian updating structure with signals potentially coming both from grades and wages - assuming that people who work and study or just study or just work are equivalent conditional on their types. The model allows for learning to be correlated across different career paths. I create counterfactual exercises that generate type-specific alternative career profiles for different information sets.

Finally, I use conditions in Blackwell (1953) to robustly Blackwell order different ability signals and verify that wages provide more valuable ability signals - wage signals Blackwell-dominate the informational content of grades. There is a significant informational gain from wages, which affects lifetime income directly, and indirectly through higher education.

Related Literature

This paper builds on seminal work by Manski (1989), who rationalizes higher dropout rates in the first year of post-secondary education as a result of degree experimentation - only after enrolment can individuals realize whether graduation is both desirable and likely. This uncertainty about one's own ability is fundamental in explaining student dropouts even when ex-post payoffs are perfectly known (Altonji, 1993).

Innate ability - known or unknown to the agent at the start of the period -

available at: https://nces.ed.gov, UK - Higher Education Statistics Agency available at: https://www.hesa.ac.uk/data-and-analysis, Brazil - National Institute for Research in Education available at: https://www.gov.br/inep/pt-br.

[4]Attrition includes dropout, stopouts and program switches, see Arcidiacono et al. (2024) for a full review.

[5]In related work, Malamud (2011) show that delayed field specialization in high school does not lead to a loss in skill accumulation and decreases the probability of future field switches.

is an important driver of long-run outcomes. Ability has a direct effect on lifetime income through work productivity and an indirect effect through human capital investment and increased sorting (Keane and Wolpin, 1997; Arcidiacono, 2004). Ability sorting arises when students are aware of their ability type at the moment they make career decisions. Imperfect knowledge about one own's ability therefore has important implications for income inequality, equality of opportunity and for the efficiency of public policy such as affirmative action or compulsory schooling (Eckstein and Wolpin, 1999).

The empirical literature has documented interesting facts about students' beliefs about the future. Their ex-ante beliefs about ability can greatly differ from their ex-post beliefs (Stinebrickner and Stinebrickner, 2012; Crossley et al., 2022). For example, by surveying students at the University of California, San Diego, Betts (1996) finds that senior students are informed than freshmen about career prospects and are able to forecast their wages more accurately. In a similar way, Baker et al. (2018) document that students without any experience in the labour market tend to make less accurate estimates of their degree returns.

Accounting for the informational value of wages is not straightforward. Wages are a convolution of employer and employee known and unknown characteristics, and the aggregate state of the world (Miller, 1984; Sullivan, 2010; Lentz et al., 2023). In Miller (1984), individuals learn about their match-specific productivity once they start the job and they get better at estimating match-specific productivities with experience. Under imperfect information, the option to readjust one's career path, plays a big role in reducing inequalities (Stange, 2012). Sullivan (2010) quantifies that shutting down occupational mobility channels would reduce wages by 31% due to workers inability to sort into better matches.

A high level of schooling ability does not necessarily imply high labour market ability, but these are positively correlated (Belzil and Hansen, 2002; Arcidiacono et al., 2010). Discrete Choice models of correlated learning are notably hard to solve due to the need to integrate over expected abilities (Miller, 1984). I follow James (2011) and Arcidiacono et al. (2024) computational solution and treat abilities as known at different steps within the Expectation Maximization algorithm.

Most similar to this paper are the multi-period learning-through-grades sequential decision models applied to the U.S. higher education system in Stange (2012) and Arcidiacono et al. (2024). Relative to Stange (2012), this paper implements correlated learning about ability while in education and also while at work. Another difference is that the labour market is not treated as an absorbing state - a key feature aimed to capture inflows of mature students back to education. This paper follows a similar structural approach to Arcidiacono et al. (2024) using finite dependence results from Hotz and Miller (1993) and Arcidiacono and Miller (2020) to compute parameters of interest without requiring the full solution of the model. The salient contribution of this paper lies on the computation and implementation of systems to measure the value of ability signals received across different experiments (Blackwell, 1953; Brooks et al., 2024) while carefully distinguishing accumulated human capital from the innate ability component that is gradually revealed over time.

Through the use of novel panels constructed from individual-level administrative datasets, this paper moves away from standard approaches that rely on strong reference group assumptions and often involve estimating lifetime incomes from different samples (Altonji, 1993; Stange, 2012; Arcidiacono et al., 2024). Most importantly, the wage information used in this paper comes from actual employee records, which makes it less prone to measurement errors.

This paper also relates to a small but growing body of literature that pays attention to the mature student phenomena (Griliches, 1980; Jamieson, 2007; Murphy and Topel, 2016; Altonji et al., 2016; Bárány et al., 2023). Late bloomers are not only responsible for narrowing the gender and race inequality gap but also play a role in the long-run rise of the aggregate college share, and, more importantly, may largely affect measures of returns to education (Heckman et al., 2008; Belzil et al., 2017; Bárány et al., 2023).

The reminder of this paper proceeds as follows: section 1.2 presents the model set-up and justification; section 1.3 briefly introduces the data used in the empirical analysis and the Brazilian higher education environment; section 1.2 discusses the model solution and estimation procedures, parameter identification and main results; while section 1.5 discusses its implications to the learning process; and section 1.7 concludes.

1.2 A sequential model of career choices

When making career decisions, individuals face uncertainty about ability, which affects future wages and academic performance. Individuals are rational - they make the best possible forecast given the information available to them at a given time (current labour market conditions, past realization of ability signals and the process that describes aggregate uncertainty).

Learning about ability is costly. When in formal education, receiving signals from grades has a direct cost in the form of tuition fees and, even in the absence of fees - through public provision of education - or when learning about ability in the labour force, the indirect cost of learning from foregone earnings can be substantial.

Human capital accumulates throughout the life-cycle both because of investments in education, and because learning-by-doing on the job leads to accumulation of work experience.

1.2.1 Model Set-up

Let $t = 1, ..., T$, with $T = 9$ periods[6]. An individual i chooses, in each period, between 4 mutually exclusive *sectors*, $d_{it} = s$ with $s \in W, B, S, I$ that represents work, simultaneously work and study, study, simultaneously not work and not study, respectively. Individuals are allowed to differ according to M types unobserved by the econometrician[7]. At any point in time, the individual chooses d_t to solve the dynamic problem below[8]:

$$V_{t,m}(\Omega_t) = \max_{d_t} u_{t,m}(\Omega_t, d_t) + \beta \mathbb{E}_t \Big[V_{t+1,m}(\Omega_{t+1}) | \Omega_t, d_t, m \Big] \tag{1.1}$$

With $\mathbb{E}_t[.]$ denoting expectation conditional on the information set available at the beginning of period t and where Ω is a vector of state variables which include age, education, work experience, choice tenure, previous year decision, idiosyncratic shocks, current skill prices and ability signals. The first 6 are deterministic, while the last 3 are stochastic. β denotes the discount factor with $\beta \in (0, 1)$.

The instantaneous linear utility function is choice specific $u_{t,m}(\Omega_t, d_t = s)$ with $s \in \{W$ (work), S (study), B (both work and study), I (inactive - neither works or studies)$\}$. The last option, $d_t = I$, is the reference alternative normalized to zero and the flow utilities associated with the other options $u^W_{t,m}$, $u^S_{t,m}$ and $u^B_{t,m}$ must be interpreted relative to this baseline.

Work $d_t = W$

If $d_t = W$, the individual sole activity is to work on a specific job. When working, individuals are paid a wage (w_t - in logs). They cannot save and, therefore, derive utility from consuming the entirety of their wages.

[6]This is motivated by the number of periods I can track individuals in the available data.

[7]For more details on unobserved heterogeneity identification and interpretation of the m types, see Appendix B.3

[8]I omit the i subscript for simplicity.

$$u^W_{i,t,m} = w_{i,t,m}$$

The log-wage function takes a Mincerian form as in Bárány et al. (2023), with the exception that I discriminate between years of experience accumulated before and after University graduation.

$$w_{i,t,m} = \beta^w_0 + \beta_t + \overbrace{\beta^w_{1,bu} exper_{i,t,m,bu} + \beta^w_{2,t,m,bu} exper^2_{i,t,m,bu}}^{\text{experience before university}} + \beta^w_3 grad_{i,t,m} +$$
$$\underbrace{\beta^w_{1,au} exper_{i,t,au} + \beta^w_{2,au} exper^2_{i,t,au}}_{\text{experience after university}} + \beta^{wx\prime} X_i + \zeta_i + \epsilon^w_{i,t} \quad (1.2)$$

Where $grad$ is a dummy that takes the value of 1 if a person has a University degree by time t and X_i is a vector of time-invariant personal characteristics such as gender, ethnicity, parents' education, etc). The mincer equation residual is composed of ζ_i, the latent innate ability signal, and $\epsilon^w_{i,ti,t}$ a choice-specific shock that assumes the common logit form. Specifically in the case of options that involve wages, $\epsilon^w_{i,ti,t}$ can be decomposed into $\epsilon^w_{i,t} = r_t + \epsilon^W_{i,t}$. Where r_t, the individual-invariant term, can be interpreted as a skill price assumed to be governed by an AR(1) process, and $\epsilon^W_{i,t}$ in the sector W-specific idiosyncratic shock.

Study $d_t = S$

When studying, the individual incurs a program j-specific cost τ_{jt}. The coefficient τ_{jt} can be interpreted as the total cost of taking part in a degree program for the average student in program j at time t, including tuition fees and the cost of dedicated hours of study. Students also derive a m-type-specific and time-permanent (dis)utility denoted by δ^S_m.

$$u^S_{i,t,m} = \delta^S_m - \tau_{j,t} + \epsilon^S_{i,t}$$

Work and Study $d_t = B$

Alternatively, the individual can both study and work in the same period. Then, they earn a wage $w_{i,t,m}$ but also incur the cost of acquiring education τ_{jt} and receive the permanent m-type-specific (dis)utility of dedicating time to both activities, δ^B_{0m}.

$$u^B_{i,t,m} = \delta^B_{0m} + w_{i,t,m} - \tau_{j,t} + \epsilon^B_{i,t},$$

1.2.2 Grades

Grades do not enter the instantaneous utility functions. They are merely outcomes of the choice of studying, defined by:

$$g_{i,t,m} = \beta^g_0 + \beta^g_t + \beta^{gx\prime} X_{i,t} + \zeta_{i,t} \tag{1.3}$$

Where X_i is a vector of time-variant and time-invariant characteristics of an individual, such as gender, ethnicity, age, program of study, years of work experience and years of education.

Human Capital transferability

It is important to distinguish between innate ability that is revealed to the individual over time and the process of human capital accumulation i.e. the acquisition of skills with experience that also evolves over time. The first assumes that there is a fixed ability level that is not fully accounted for by the decision maker in the initial periods but affects their outcomes, and eventually, it is fully incorporated into individuals' information sets. On the other hand, accumulated human capital and the process from which it evolves are known to the individual before a decision is made.

I account for 3 different types of human capital as states in the model. The first are the empirically observed measures of human capital, years of experience and years of schooling. Both evolve by a unit every t if the agent was employed or enrolled in education in $t-1$. Years of experience are fully transferable. If experience is sufficiently transferable, one can flow from one sector to the next even in the absence of aggregate shocks. Finally, choice tenure $ten_{i,t}$ accumulates only when $d_{t-1} = d_t$ and it captures the part of human capital that fully depreciates with a sectoral switch.

1.2.3 Information sets

Individuals face uncertainty about idiosyncratic shocks ϵ, future skill prices r, and their innate ability.

Idiosyncratic shocks and skill prices

Let $F(\epsilon_{i,t+1}, r_t|\Omega_{a,t})$ denote the distribution of these shocks in the next period, with $\epsilon_{i,t} \equiv (\epsilon^W_{i,t}, \epsilon^S_{i,t}, \epsilon^B_{i,t}, \epsilon^I_{i,t})'$ and r^w_t.I assume the processes for the idiosyncratic shock and the future skill prices are mutually independent[9]:

$$F(\epsilon_{i,t+1}, r_t|\Omega_{i,t}) = F^\epsilon(\epsilon_{i,t+1})F^r(r_t|\Omega_{i,t})$$

$F^\epsilon(.)$ is Generalized Extreme Value (GEV) with a scale parameter ρ.

Idiosyncratic shocks ϵ^s are assumed to be independently and identically distributed across individuals, and uncorrelated with individual and aggregate characteristics. The variance of the random shock σ^2_s is correlated across choices, but independent of all other state variables. Some unobserved characteristics (abilities, preferences, credit constraints) that are likely to affect labour market entry or enrolment may be at play across the different s sectors.

Beliefs about ability

At each period, individuals use their realization of grades and/or wages to update their beliefs about their true innate ability, a_i. Denote $\Lambda_{i,t} \equiv [\Lambda^w_{i,t}, \Lambda^g_{i,t}]$ the vector of potential signals received by individual i at a given point in time. Let the signal associated with each value of wages and grades, the ζ from equation 2 and (1.3), represent a noisy measure of their true ability a_i, respectively:

$$\Lambda^w_{i,t} = \zeta^w_{it} \tag{1.4}$$

$$\Lambda^g_{i,t} = \zeta^g_{it} \tag{1.5}$$

This means wage signals Λ^w are only received when individuals are employed, while grade signals Λ^g are only received when individuals are enrolled. When a person is simultaneously working and studying, they may receive both signals. If an individual is neither studying or working, they receive zero ability signals.

At $t = 0$ individuals' beliefs are given by the population distribution of ability, $\overline{a}$. At $t > 0$ individuals' update beliefs in a Bayesian fashion following (De Groot, 1970). Let $E_t[a_i]$ denote posterior ability and $\Sigma[a_i]$ denote posterior covariance, while $\overline{\Lambda}_{it}$ is the $|\mathcal{S}| \times 1$ vector of signals corresponding to the choice made at t. D_{it} is the $|\mathcal{S}| \times |\mathcal{S}|$ matrix of decisions.

[9]This is consistent with the notion of atomistic individuals in Altuğ and Miller (1998)

$$E_t[a_i] = (\Sigma_{t-1}^{-1}[a_i] + D_{it})^{-1}(\Sigma_{t-1}^{-1}[a_i]E_{t-1}[a_i] + D_{it}\overline{\Lambda}_{it})$$

And:

$$\Sigma_t[a_i] = (\Sigma_{t-1}^{-1}[a_i] + D_{it})^{-1}$$

Next, I describe the notion of finite dependence used to avoid full-solution methods in recovering the model parameters and how this maps into the empirical choice probabilities and the panel dimensions of the data.

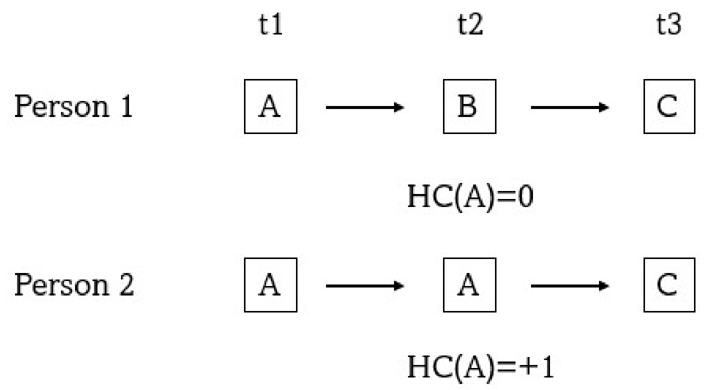
1.3 Data and Finite Dependence

The identification of flow utilities under correlated learning relies on observing sequences of decisions and outcomes that can undo the dependence of the state variables on the previous choice (Hotz and Miller, 1993; Arcidiacono and Miller, 2020). The intuition between finite dependence is analogous to that of differences-in-differences in the sense that identification relies on decision switchers. Consider the case when human capital is fully transferable (it does not depreciate). If two identical[10] individuals $i = 1, 2$ are in different sectors at time $t-1$, say $s_{1,t-1} = A$ and $s_{2,t-1} = B$, switch to the same sector $s_{i,t} = C$. Then, at t, they will be identical in the eyes of the model. A similar argument applies to the case when human capital is only partially transferable. Denote $h = 1$ the length of time that the partially transferable human capital takes to depreciate.

The key to restore state independence is to use all observed pairs of sectoral moves to construct h-length choice deviation paths. For instance, say now person 1 and 2 both start their careers in A, then at $t = 2$ person 2 makes a career switch to B, and at $t = 3$ they finally both switch to C as illustrate in Figure 1. In that case, the partially transferable part of human capital associated with sector A fully depreciates by the time that person 1 switches to sector C so the difference between person 1 and 2 is now only due to the additional unit of human capital accumulate by person 2 by staying in A at $t = 2$. This intuition applies to sequence of choices of any length as long as $T \geq h + 2$ and the human capital accumulation process is known.

[10] In observed (age, experience,...) and non-observed characteristics.

Figure 1 – A simple illustration of Finite Dependence



Note: A three-period illustration of Finite Dependence. Persons 1 and 2 can choose between options A, B, C*.* HC *stands for Human Capital.*

For this purpose, this paper combines individual-level data from the Annual Census of Higher Education (ACoHE), the Annual Register of Employment Records (RAIS) and administrative data from two major student examination: the National High School Exam (ENEM) and the National Exam of Student Performance (ENADE).

The ACoHE assembles all student records from all Brazilian higher education institutions. From the panel version of the ACoHE, it is possible to monitor the progress of each Brazilian student since 2008, as they are assigned an identification number kept throughout their university career. This allows me to link the socioeconomic characteristics of individuals from the moment which they enrol until they graduate or leave their chosen university degree.

The RAIS is a matched employer-employee dataset that compiles the population of formal jobs in Brazil. Employers must fill their RAIS declaration every fiscal year[11], mandatorily. RAIS contains particularly useful information on individual and job characteristics such as monthly wages, employment status, occupation, level of education and place of residence. I use the identified version of RAIS from 2009 to 2019 to match individuals across datasets and construct a map of their professional and educational transitions.

Additional administrative datasets allow me to observe student performance prior and during their trajectory in higher education. From the ENEM Dataset, I can recover students' performance in a standardized test taken by most students at the end of high school. This exam serves as an entry exam for all public and most private universities, and it is a requirement for application to government-sponsored financial aid programs. The ENEM application form also contains detailed informa-

[11]In Brazil, the fiscal year coincides with a calendar year.

tion about the applicant's socioeconomic characteristics. During their university careers students required to take ENADE. The ENADE is a standardized exam taken nationally by students in different stages of a degree. The exam comprises a block of questions of general knowledge and a second block of subject-specific questions. The exam is compulsory - students are required to take ENADE in order to graduate. Brazil's Ministry of Education (MEC) uses ENADE results to evaluate the quality of degree programs nation-wide.

The above datasets are combined in a panel that tracks individuals annually from 2009 to 2019, mapping their characteristics, work and schooling decisions, experience and tenure, wages and test scores over time. Appendix A offers a detailed description of the data, including timing of data collection of the different data sources, procedures used to match individuals across datasets, variable definitions and descriptions of the cohort.

1.4 Identification and Estimation

Because the idyossincratic shocks are assumed mutually and serially uncorrelated, a version of the model without unobserved heterogeneity can be estimated sequentially. This assumption makes it possible to integrate out the initially unknown heterogeneity parameters from the individual likelihood contribution. The likelihood contribution can then be written as an integral where $F(A)$ is the probability density function of the initially unobserved ability term.

Because choices depend on ability A solely through the observed signals we can separate this expression between choices and outcomes using the law of successive conditioning. The contribution of the sequence of decisions L_{id} is given by the product of the decision probabilities obtained in a discrete choice problem with Generalized extreme value distribution over time and, in a similar manner, the likelihood contribution from the sequence of outcomes - grades g and wages w - can be written as [12]:

$$L_{i,g,w} = L(g_{i1}|d_{i1}A)...L(g_{iT}|d_{i1},...d_{iT},A) \times L(w_{i1}|d_{i1}A)...L(w_{iT}|d_{i1},...d_{iT},A)F(A)d(A) \quad (1.6)$$

Where the normal conditional pdfs for each outcome is given by $L(Y_{it}|d_i,...,d_{it},A)$. With the above likelihood structure in mind, it turns out that the the choice component and the outcome component of log-likelihood are additively separable.[13]

[12] For the full derivation see Appendix B.2.

[13] As noted by Arcidiacono and Miller (2011), adding unobserved heterogeneity compromises the additive separability of the log-likelihood. However, as in the EM algorithm, the unobserved types are treated as known for the maximization step, additive separability is restored.

I then use the EM algorithm (Dempster et al., 1977) for likelihood maximization. The estimation iterates over the E (expectation) and M (maximization) steps until convergence.

- E-step: For each iteration k, the ability posteriors for each individual n is updated using the observed grades and wages. Then, the population covariance matrix of the ability distribution is updated following:

$$\Delta^k = \frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_i^k(A) + E_i^k(A)E_i^k(A)'\right) \tag{1.7}$$

  Where $E_i^k(A)$ is the posterior ability mean and $\Sigma_i^k(A)$ is the previously computed posterior ability covariance.

- M-step: Given the parameters of the the posterior distribution of ability obtained in the E-step, the M-step maximizes the log-likelihood of the outcome data or the Expected Log-likelihood El_i^k:

$$El_i^k = \int ln\left[L(g_{i1}|d_{i1}A)...L(g_{iT}|d_{i1},...d_{iT},A) \times L(w_{i1}|d_{i1}A)...L(w_{iT}|d_{i1},...,d_{iT},A)\right]F(A)d(A) \tag{1.8}$$

  The additive separability of the log-likelihood is present not only over choice options but also over time. It follows that for each time period t and the normality assumption on the grade shocks that:

$$\int \ln\left[L(g_{iT}|d_{i1},...,d_{iT},A)\right]F(A)d(A)] = -\frac{1}{2}\ln\left(2\gamma\sigma_t^2\right) - \frac{1}{2\sigma_t^2}\left(\Sigma_i^k(A) + (\zeta_i^g)^2\right) \tag{1.9}$$

  The remaining parameters are updated by solving the following minimization routine:

$$\min\sum_{it}\left(\ln(\sigma_t^2) + \frac{1}{\sigma_t^2}\Lambda_i^k(A)(\zeta_i^g)^2)\right)$$

  The log-likelihood of the work outcomes - wages ω - are obtained in a similar fashion.

Once the parameters from the wage and grade equations, the flow payoffs can be estimated using finite dependence (Arcidiacono and Miller, 2020) through the following steps:

1. Estimate Conditional Choice Probabilities (CCPs) through a multinominal logit in a first stage;
2. Obtain the difference in expected values from the finite dependent paths;
3. Express the flow utility parameters as a function of the CCPs (Hotz and Miller, 1993) and estimate it as a multinomial logit.

1.5 Model Estimates

Next, I present the results from the model estimation. I leave description of the sample to the Appendix. All tables are collected at the end of the paper for readability. The model was estimated for $M = 4$ heterogeneity types unknown to the econometrician but observed by the agent[14]. First, I examine the parameters of the outcome equations in 1.2.1 and (1.3) presented in Table 1.

Estimates from the outcome equation from Table 1 indicate that being non-white has no statistically significant effect on grades but is associated with a 1.25% decrease in wages, while being a female is associated with 2.66% higher grades but 2.81% lower wages. Those with parents with a higher education degree tend to earn higher wages although they score lower grades while at University. Previous record of exam performance given by the National High School Exam (ENEM) is positively correlated with both grades and wages. Regional differences in outcomes are also in line with the expected with lower wage and grades performance in the North and higher performance in the Southeast and South. An additional year of work experience is associated with a 8.1% increase in wages, while an additional year of education increases wages by 6.2%. The graduation wage premia is 15.8% which is slightly above the number documented in the US by Arcidiacono et al. (2024) but in line with evidence from Brazil and other developing countries (Dix-Carneiro and Kovak, 2015).

Another product of the empirical application of the model is the estimated variance-covariance matrix for the gradually revealed unobserved ability. I can summarize the variance-covariance matrix to obtain measures for the correlation between ability revealed through different career paths. Table 2 displays these results. Work (W) and work and study (B) are, unsurprisingly, the channels with the highest revealed ability correlation (0.79). By construction, individuals who work and study receive signals from both wages and grades. As long as work and study individuals can access jobs with similar informational value to those who solely work, the correlation between B and W should be high. The ability correlation

[14] Appendix B.3 details the estimation procedure.

between work and study however, is positive but much lower (0.34). This number is similar to what was found by Belzil and Hansen (2002) (0.28) and by Arcidiacono et al. (2024)[15].

Complementary to the outcome equation regression in Table 1 are the flow utility estimates in Table 3. Non-whites receive positive utility from work and study and some disutility from solely working, in contrast to their white counterparts. In contrast, females tend to receive more utility than males when involved in study activities, and less utility when solely working. High school performance seems to affect positively all flow utilities although (revealed) abilities only impact utility indirectly through wages. Academic performance does not affect current utility and therefore ability does not affect flow payoffs through grades. The coefficients for previous choices reveal statistical significance of switching costs between work-to-study switches and transition net benefits for switches that contain $s = B$ (work and study). This is expected since for individual who already work and study, switching to just study or just work barely has a cost, and vice-versa. Looking at permanently unobserved types, it seems that only High-High comparative advantage types display significant results - placing higher return to specializing in a sole activity, with higher utility associated with only studying.

I now examine the model's fit by using the parameter estimates and the empirical CCPs. An ability vector is drawn for each individual from a multivariate normal distribution previously defined. Each individual is then assigned an unobserved type from a categorical distribution. Next, I draw the idyossincratic and skill price shocks and compute choice probabilities based on observed states Ω and the draws, flow utility estimates, and the empirical CCPs. Implied ability beliefs are then computed using these shocks and ability draws, updating the state space and repeating for $t = 1, ..., 10$ periods across each individual in the estimation sample.

Table 4 displays how well the model aligns with moments in the data. The model-predicted and actual empirical moments are pooled across the 10 periods. The baseline model does a good job at reproducing most features of the empirical data, corroborating with the idea that information about ability is gradually revealed through both education and work experiences. I also simulate a full-information version of the model where individuals are fully aware of their ability at the time decisions are made. In line with the canonical model, the full-information model predicts enrolment at a much earlier age, fewer work-to-school transitions and low stop-out rates. A third simulation is run by sitting down the acquisition of

[15] Arcidiacono et al. (2024) estimate the correlation of unobserved ability (in patrenthesis) between studying for a science degree and working in white-collar occupation (0.28) and blue-collar occupation (0.23).

information about ability through wages. All ability revelation must now come from grades. Such a model overestimates time to graduate, due to higher experimentation at University. It also predicts higher enrolment rates as now the enrolment option carries all the information value in line with Stange (2012).

The most intuitive counterfactual exercise designed to measure the effects of the gradual revelation of ability is to compare the baseline model from Section 1.2 with simulations from an alternative model where abilities are fully observed by the agent before the decision (Stange, 2012; Arcidiacono et al., 2024). An additional exercise , could be to shut down the learning process for only one type of signal and simulate counterfactual scenarios to try to measure the impact of different types of information acquisition.

The full-information comparison: Comparing Column 2 (full-information) to Column 1 (baseline) in Table 4 reveals a 25 percentage point increase in graduation rates if individuals could perfectly observe their ability type at the moment the career decision is made. This rise in graduation rates occurs despite the 9 p.p. increase in enrolment rates, hinting at increased sorting. These changes result from individuals who, under imperfect information, enroll in college, discover a poor fit, and then drop out. With full-information, these individuals would not enroll initially. Conversely, some students who do not enroll under imperfect information choose to attend college continuously when they know their ability to be of high type.

The learning through grades only comparison: A second significant difference between the full-information counterfactual and the baseline. When comparing the baseline model to one where only signals from wages are removed (Column 3), there is an additional decrease in college graduation rates. This decrease results from two opposing forces. When learning from wages is shut down, the only way to discover information about one's own ability is through enrolment. Mechanically, the value of the enrolment option increases and more people start studying in the early periods. However, with the low informational content of grades, compared to that of wages, ability revelation is slower and, more people with low propensity to graduate join higher education, leading to a net decrease in graduation rates in comparison to the full-information scenario.

The impact on the graduate premia: Table 5 presents the variation in average graduate wage premia across different age groups for the baseline model, full-information, and learn-through-grades counterfactuals. The graduate wage premia are concave over age, with maximum premia peaking at around age 35 in the baseline, full-information and learn-through-grades scenarios.

The dynamics of the graduate premium change significantly across scenarios.

The full-information scenario displays higher wage premia overall, but particularly in the beginning of individual careers. These differences are largely due to high-ability type individuals being more likely to enrol in higher education early in their careers under full-information. At the same time, low-ability individuals are less likely to enrol in education and, consequently, graduate. In the baseline model, most high-ability individuals also graduate, eventually, but this happens at a slower pace than in the full-information counterfactual, generating some ability "misallocation" in the short-run. This is carried through to the long run via less experience accumulation (in comparison to the full-information counterfactual).

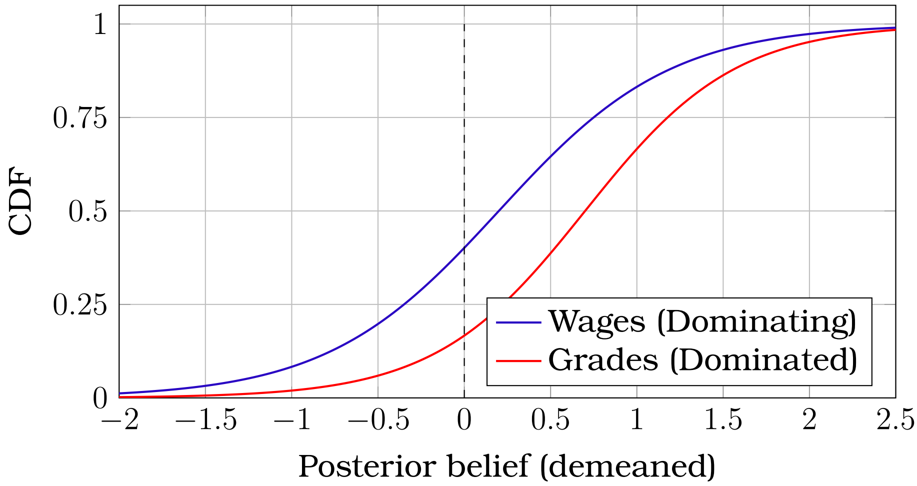
Simulations from the learn-through-grades scenario, suggest slower revelation of ability over time than the baseline, generating ability "misallocation" with higher persistence. This further closes the graduate premia compared to the full-information scenario, highlighting the role of information and speed of information revelation in determining wage inequality.

1.6 Ranking different signals

I now turn to the ranking of the two sources of information of ability presented in this model. I use notions of Blackwell dominance to compare two signals in isolation. A signal A is considered more valuable than B if it leads to a better distribution of beliefs regardless of the specific decision problem. This condition does not account for interactions between A and other signals in the environment. If an agent already possesses a signal similar to A, B might become more valuable because of the complementarity or redundancy with existing information. However, this condition does apply to when the agent's preferences are unknown. In my setting, this would translate to the individuals of a certain m-type who really enjoy working but dislike school - thus preferring to receive wage signals rather than grade signals, for example.

Blackwell (1953) proofs that saying that a signal A Blackwell dominates signal B is equivalent to saying that the posterior belief distribution of A First-Order Stochastically Dominates that of B. Having identified the revealed individual ability from previous steps and the signal sequences associated with grades and wages and their respective posterior beliefs from our structural assumptions, I can use it to test some properties of the posterior belief distribution. I approximate the cumulative density functions from the empirical values of grades and wages and the functional forms specified in section 1.2. The resulting cdfs are displayed in Figure 2. The First-Order Stochastic Dominance of wage signals is clear from the graph - indicating that wages provide higher-quality signals of ability regardless of preferences for a

Figure 2 – CDF comparisons



Note: Comparison of the two approximated CDFs, where CDF of posterior beliefs of wage signals stochastically dominates the posterior beliefs of CDF of grades. Success is defined by posteriors correctly reporting the ability revealed at T.

specific action.

1.7 Conclusion

The main goal of this paper was to quantify the effects of the revelation of information about ability on individuals' career decisions and consequent work and school performance. When individuals face uncertainty about their own ability, the higher informational value of wages relative to grades shines a light on alternative career paths: individuals may experiment with the labour market before they join University.

I estimate a dynamic discrete choice model where individuals update their ability priors from new information received by their realization of grades and wages. In line with Arcidiacono et al. (2024), I find that the process of information revelation is slow-paced, contrary to the idea of the full-information revelation in the first period proposed by Manski (1989). In my model, higher first-year dropouts are not entirely a product of information revelation, *per se*, but of the realization that there is little informational value added to the schooling option. When the low value added is realized, individuals prefer to continue to acquire information about ability in the labour market, where the costs associated with receiving ability signals are relatively low.

The novelty of this paper lies on the comparison of the informational values added of the two different ability signals - grades and wages. In order to rank the informational value of grades and wages I rely on i) the parametric assumptions on the wage and grade functions, ii) the availability of data on the history of signals received by individuals and, most importantly, of individuals who receive both signals simultaneously and iii) Blackwell order conditions that are robust to mispecification of the belief updating process (Blackwell, 1953; Brooks et al., 2024). I find that ability signals from wages dominate ability signals from grades - concluding that on-the-job experience tends to be more informative about one's own ability than formal education.

These findings have important implications for policy relating educational attainment to income inequality. First, education does not (privately) hurt low-income low-ability types. The slow revelation of true ability somewhat counteracts the rise of income inequality from increased sorting - generating ability misallocation in the short-run. This is due to some of the low-income low-ability individuals"misallocated" to the educational sector graduating and receiving higher wages than their peers without a degree during the first couple years of work after graduation. However, this difference in earnings wears off over time as these individuals gain experience on the job.

At the same time, high-ability individuals who did not participate in higher education in the first periods tend to earn less than their graduate peers in the short run. But in the long run, early graduate and non-early graduate earnings tend to converge. The speed of convergence, however, depends on whether the individual re-entered higher education or not. Which may be related to graduation carrying higher informational value to the employer than to the individual. In such a setting, educational policy of the likes of affirmative action or financial aid that focus on older individuals, who tend to have higher self-knowledge, is efficient in the way that it assures places in higher education to those who can achieve higher private returns to their human capital investment.

Table 1 – Estimates of wage and grade outcomes

	(1) Grades	(2) Wages
Non-white	-0.0004	-0.0125 **
	(0.004)	(0.008)
Female	0.0266***	-0.0281**
	(0.003)	(0.002)
Public schooled	-0.0006	-0.0002
	(0.002)	(0.003)
Mother educ = Low	0.0001	-0.0001
	(0.007)	(0.005)
Mother educ = Med	-0.0134	0.0092
	(0.007)	(0.005)
Mother educ = High	-0.0347***	0.0465***
	(0.007)	(0.006)
Enem score	1.052***	0.940***
	(0.161)	(0.127)
Rural area	-0.0025*	-0.2021***
	(0.005)	(0.003)
Region:		
Northeast	0.0130	0.0220*
	(0.006)	(0.009)
North	-0.0263***	-0.0139**
	(0.007)	(0.006)
Southeast	0.0328***	0.273***
	(0.006)	(0.005)
South	0.0563***	0.158***
	(0.007)	(0.006)
W-S transition	0.0064	
	(0.003)	
Switched courses	0.0117	
	(0.003)	
Experience pre-grad		0.081***
		(0.004)
Experience post-grad		0.082***
		(0.002)
Years of educ		0.062***
		(0.005)
Graduate		0.158***
		(0.008)
Unobserved type:		
LH	0.002*	0.0563***
	(0.001)	(0.00791)
HL	0.165***	0.1126***
	(0.002)	(0.00243)
HH	0.392***	0.5621***
	(0.006)	(0.00787)
Observations	1,105,821	15,689,456
R^2	0.116	0.253

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Regression includes time fixed effects, age dummies, and a constant. Baseline category for region is Central. Baseline category for mother's education is non-alfabetized.

Table 2 – Ability Correlation Matrix

	Work	Study	Work and Study
Work	1.000		
Study	0.337	1.000	
Work and Study	0.794	0.426	1.000

Table 3 – Flow Utility Parameter Estimates

Variable	Study	Work and Study	Work
Non-white	0.063	0.025	-0.011
	(0.026)	(0.008)	(0.001)
Female	0.125	0.063	-0.021
	(0.005)	(0.007)	(0.002)
ENEM score	0.482	0.257	0.223
	(0.015)	(0.068)	(0.017)
Mother educ = High	-0.237	-0.002	0.125
	(0.010)	(0.0014)	(0.007)
Previous study		0.104	0.007
		(0.028)	(0.079)
Previous work and study	0.225		0.628
	(0.022)		(0.012)
Previous work	-0.029	0.021	
	(0.007)	(0.010)	
Graduate			0.128
			(0.002)
Experience pre-grad	0.407		
	(0.009)		
Experience post-grad	0.572		
	(0.002)		
Unobserved type			
LH	-0.028	-0.017	0.024
	(0.027)	(0.073)	(0.010)
HL	0.012	0.010	0.022
	(0.013)	(0.033)	(0.015)
HH	0.352	-0.111	0.128
	(0.049)	(0.008)	(0.031)
Log likelihood	-284,583		
Observations	4,572,132		

Note: Regression includes time fixed effects, age dummies, controls for geographic region, whether individual studied in public schools, mother's level of education, lives in a rural area, study program fixed effects

Table 4 – Model Fit

	Empirical	Baseline	Full info.	Learn through grades
Years to graduate	3.31	3.29	3.1	3.9
W-S transitions	0.67	0.71	0.11	0.16
Graduation rate	0.47	0.45	0.7	0.58
Stop out rate	0.12	0.09	0.03	0.05
Enrolment rate	0.2	0.17	0.23	0.21
Age at enrolment	29.1	29.0	18	23

Note: Model frequencies are constructed using T=10 simulations of the structural model for each individual included in the estimation. Moments are constructed pooling the 10 periods. "Full info" refers to a counterfactual where individuals have complete information about their abilities. "Learn through grades" refers to a model specification where individuals only update their beliefs over abilities through grades.

Table 5 – Average wage premia in baseline and counterfactual models

	Graduate wage premium			Change in graduate wage premium	
	Baseline	**Full info.**	**Learn through grades**	**Full info.**	**Learn through grades**
At age 30	0.368	0.507	0.286	0.136	-0.082
At age 35	0.302	0.511	0.272	0.209	-0.003
At age 45	0.249	0.421	0.265	0.172	0.002
At age 50	0.244	0.404	0.253	0.160	0.001

Note: "Graduate wage premium" is the difference in average log wages between college graduates and non-graduates. "Change in graduate wage premium" refers to the difference in wage premia relative to the baseline model results.

2 THE ROLE OF FINANCIAL AID FOR LOW-INCOME LOW-ACHIEVERS

I use a series of discontinuities in policy eligibility to uncover the impact of different types of financial aid directed to low-income students, along their exam score distribution. Results show that low-interest loans are more effective than full and 50% grants in securing College completion for students of similar low socioeconomic backgrounds. I argue this is because loans act as a commitment device by not only relaxing present financial constraints but introducing a penalty in case of dropout.

2.1 Introduction

It is a general consensus that education is an essential channel for improving social mobility and promoting economic growth (Becker, 1965; Lucas, 1988; Romer, 1990; Manski, 1989; Moretti, 2004). It is also known to produce other desirable social effects such as reducing crime (Lochner and Moretti, 2004) and improving health levels (Clark and Royer, 2013). Still, not everyone has access to the same educational opportunities. Inequality in access to higher education arises from disparities in schools, neighbourhoods, and other environmental factors that accumulate since birth (Heckman and Carneiro, 2003; Fryer and Levitt, 2013; Chetty et al., 2011). While early-life interventions are extremely important, late-life interventions are shown to be effective in increasing intergenerational mobility and reducing segregation (Deming and Dynarski, 2010; Dale and Krueger, 2014). For instance, Chetty et al. (2020) find that relaxing merit-based selection criteria for low-income students increases social-economic mobility, without changing colleges' educational programs' quality.

A common instrument to equalize access to higher education is financial aid. Plenty of empirical studies have focused on financial aid programs with mixed results. Because financial aid assignment is usually merit-based, evidence often comes from a population of high-achievers (van der Klaauw, 2002; Avery and Turner, 2012; Modena et al., 2020). Perhaps for that reason, empirical evidence on the impact of grants on student achievement is mixed (Cohodes and Goodman, 2014; Deming, 2023) since high-achievers may be less likely to change their decisions conditional on financial aid than their lower-achieving counterparts (Dynarski, 2005; Dale and Krueger, 2014; Hoxby and Turner, 2015). The great majority of low-income students, however, are not high-achievers. These students in particular, are more sensitive to how the subsidy information is communicated and more debt-averse (Caetano et al., 2011; Hoxby and Turner, 2015; Baum and Schwartz, 2015; Belzil and Hansen, 2002).

Despite the popularity of the theme, there is no consensus on whether grants are more efficient than loans, or even other nudges, in promoting higher education attainment. Or a conclusion about what is a more efficient policy. This is salient since I know that small changes in the design of educational policies can significantly affect the results for high-performance and low-income students (Hoxby and Turner, 2015); (Peterson et al., 2003) and transform the long-term distribution of income (Chetty et al., 2017).

This study aims to investigate the impact of two government-sponsored higher education financial aid programs -grants and loans, on student enrolment and attainment in Brazil. The Brazilian government is the largest student sponsor in the country, being responsible for awarding nearly 76% of all student loans 93% of all student grants. State-sponsored financial aid is awarded according to student rankings in the National High School Exit Exam (ENEM). Assignment thresholds are determined for each financial aid/degree program/education institution triple. By leveraging multi-cutoff regression discontinuity techniques, I separately estimate average treatment effects by type of financial aid and financial aid effect across the distribution of student exam scores.

For some modalities of grants and loans, the eligibility cutoffs are the same. This allows me to leverage variation on the effects of different modalities of financial aid over the same and across different exam score cutoffs. Overall, results show that loans are more effective than grants in securing college completion. I argue that this could be due to a discipline mechanism: while grants affect students' choices through a reduced cost of education, loans reduce the value of opting out.

Related Literature

A considerable body of applied literature concentrates on the role of financial aid in relaxing current financial constraints and how it affects schooling decisions. It is generally expected that loans or grants should increase college take-up, but these policies are costly and often inefficient as they result in high enrolments with higher dropout rates.

In a classic setting, the impact of financial aid in student attainment is straightforward. Financial aid increases the likelihood of enrolling in a higher degree by lowering the expected costs of acquiring education or relaxing current budget constraints (Avery and Turner, 2012; van der Klaauw, 2002; Epple et al., 2006). However, empirical evidence on the direction, significance and size of the financial aid effect is puzzling.

A stream of literature documents that the effects of financial aid seem to depend heavily on how this aid is introduced. The "flattering effect", for instance, happens when a student prefers a scholarship with a name associated to social distinction to another generically named award of higher nominal value (Benhassine et al., 2015). Other students show higher aversion to loans depending only on the complex appearance of how the scheme is labelled (Avery and Hoxby, 2003; Angrist et al., 2010).

A second strand documents how lower the family income, the greater is the impact of how information on costs and financial aid is communicated in their decisions (Hoxby and Turner, 2015). Low-income students and their parents are more likely to overestimate costs of attending college (Avery and Kane, 2004) and are more reluctant to finance college through loans (Baum and Schwartz, 2015). In Latin America, Caetano et al. (2011) found that low-income students applying for financial aid are less likely to choose arrangements labelled as "debt" or "loan", as opposed to other financially equivalent contracts without these labels.

At the same time, Dynarski (2005) shows that some individuals (especially high-achievers) are unlikely to change their decision to complete a degree conditional on receiving financial aid or not - independent of family income. However, large evidence is found that direct cost of school is not the only cause of non-completion since even in the case of free tuition (Cameron and Heckman, 1998; Sauer, 2004), drop-out rates are still considerably large[1].

Novel evidence seems to point in the same direction. Under highly plausible exogeneity conditions[2], Bulman et al. (2021) find that large cash transfers have no effect on enrolment of low-ses students. At the same time, the reevaluation of large programs such as the Pell Grants (Eng and Matsudaira, 2021) using Regression Discontinuity techniques, such as the ones in this paper, revealed smaller effects in student outcomes than previously obtained (Deming and Dynarski, 2010).

Specific to the Brazilian case, Vieira and Arends-Kuenning (2019) explore the adoption effect of Affirmative Action Policies by public universities in Brazil on the enrollment rate of the beneficiary groups. Their results suggest that race-oriented policies may be more effective in increasing enrollment of students from lower socioeconomic strata than policies neutral in ethnicity with the same income

[1]A potential explanation to this phenomena is the role of information revelation, treated in detail in Chapter 1.

[2]Bulman et al. (2021) study the effect of lottery prizes on student attainment.

eligibility requirements. Herskovic and Ramos (2019) model the welfare impacts of Affirmative Action Policies in welfare using Brazilian data. Even with higher dropout rates for policy-eligible minorities, their results indicate that reaching low-income candidates is effective in reducing the inter-generational persistence of income inequality and improving well-being and aggregate production. Affirmative action increases the number of admitted socially disadvantaged students but distorts incentives for educational investments. They find evidence that introducing fees at public universities would not lead to a loss in well-being.

Studying the same loan program, Barahona et al. (2022) uses a general equilibrium model to measure the effects of the subsidized loans on tuition prices, enrolment and completion. They find that increasing the supply of subsidized loans increases tuition fees for all students. The overall effect of the policy still raises enrolment relative to not offering loans at all - an effect for which I do not find support.

Clearly, there is still much to be learned about who benefits from large education financial aid programs, even though there is a large body of literature dedicated to the topic. As evidence becomes conflicted, it is fundamental to dig deeper into the (implicit) assumptions involved in the identification of such effects and given these, what is their correct interpretation. This paper attempts to stretch this small but treacherous battlefield.

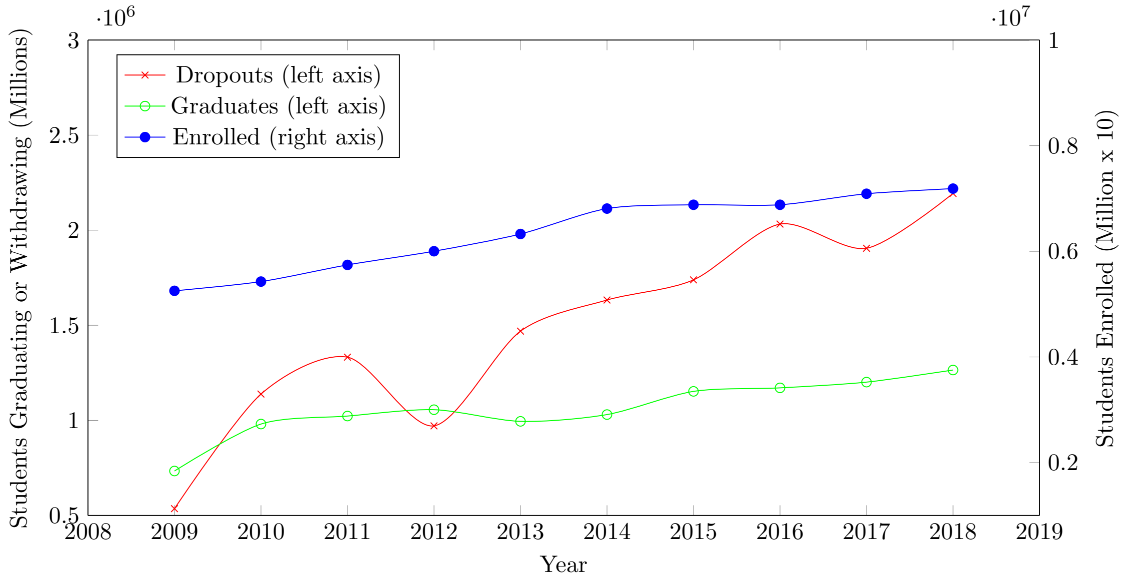
The remainder of this paper is structured as follows: Section 2.2 presents the policy environment and data, Section 2.3 develops the estimation procedure, Section 2.4 discusses the estimation results and Section 2.5.

2.2 Policy background and Data

Brazil is an ideal environment to study the impacts of different types of financial aid. With a general quality gap between public and private schools, access to higher education is highly unequal. Particularly in the last six years, Brazil faced a substantial increase in University dropout rates. Around 50% of every cohort ends up not graduating. Yet, access to financial aid and overall access to higher education has gone up over the last 20 years. In 2019, 2,448 institutions hosted 8.5 million students, from which 3.5 million were new enrolments. Currently, there are 286 public institutions active versus 1,962 private institutions and they are highly heterogeneous. Figure 3 illustrates the size of the higher education sector in student numbers over time.

The expansion of access to higher education started in the late 1990s, which

Figure 3 – Numbers of students enrolled, graduating and withdrawing from higher education by year



Note: Absolute figures of students enrolled, graduating and withdrawing from higher education by year as taken from ACoHE.

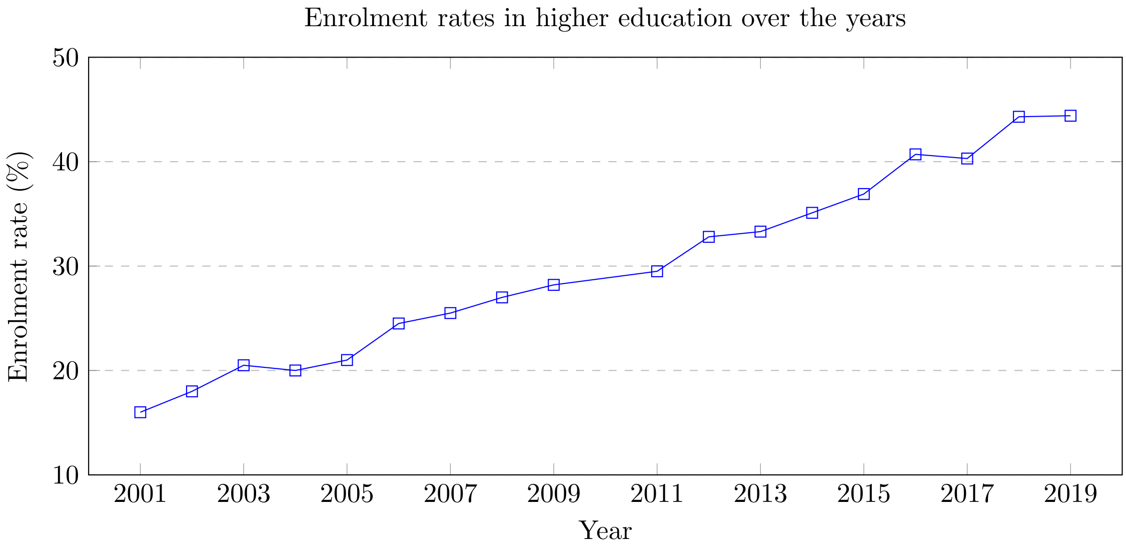
coincided with Brazil's economic opening. The rate of enrolment in higher education for young adults aged 18-24 reached its peak at 50% in 2018 (Figure 4). The rate of enrolment amongst young adults in urban areas is considerably higher than in rural areas, and the gap between the two becomes wider as time goes by (Figure 5). The same applies to income percentile comparison; the gap between the richest and the rest of the income distribution has gotten wider since 2016. In the meantime, the supply of higher education has increased, and government programs that target underrepresented minorities have been offered.

To reduce inequality of opportunity in higher education, the Brazilian government launched the "University for All" initiative in 2012. Before that, some public universities had already adopted Affirmative Action in admissions, beginning in 2003. The Brazilian higher education environment presents a diverse set of affirmative action policies that target different types of low-income students coming from public schools. Figure 6 illustrates the eligibility cutoffs for each program across 2 running variables: income and performance in the National High School Exam (ENEM).

Figure 6 takes into consideration 6 of the largest higher education government-sponsored programs j awarded according to their own specific income and exam score cutoff distribution. Exam cutoffs vary at the degree and institution level, and Figure 6 illustrates the general ordering of the average program cutoff[3]. From right

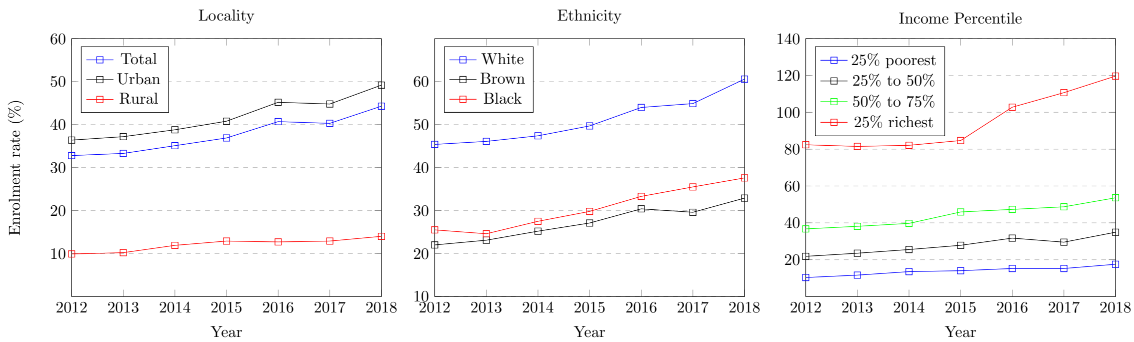
[3]Cutoffs are obtained at the degree-institution-program level, here I average them over degree-

Figure 4 – Enrolment rates over time



Note: Percentage of new enrolments in higher education over population aged 18 to 24 by groups. "University for All" initiative was implemented in 2012.

Figure 5 – Enrolment rates by group

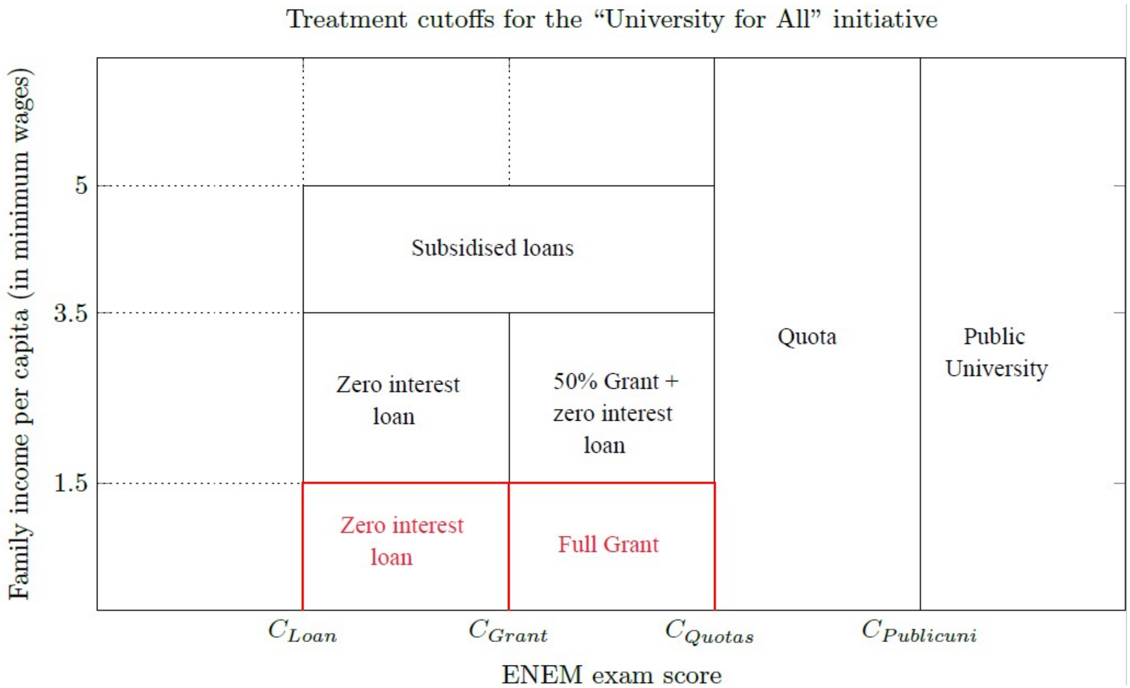


Note: Percentage of new enrolments in higher education over population aged 18 to 24 by groups. "University for All" initiative was implemented in 2012. Since then, income and locality inequality gaps in enrollment seem to have increased.

to left, the first treatment is characterized by receiving a university education, free of tuition, in a public university, after participating in the regular selection process. The second treatment is characterized by receiving the same university education, free of tuition, in a public university but through a place reservation program (also known as quotas). Students are eligible to enter university through quotas if they completed the previous 3 years of high school in a public school. Through this place reservation program, 50% of all places in public universities are reserved for students from public schooling. Below the quota exam score cutoff, students can no longer receive publicly provided higher education. Then, according to their income eligibility, they receive different types of government-sponsored grants and loans.

institution pairs.

Figure 6 – Policy Eligibility



Note: Eligibility cutoffs for ENEM exam score and income (measured in minimum wages). Quadrants represent possible beneficiaries of the program (treatment receivers). The type of treatment/program such an individual may receive is specified in each quadrant.

Students who perform above the C_{Grant} cutoff in ENEM can be assigned to 3 different treatments. The Brazilian Government offers grants through the PROUNI program. The PROUNI offers grants of up to a 100% of tuition fees for low-income students attending eligible private institutions. If their family income per capita amounts up to 1.5 minimum wages a month, they are assigned a grant of 100% of the eligible private institutions' tuition fee. If their family income per capita is greater than 1.5 but less or equal to 3 minimum wages a month, they are assigned a grant of 50% of the eligible institution's tuition fee.

If a student's family income p.c. is greater than 3 but less or equal to 5 minimum wages a month, they no longer qualify for a grant even if they score above the grant exam cutoff, but could qualify for a subsidized loan for tuition fees (also called P-FIES).

Moving further down the cutoff score distribution, the lower ENEM performance cutoff C_{Loan} classifies students for different state-sponsored loans, in a program called FIES. FIES is a program that grants low-interest loans for low-income students to study in eligible private institutions. FIES amortizations begin upon completion or withdrawal of the selected degree. There are 2 types of loans awarded through this program, FIES and P-FIES. The former grants zero-interest rate loans, while the latter grants loans with subsidized rates of 2-4% a year. To apply for zero-rate loans, the candidate must have a monthly gross family income per capita of up to 3 minimum wages. If their family income p.c. is greater than 3 but less or equal to 5 minimum wages a month, they are assigned the subsidized interest rate loan for tuition fees (P-FIES). Finally, a student that qualifies within the zero-interest loan rule and scores above the grant cutoff can jointly take up the partial grant and the loan for the reminder of the tuition fee amount.

To illustrate the size of each financial aid program, in the 2015 cohort 142,612 students received full tuition fee grants and 252,430 students received grants. At the same time, 382,485 students received loans at interest rates ranging from 0 to 4% per annum from FIES.

In this paper, I look at the effects of two programs that are part of “University for All”: PROUNI grants and FIES (the red area in Figure 6), which offer partial grants and zero-interest loans (respectively) with the same eligibility criteria, in this case, family income up to 3 minimum wages.

Program application and treatment assignment

Students apply for financial aid online, and program availability is advertised nationwide through television and social media. At the moment of application, the candidate must pick 2 preferred institution/degree combinations. If a student’s ENEM score is above this institution/degree threshold, they receive an offer for financial aid. Student financial aid selection then proceeds as follows 2 rounds of offers:

1. If a student receives an offer in the 1st round, they must accept or reject it at round 1. Students that did not receive an offer, proceed to round 2.
2. If a student did not receive an offer in the second round, then the procedure follows as 2. However, this is the final round of offers.

The Brazilian government decides the number of grants or loans that can be assigned to each institution/degree combination using criteria from the National Education Plan, which is updated every 5 years and aims to prioritize areas of knowledge that are strategic for the country, promote equal opportunities within different regions and provide high-quality education. Neither students nor institutions can determine perfectly forecast C_j.

Once students take ENEM and apply for the program, best-ranked students are assigned financial aid until exhaustion of program funds as previously determined. In this case, treatment assignment depends on the exam scores S in a deterministic way. At this moment, I assume that students choose the program offer that optimizes their present value, yielding preferences over financial aid programs of the form:

$$\textit{PROUNI 100\%} \succ PROUNI\ \ 50\% \succ FIES \succ \textit{P-FIES} \tag{2.1}$$

This means that students who qualify for all financial aid programs - who hold an exam score above the highest cutoff and qualify below the 1.5 minimum wage income threshold, would always prefer full grants to other types of financial aid.

Data

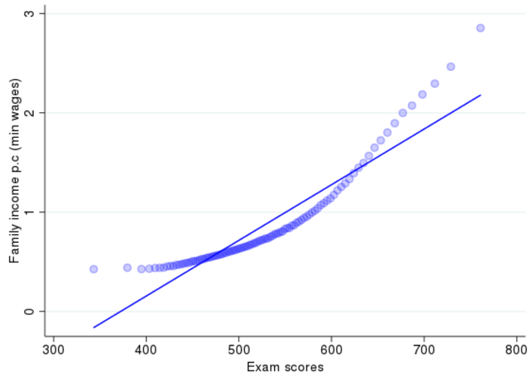
This paper uses individual-level data from the Annual Census of Higher Education (ACoHE). With the ACoHE, it is possible to monitor the progress of each Brazilian student since 20, as they are assigned an identification number kept throughout their university career. This allows me to link the socioeconomic characteristics of individuals and their families as well as government programs they benefit from to the exact date on which they enroll, graduate or leave their chosen university degree.

Other complementary datasets allow me to observe student performance prior to and during their trajectory in higher education. From the National High School Examination Dataset (ENEM), I can recover students' performance in a standardized test taken by most students at the end of high school. This exam serves as an entry exam for public and most private universities and is required for applying to government-sponsored financial aid programs. The final cohort used in this study consists of every individual who took the National High School Exam (ENEM) in 2015, tracked over a period of 5 calendar years using the same algorithm applied in Chapter 1. Appendix A details the steps to obtain the cohort panel used in this study.

Cohort Descriptives

The policy assignment rule awards students on the basis of their performance in a standardized exam. A natural question, then, is what drives students' exam performance? A first thing to note is that indeed exam performance and family income are correlated, as illustrated in Figure 7. Even amongst those who qualify for financial aid, there is nuance in income variation that may be driving student performance. Appendix C.1 contains balance checks between financial aid takers

Figure 7 – Family Income and ENEM exam performance



Note: Binned scatter plots of ENEM exam scores vs. family income per capita (4,986,038 observations in 99 bins).Minimum ENEM score is 0 and possible maximum score is 1000. Family (household) income per capita is measured in minimum wages. Fitted lines correspond to OLS regressions of the y-residuals on x-residuals.

and non-takers as well as estimates of the determinants of ENEM exam performance. Exam performance is highly correlated with a series of individual characteristics, most importantly, parents education seem to be a significantly large determinant of exam performance. This highlights the need for an econometric approach that reflects the heterogeneity of individuals at different location of the exam score distribution.

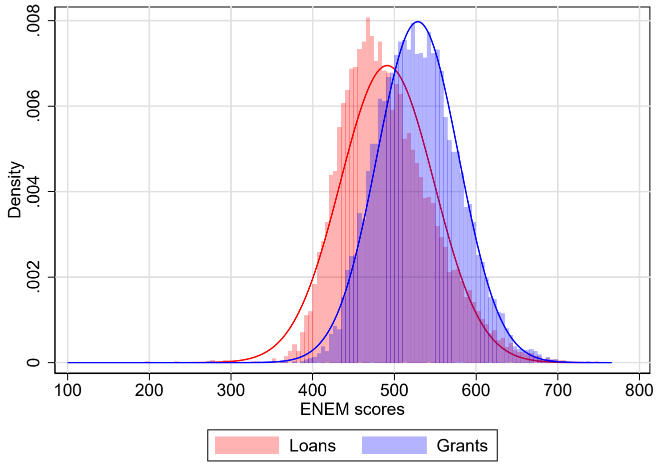
2.3 The Regression Discontinuity Design

The standard Regression Discontinuity Design (RDD) estimates the causal effect of an intervention by exploiting a cutoff or threshold in a continuous assignment variable. Individuals on one side of the cutoff receive the treatment deterministically or stochastically, while those on the other do not. The regular assumption is that individuals just below and just above the cutoff are comparable, and any difference in outcomes can be attributed to the treatment effect.

In contrast, a cumulative multi-cutoff RDD extends the standard RDD by allowing for multiple thresholds, or cutoffs, across different groups or contexts. This approach leverages the cumulative and non-cumulative information from these multiple thresholds to enhance statistical power and generalizability.

The research goal is to identify the effect of loans and grants on student attainment. My approach utilizes overlaps on the financial aid eligibility thresholds over 2 dimensions. For each type of financial aid program j, there will be $p = 1, ..., N$ degree-institution cutoffs. Figure 8 displays the empirical distribution of admission cutoffs by type of program. Degree-institution cutoffs are on average higher for grant awards than for loan awards. The distribution of cutoffs associated with grants also

Figure 8 – The distribution of degree-institution admission cutoffs by program type



Note: N=6829 degree-institutions. Histogram bin width=10. Solid lines represent the approximated PDFs.

display higher kurtosis meaning that grant awards tend to be selected on more a more homogeneous exam score basis than loans. Nevertheless, the Figure shows a substantial overlap of cutoff scores across the different types of financial aid award.

As exam score cutoffs vary for each institution and degree combination (j, p), so varies the probability of receiving a specific type of treatment. The heterogeneity of admission thresholds and treatments allows me to explore Multi-cutoff RDD techniques (Cattaneo et al., 2018).

i) To obtain estimates that take into account the 2 different types (or doses) of financial aid available I use Cumulative Multi-Cutoff RD to obtain treatment effects at each cutoff $c_{j,p}$.

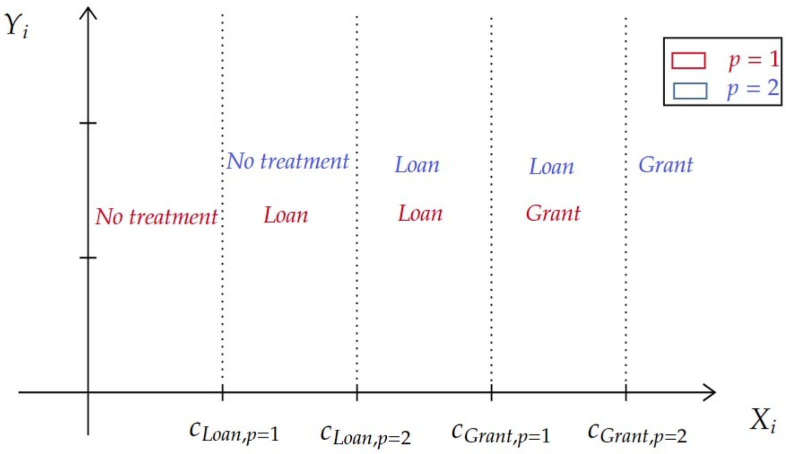
ii) Then, I obtain average treatment effects of financial aid (loans and grants are assumed homogeneous) over the distribution of Exam scores.

iii) I use the results from (ii) as empirical weights combined with (i) to obtain the overall treatment effects of grants and loans.

Set-up

First, some notation needs to be introduced. Let $Y_i(0), Y_i(1)$ and $Y_i(2)$ represent student i outcomes for receiving no treatment, being awarded a loan and being awarded a grant, respectively. Let X_i represent the running variable, ENEM test scores, $T_i = \mathbb{1}[X_i \geq c]$ represent the treatment rule and D_i denote actual treatment

Figure 9 – An illustration of the two-treatment two-program multi-cutoff case



Note: Color red denotes the treatment an individual who falls within the explicit $\underline{x} < X > \overline{x}$ interval would get in program 1 and blue denotes the treatment an individual who falls within the explicit $\underline{x} < X > \overline{x}$ interval would get in program 2.

assignment[4]. Define $\boldsymbol{C}$ as the set of ENEM score cutoffs $c_{j,p}$ to be awarded a financial aid program j at the degree-institution p.

Next I separately defined the estimators used in step (i) and (ii).

Financial-aid program cutoffs, j

The cumulative multi-cutoff is a special case of multi-cutoff RDD in which an individual observation can be exposed to multiple cutoffs. For example, take a two-treatment two-program case illustrated in Figure 9. Assume there is an individual observation with $X_i = x$ that may be above $c_{Loan,p=1}$ and below $c_{Loan,p=2}$.A cutoff-specific analysis at $c_{Loan,p=1}$ would treat this individual to a loan, while an analysis at $c_{Loan,p=2}$ would categorize this observation in the control group, potentially correlating effects across cutoffs and complicating treatment interpretation.

Cattaneo et al. (2018) provide a way to address this by computing a midpoint between the two cutoffs using a median, for example. This midpoint is then used to reassign treatment and control groups. Another option is separating control and treatment groups through optimal bandwidth procedures.

A more formal estimation definition follows from standard RD arguments. Denote the treatment categories as, $j = 0, 1, 2$ for a fixed p, the program j-specific treatment effect is given by:

[4] There is imperfect treatment assignment in this setting, characterizing a Fuzzy RDD.

$$\tau_j = \mathbb{E}\left\{Y_i(d_j) - Y_i(d_{j-1}) \mid X_i = c_j\right\} \tag{2.2}$$

Thus, $\tau_j(c)$ maintains the same interpretation of standard RD but with treatment effects now cumulative. At the end of the estimation procedure, the econometrician obtain as separate $\tau_j(c)$ for each c_p. It follows that individuals exposed to multiple cutoffs can be used to estimate consecutive treatment effects. For instance, a unit with $c_j < X_i < c_{j+1}$ receives dosage t_j, serving as a treatment unit for c_j and a control for c_{j+1}. Cutoff-specific estimators may be correlated, but dependence weakens as bandwidths diverge.

Where each individual falls around each cutoff may be due to its own individual characteristics, observable and non-observable. So each cutoff interval could be interpreted as an individuals type for which $E[Y_{1i} - Y_{0i}|X_i]$ recovers the average treatment effect conditional on X_i. Under Bertanha (2020), average treatment effects are portrayed by equation 2.3, obtained by first identifying local average treatment effect τ_{jp} (at each each cutoff) and weighting it by ω_{jp}. This approach uses the idea of self-revelation, as introduced by Dale and Krueger (2014). The cutoff region around where each observation falls is informative of students characteristics, observed and non-observed.

Because of imperfect compliance, additional assumptions are needed. The propensity score function of being treated must preserve a discontinuity at some point $X = x$.

$$\tau_p = \sum_{j=1}^{J} \omega_{jp} \tau_{jp} \tag{2.3}$$

Differently from the usual normalizing cutoff assumption, when ω is simply obtained by the probability of each cutoff occurring, now $omega$ weights are a counterfactual probability mass to be jointly estimated according to the procedure below.

Estimation is implemented in 2 steps. First, for every (j, p) pair, a cutoff specific treatment effect, τ_{jp}, is obtained using the optimal bandwidth procedure described in Calonico et al. (2019) and Bertanha (2020). Separately, non-cumulative cutoff RD analysis is conducted across all cutoffs, assuming treatments j are homogeneous. From this step, one can obtain regular “normalized cutoff weights” weights that are used in an iterative procedure to solve for τ_p over the support of X_i and propensity of individuals to get treated once they are assigned these same cutoffs. Finally, program-specific average treatment effects over j, described by τ_p, are obtained.

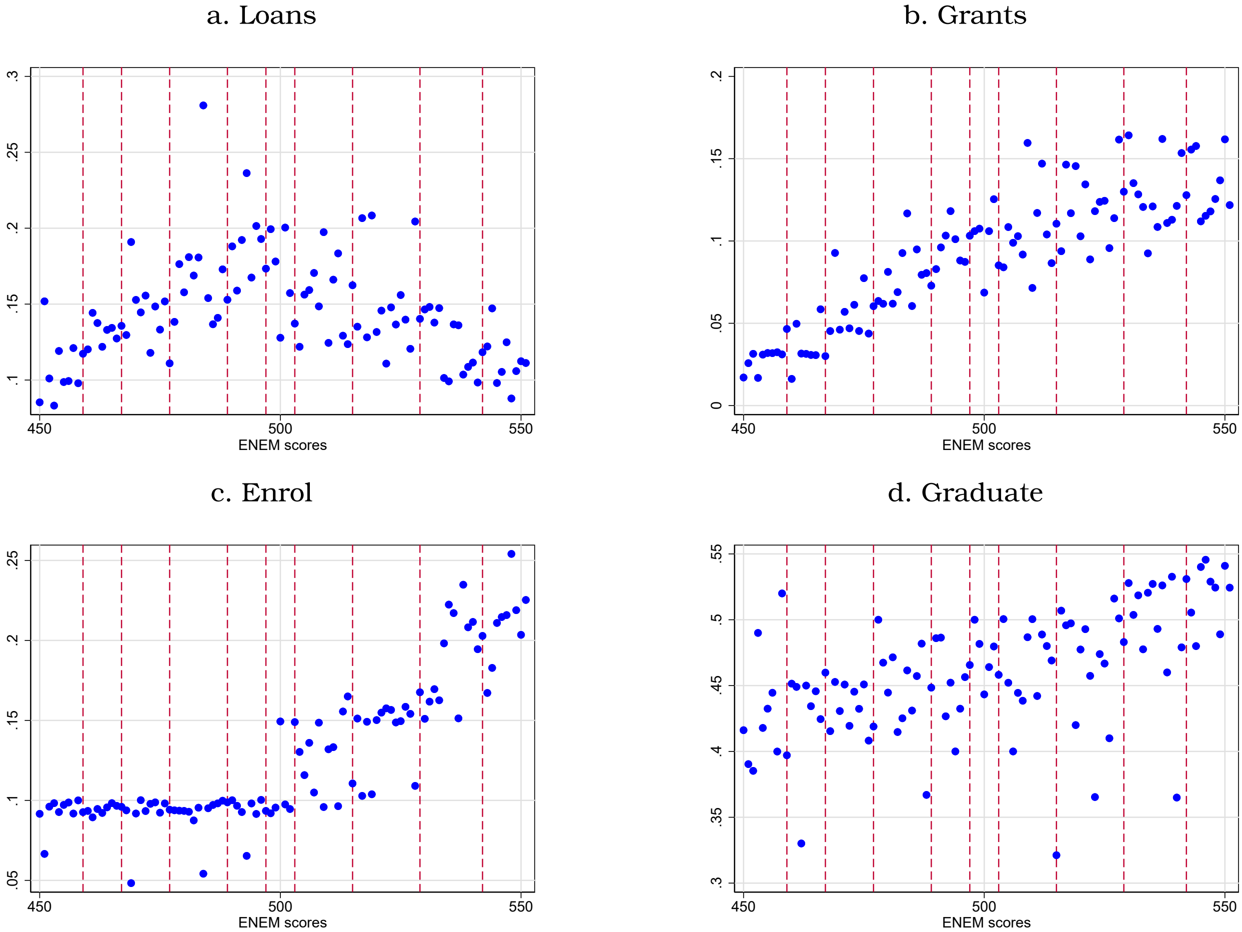
2.4 Results

This Section presents the results from the RD regressions. For the sake of brevity, I leave the discussion of balance and manipulation checks to Appendix C. For now, suffices to say that the sample is well-balanced for all observables and test do not support the hypothesis of cutoff manipulation.

Figure 10 illustrates how treatment and outcomes respond to discontinuities along the running variable, ENEM test scores. The first 2 sub-figures represent the probability of receiving treatment for a group of students with similar test scores. The probabilities of receiving either treatment display visible discontinuities at the degree-institution cutoffs. The relationship between the probability of receiving a loan is concave in ENEM scores reaching its maximum around the means student score (501.98) and decreasing after that. This is possibly due to the fact that the higher the student's performance in the exam, the higher the probability of also being awarded a grant and never taking the loan. One can confirm this relationship through sub-figure (b): the probability of receiving a grant is increasing throughout the ENEM score interval, but the discontinuities become less contrasting as scores grow - meaning that the difference between the probability of not being awarded a grant is smaller at the upper end of the score distribution. In the bottom row, sub-figures (c) and (d) represent the probabilities of enrolling in higher education and graduating (conditional on enrolment). Even though the probability of enrolment displays clear discontinuities at the cutoffs below the average ENEM score, these are relatively small compared to the jumps of almost 5 p.p. in enrolment probabilities for those placed at cutoffs above the mean score. This is not the case, however, for the probability of graduation: although the probability of graduating and exam scores are positively correlated, the discontinuity across cutoff intervals is less clear than for the other outcomes and treatments.

Next, I examine the first-stage regression results associated with the first row of sub-figures in Figure 10, shown in Table 6. The first stage estimation is run p for each treatment resulting in different estimates for each degree-institution cutoff. Table 6 displays results that corroborate with the previous visuals. The effect of crossing a cutoff on the probability of receiving a loan starts positive and significant and becomes non-significant as the cutoff approaches the mean exam score, after that it becomes negative but still significant. First-stage results for receiving a grant

Figure 10 – Treatments and Outcomes Across cutoffs



Note: Binned scatterplots averaged over 100-students. Exam score interval selected for exposition purposes. Sub-figure (a) measures the probability of being awarded a subsidized loan; (b) measures the probability of being awarded a grant; c) measures the probability of enrolling in higher education the year after the exam, and (d) measures the probability of graduating within 5 years of enrolment (conditional on enrolment). Dashed lines represent degree-institution cutoffs.

show significant and positive effects on the probability of receiving the award after crossing a cutoff. These cutoff specific coefficients appear more constant across different cutoffs (oscilatting at around 1.6 p.p.) than the ones for loans. Overall, first-stage results appear significant, ensuring the strength of the first-stage.

the impact of loans and grants on University enrolment and completion. According to Table 8, the overall effect of grants on enrolment is positive and significant; the estimated coefficient indicates a small increase of 0.04 percentage points (p.p.) in enrolment due to grants when compared to loans. In contrast, the overall effect on completion significantly negative, suggesting that grants may reduce the likelihood of completion by 1.1 p.p. When analyzing specific thresholds, there are mixed results. At the 10th-ranked degree-institution threshold, there is no significant effect on either enrolment or completion. The same is observed at the 6000th degree-institution threshold. Grants do not seem to exert a statistically significant effect on enrolment for the majority of thresholds when compared to loans. On the other hand, receiving a grant seem to exert statistically significant negative effects on completion when compared to loan, across most degree-institution thresholds. Overall, these results suggest that grants have a modest positive effect on enrolment but may negatively affect completion rates.

Table 6 – First-stage results

	Loans	Grants
at $c = 459$	0.024**	0.016**
	(0.009)	(0.006)
at $c = 477$	0.027**	0.012**
	(0.011)	(0.005)
at $c = 503$	0.003	0.017**
	(0.004)	(0.007)
at $c = 571$	0.001***	0.009*
	(0.000)	(0.004)
at $c = 642$	-0.014**	0.013*
	(0.004)	(0.005)
Min Obs within bandwidth	301	254
Max Obs within bandwidth	1,450	1,118

Note: Coefficients displayed are evaluated at arbitrary values for exposition purposes. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, Tables 7 and 8 show the estimates of Local and Average Treatment effects by financial aid program. I start by discussing the estimates of τ_{jp} for grants. Grants' Local treatment effects on enrolment are mostly positive, significant and reasonably constant across cutoffs. Over all cutofffs, being eligible for a grant increase student enrolment by 4.1 p.p. But even though receiving a grant may drive people into education, being eligible for grant doesn't seem to be enough to guarantee degree completion as evidenced by statistically insignificant overall treatment effects on graduation. In fact, receiving a grant may reduce the probability of completion by 0.2-0.05 p.p. for those in the middle of the exam score distribution.

Finally, I examine the estimates of Local and average treatment effects for being eligible for a subsidized loan on student enrolment and completion. Table 8 displays a reverse relationship than that of grants. Being eligible for a loan does not seem to significantly affect enrolment. On the other hand, it has a positive and statistically significant effect on completion. Being eligible for a loan increases the probability of overall graduation by 3.1 % p.p. while its effects are mostly concentrated on those who fall on the lower end of the exam score distribution.

Table 7 – Treatment Effect of Grants: Enrolment and Completion

	Enrolment	Completion
Overall effect τ_j	0.041***	0.007
	(0.001)	(0.004)
τ_{jp} at $c = 459$	0.001**	0.010
	(0.000)	(0.009)
τ_{jp} at $c = 477$	0.003**	-0.005*
	(0.001)	(0.002)
τ_{jp} at $c = 503$	0.052***	-0.002*
	(0.001)	(0.001)
τ_{jp} at $c = 671$	0.047***	0.007
	(0.003)	(0.011)
τ_{jp} at $c = 642$	0.041*	0.013*
	(0.018)	(0.007)
Min Obs within bandwidth	347	286
Max Obs within bandwidth	1,623	1,215

Note: The Local Average Treatment Effects at each cutoff, τ_{jp} are evaluated at arbitrary values for exposition purposes. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8 – Treatment Effect of Loans: Enrolment and Completion

	Enrolment	Completion
Overall effect τ_p	0.004	0.031***
	(0.007)	(0.003)
τ_{jp} at $c = 459$	0.002*	0.032*
	(0.001)	(0.016)
τ_{jp} at $c = 477$	0.001	0.042**
	(0.001)	(0.017)
τ_{jp} at $c = 503$	0.007	-0.001*
	(0.013)	(0.000)
τ_{jp} at $c = 571$	0.005	-0.001
	(0.009)	(0.003)
τ_{jp} at $c = 642$	0.003	0.002
	(0.019)	(0.007)
Min Obs within bandwidth	303	226
Max Obs within bandwidth	1,587	1,115

Note: The Local Average Treatment Effects at each cutoff, τ_{jp} are evaluated at arbitrary values for expositional purposes. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.5 Discussion

The two largest higher education financial aid programs in Brazil, PROUNI and FIES, benefit similar audiences. Nevertheless, the two present very different effects. State-sponsored financial aid is distributed based on students' rankings in the National High School Exit Exam (ENEM), with assignment thresholds determined for each combination of financial aid type, degree program, and educational institution. Using multi-cutoff regression discontinuity methods, I estimate the average treatment effects separately by the type of financial aid and analyze how these effects vary across the distribution of student exam scores.

I find that while the grant program increases general enrolment, this is mostly for those at the higher end of the test score distribution and has no impact on student retention. Meanwhile, the loan program has no effect on enrolment but increases student retention. Although both programs were created to relax current budget constraints that may prevent certain students to pursue higher education, they are virtually different. A grant is an unconditional transfer given to the student at the moment of enrolment - making it cost-free to enrol in a degree. The loan, even though awarded in a similar manner, introduces a penalty in case of discontinuation of studies. The student must start the loan amortization payments as soon as it stops studying, whether due to conclusion of the degree or withdrawal,

Are loans more effective than grants in retaining students until graduation? Or are loans only attractive to students that place a higher value to education? If the cutoffs around which different students fall are a reflection of the student's decision process, then using this empirical strategy may help control for the role of unobserved heterogeneity in students' preferences, risk aversion and effort levels. The Iterated Deferred Acceptance mechanism used to assign student to programs provides and the information revelation that happens in its different steps provide a setting for increased sorting, in a way that the assigned cutoff may be revealing of students unobserved characteristics. If that is the case, than we can rule out selection and determine that grants are more effective than loans in increasing enrolment but loans are more effective in securing degree completion, conditional on enrolment.

Nonetheless, the evidence from this paper highlights the presence of heterogeneity in treatment effects across the distribution of exam scores. Interposing an argument in disfavour of the standard "normalize-and-pool" Multi-cutoff Regression Discontinuity techniques. Particularly, I find evidence against the positive enrolment effects of subsidized loans documented by Barahona et al. (2022) using this standard RD approach.

3 SEARCH AND REALLOCATION IN THE COVID-19 PANDEMIC: EVIDENCE FROM THE UK***

Abstract

The impact of the pandemic on the UK labour market has been extremely heterogeneous across occupations and industries. Using novel data on job search, we document how individuals adjust their job search in response to changing employment patterns across occupations and industries in the UK. We observe that workers changed their search direction in favour of expanding occupations and industries as the pandemic developed. However, non-employed workers are more attached to their previous occupations and workers with low education are more likely to target declining occupations. We also observe workers from declining occupations making fewer transitions to expanding occupations than those who start in expanding occupations, despite targeting these jobs relatively frequently. This suggests those at the margins of the labour market may be least able to escape occupations that declined during the pandemic.

3.1 Introduction

It is well known that different sectors of the economy react differently to the business cycle. The Covid-19 pandemic highlighted this feature very clearly.[1] For example, the pandemic and related lockdown measures applied in the UK meant that, by the end of 2021, the Accommodation and Food industry lost close to 20% of its pre-pandemic employment. At the same time, Public Administration employment grew by about 8%. A similar feature occurred across occupations. Elementary occupations lost about 10% of their pre-pandemic employment, while Administrative occupations gained 10% (see Figure 14, below). These differences left a large number of individuals, mostly from the worst affected sectors, unemployed or at risk of

***This chapter is joint work with Carlos Carrillo-Tudela, Alex Clymo, Annette Jackle, Ludo Visschers and David Zentler-Munro and it was originally published in Labour Economics, vol.81, April 2023.

[1]A unique feature of the pandemic recession relative to other recent recessions was the speed at which shocks impacted the different sectors and the unprecedented policy measures governments implemented to ameliorate the impact.

unemployment. As evidenced by the labour shortages afflicting many economies after the pandemic, the speed and strength of the economic recovery not only depends on renewed vacancy creation but also on workers' willingness and ability to reallocate from harder hit sectors to those that are booming, as well as firms' willingness to hire them.[2]

In this paper we investigate how workers adapted to the rapid structural shifts in demand from different industries and occupations during the Covid-19 pandemic. We tackle two questions in turn. Did workers adjust their job search during the pandemic and target jobs in expanding occupations and industries? Did any adjustment translate into labour reallocation across sectors? A key contribution and innovation of the paper is that we collected data, through the COVID-19 Study of the UK Household Longitudinal Study (UKHLS), on which occupations and industries job searchers were targeting during the second half of 2020 and January 2021. Since we collected data at different points during the pandemic, these data allow us to investigate the extent to which job seekers were reacting to the evolving occupation and industry differences arising from the pandemic and lockdown policies. A further advantage is that we are able to merge for each individual surveyed rich information about the previous labour market history of respondents. Thus our main analysis takes into account not only observable but also unobservable individual characteristics. We complement these data with the UK Labour Force Survey (LFS) in order to investigate the evolution of aggregate job search during the first two years of the pandemic as well as the likely implications of individuals' job search behaviour for aggregate reallocation flows.

Our starting point is to document the heterogeneity in the shocks to employment by occupation and industry, before turning to the responsiveness of job search to these shocks. We observe large heterogeneity in employment changes across occupations during the pandemic in contrast with that seen during the Great Recession, where occupations experienced less dispersed employment changes. Moreover, occupation shocks are not simply a reflection of underlying industry shocks. We show that declining occupations saw employment falls for occupation specific reasons which were not driven by changes in between-industry composition. These shocks have tended to accelerate the longer term trends in the labour market by industry and occupation. An important question is therefore whether workers adjusted their job search to the specific nature of the shocks during the pandemic, or based their search on longer-term trends.

At the extensive search margin, we observe that unemployed workers from

[2] Liu, Salvanes and Sorensen (2016) present evidence that the degree of mismatch between workers and jobs is a key driver of the scarring impacts of recessions.

declining sectors were more likely to quit their job search in the first half of 2020 and were more likely to resume job search as the economy recovered. At the intensive search margin we document three novel facts. (i) Workers changed their direction of search in favour of expanding occupations and industries as the pandemic progressed, which suggests job searchers were responding to occupation-wide and industry-wide conditions.[3] Nevertheless a large proportion of workers continued targeting declining occupations and industries. (ii) The individuals most likely to target declining occupations were those at the margins of the labour market: those with the lowest education levels and, most significantly, those coming from declining occupations and industries due to attachment to previous jobs. (iii) There is also a substantial mismatch between targeted and realised transitions. Among those targeting an occupation switch, the proportion of workers actually making an occupation transition into expanding occupations was substantially lower than the proportion of job seekers targeting a switch into an expanding occupation, particularly for those individuals coming from declining occupations. This suggests substantial impediments to reallocation across occupations during the pandemic.

Our analysis further shows that worker reallocation was occurring at an aggregate level, evidenced by the large rise in net mobility across industries which was double the level observed during the Great Recession. This finding is important in light of the Job Retention Scheme (or "furlough") introduced by the UK Government at the start of the pandemic.[4] Some commentators raised concerns that the furlough scheme, which mediated the nature of the pandemic shocks on occupations and industries in the UK, was going to hold back Schumpeterian forces of "creative destruction" associated with labour market churn and reallocation.[5] The balance of evidence suggests the furlough scheme had a stronger impact in limiting job destruction than in holding back job creation or mobility across industries.

Across occupations this dynamism, however, was much more subdued. We find that net mobility flows across occupations remained broadly stable in line with the experience during the Great Recession. This is driven mainly by a combination

[3]This is also suggested by the fact that job searchers who were in occupations that expanded during the pandemic seeked to switch occupations less frequently than those in declining occupations. The growing occupations were those which typically require higher skills, offer higher wages and provide more opportunities to work from home.

[4]The JRS, or "furlough" scheme, provides furloughed workers with 80% of their pre-furlough wages, up to a limit of £2,500 per month, on the condition they remain on the employer's payroll but no longer working. At peak usage (April 2020) around one third of the UK's workforce was fully furloughed.

[5]"The scheme could even be economically damaging if it dissuades people from searching for new jobs or helps 'zombie' firms to survive for longer. Reallocation of workers and capital to more productive sectors with better prospects is in normal times an important vehicle for economic growth and retaining defunct employer-employee relationships risks slowing this down", Institute For Government (2020).

of workers in declining occupations continuing to target their previous occupation, and not being able to access targeted jobs in expanding occupations. This suggests a pattern of segmentation, where there was a strong attachment to previous occupations during the pandemic and those targeting an occupation change found it hard to break into expanding, higher skill and better paying occupations unless they start from one. As this segmentation did not occur across industries, policies that attempted to force reallocation from the declining low skilled jobs to expanding high skilled ones would appear to have little effect in the short run. Instead, medium term re-training policies would be more effective. This is important in light of the labour market policies the UK government enacted at several stages throughout the pandemic to incentivise job seekers to search for employment outside their occupations.[6]

The above evidence shows that workers' search behaviour *reacts* to employment changes by industry and occupation. This naturally implies that their behaviour must then *contribute* to the evolution of the labour market. The aggregate trends show that the pandemic initially discouraged job search among those who lost employment due to the lockdown measures. There was a sharp rise in the number of individuals out of the labour force flowing from employment and unemployment to inactivity. This resulted in a much larger increase in inactivity than experienced in the Great Recession, and provides another clear indication that the extensive margin of job search was a relevant channel of adjustment during the pandemic. During the second half of 2020, however, more individuals re-engaged with job search as vacancy posting began to recover, resulting in a higher unemployment rate and a slower rise in inactivity. The subsequent drop in unemployment during 2021 then led to the recovery in the employment stock. During the recovery, job-to-job transitions also increased back to and even above their pre-pandemic level. However, the recovery was marked by a divergence in gross reallocation across industries and occupations, with gross mobility across industries recovering more rapidly than in the Great Recession, while gross mobility across occupations stagnated.

Related Literature

This paper contributes to the large literature that developed during the Covid-19 pandemic analysing the impact of lockdowns and other social distancing measures on labour market outcomes. Like our paper, Albanesi and Kim (2021) and Jones,

[6]These policies were implemented through re-training subsidies or unemployment benefits cuts to individuals who do not actively search for jobs outside their occupations after three months into their unemployment spell. These types of policies are not new, however. The German Hartz reforms, for example, imposed severe penalties on the level of unemployment benefits individuals can claim if they reject a suitable job offer irrespectively of the industry/occupation.

Lange, Riddell and Warman (2021) investigate aggregate changes in the stocks and flows of inactive, unemployed and employed workers in the US and Canada respectively. A common finding in these studies and ours is that there was an initial increase in outflows from both employment and unemployment to inactivity in the onset of the pandemic, followed by a reversal of these outflows as the economy recovered. This suggests that, in these countries and the UK, the decision to participate in labour market search is indeed sensitive to aggregate labour market conditions.

Changes in the extensive margin of job search are important as they inform the degree of tightness in labour markets, a key ingredient in search and matching models (see Pissarides (2001)). Faberman, Mueller and Sahin (2022) use the Aggregate Hours Gap measure developed by Faberman, Mueller, Sahin and Topa (2020), which is shown to be highly correlated with the extent of job search, and document that the US labour market was tighter than suggested by more conventional measures based on the unemployment rate.[7] A key distinction between the US and the UK labour markets during the pandemic was the sharp rise in temporary laid-off workers and how they affected the evolution of the unemployment rate. Hall and Kudlyak (2022) and Forsythe, Kahn, Lange and Wiczer (2022) document that this rise led the unemployment rate to jump to 14.7% in April 2020. In the UK, the unemployment rate did not exceed 5% at any point during 2020 and 2021. Similar to temporary layoffs, however, the UK furlough scheme prevented the destruction of search capital by preserving a large number of worker-firm matches and hence keeping unemployment from rising to unprecedented heights.[8]

Our results complement the findings of studies that focus on changes in the intensive margin of job search during the pandemic. For example, Balgová, Trenkle, Zimpelmann and Pestel (2022) using number of applications as a measure of job search intensity, find that in the Netherlands the unemployed searched less intensively for jobs than was the case in the Great Recession. Their contribution is distinct from our focus on the *direction* of job search, which is crucial for understanding how job search both reacts and contributes to shocks that are heterogeneous by sector and occupation. Adams-Prassl, Boneva, Golin and Rauh (2022) instead investigate

[7] Marinescu, Skandalis and Zhao (2021) use data from an online jobs board—this time in the US—to look at the impact of unemployment benefit increases on job search during the pandemic. They find that the Federal Pandemic Unemployment Compensation (FPUC) causes a 3.6% decline job applications but did not decrease vacancy creation. It therefore raised labour market tightness which was otherwise depressed during the pandemic.

[8] Adams-Prassl, Boneva, Golin and Rauh (2020) construct a representative survey in the UK to investigate the characteristics and behaviour of workers on the furlough scheme. They find that workers in occupations and industries where social distancing may be more difficult are less willing to return to work. Furloughed workers in jobs with employer provided sick-pay were 13% points more likely to want to return to work than those without access to sick pay. These concerns likely also play a role in shaping the search behaviour of workers and, consistent with these findings, we find workers have a strong tendency to target higher skill jobs where working from home is easier.

perceived returns to job search among employed and unemployed job searchers in the UK and how these perceptions varied during the pandemic. Among their several findings we highlight that job searchers tend to be over-optimistic in their probability of finding a job. This is in line with our finding that workers appear over-optimistic when targeting jobs in different occupations. This is evidenced be the relatively large discrepancy we document between targeted and realised occupational mobility during the pandemic. In addition, both our Job Search Module and the survey implemented by Adams-Prassl, Boneva, Golin and Rauh (2022) collect information on the desire to change occupations. While they emphasise the role of occupational change due to working from home and other job characteristics, we emphasise the determinants of desire reallocation towards expanding and contracting occupations.

Closest to our paper is Hensvik, Le Barbanchon and Rathelot (2021). These authors investigate how the direction of workers' job search changed during the pandemic in Sweden using a widely used online job search platform. They find that jobs in high home-working occupations, or in occupations where vacancy creation has been more resilient, see increases in clicks per vacancy. This is broadly consistent with the evidence we uncover: for example, that workers target expanding occupations, which generally had higher home-working ability, and this tendency increases over the pandemic (see also Adams-Prassl, Boneva, Golin and Rauh (2022)). Further, Bauer, Keveloh, Mamertino and Weber (2020) use LinkedIn data to investigate changing patterns of job applications by industries in Germany. A key point of departure with these papers is that our approach additionally looks at the realised occupations and industry transitions of workers and compares these to targeted transitions.[9]

Finally, our results inform the growing literature of multi-sector business cycle models based on Lucas and Prescott (1974) in which worker reallocation takes centre stage (see Wiczer (2015), Carrillo-Tudela and Visschers (2020) and Pilossoph (2022)). The large observed discrepancy between targeted and realised occupation/industry mobility and the large proportion of workers that remain attached to their occupations/industries even though these are performing badly suggest that when modelling occupation/industry reallocation one needs to take into account a degree of occupation/industry attachment and the existence of significant impediments to reallocation.

[9]Our analysis also complements Carrillo-Tudela, Hobijn, She and Visschers (2016) who document the cyclical changes of occupations and industry mobility in the UK using LFS data, but do not analyse its evolution after 2012. More recently Pizzinelli and Shibata (2022) compare occupation and industry mismatch indices in the UK and the US during the pandemic. They show that mismatch increased during the pandemic, but this was short lived and smaller than the one observed during the Great Recession. In contrast to our paper, they cannot analyse the occupations and industries targeted by workers but construct their mismatch index based on realised transitions.

The rest of the paper proceeds as follows. Section 3.2 briefly describes the data we use. In order to understand the context of the UK labour market during the pandemic, Section 3.3 examines changes to aggregate labour market stocks and flows. Section 3.4 presents our main results where investigate the nature of jobs targeted by workers and reallocation of workers by occupation and industry. Finally, Section 3.5 discusses future labour market prospects, again with a focus on search and reallocation.

3.2 Data

Our analysis is based on two primary sources: the UK Household Longitudinal Study (UKHLS) and the UK Labour Force Survey (LFS). The UKHLS is a long-term panel of household in the UK that started in 2009, replacing the much smaller British Household Panel Survey.[10] Since 2009, a sample of 40,000 households have been asked questions about the changing characteristics of their household and individual circumstances, including their employment and earnings history. In April 2020 the COVID-19 Study was introduced as a new (temporary) module of the UKLHS.[11] Its aim was to measure the impact of the pandemic on individuals' and households' lives. All UKHLS active sample members $(n = 42,207)$ were invited to complete an online questionnaire and 17,761 individuals completed the first wave. Between April and June 2020 the COVID-19 Study was conducted in monthly waves. From July 2020 to March 2021 it was conducted every two months and after a hiatus the last wave was conducted in September 2021. Given that the UKHLS individuals' identifiers were also used in the COVID-19 Study, one can link the information collected through the latter to each individual's employment and earnings history collected in the annual interviews. In this way we are able to estimate individual wage fixed effects using a Mincer wage equation, compute measures of past employer, occupation and industry mobility as well as know individuals' employment status, occupation and industry during 2019.[12]

[10]University of Essex, Institute for Social and Economic Research. (2021). Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 14th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-15.

[11]University of Essex, Institute for Social and Economic Research. (2021). Understanding Society: COVID-19 Study, 2020-2021. [data collection]. 11th Edition. UK Data Service. SN: 8644, 10.5255/UKDA-SN-8644-11.

[12]Individual fixed effects are obtained by regressing real hourly wages on education categories, a quadratic in age, dummy variables indicating whether the individual was currently in permanent or temporary employment, he/she was married/cohabiting or single, in full or part time employment, his/her job was in the private or public sector, as well as the number of times the individual became non-employed and the number of time he/she made an occupation change since he/she entered the UKLHS, with additional controls for one-digit industries and one-digit occupations and year dummies.

Following the COVID-19 Study open call for content, we proposed a set of questions that aim at measuring individuals' job search strategies during the pandemic. These questions comprise the Job Search Module and were implemented in June and September 2020 and January 2021 (waves 3, 5 and 7, respectively).[13] We asked employed and non-employed individuals who said they were actively searching for jobs to name up to three types of jobs they were targeting, starting with their preferred one. We asked them to provide the exact job title and describe fully the sort of work they are looking for. This information was then coded (by professional coders) into the corresponding occupations using the Standard Occupation Classification (SOC) 2010. For each job we also asked individuals to report whether this is a job they are currently performing, have done in the past or have never performed. We also collected information on which industries they were searching for each of the three jobs. We provided the industry labels as described in the 1-digit Standard Industry Classification (SIC) 2007, which was available to respondents in a drop-down menu for each targeted job. Most respondents said the were only looking for one type of job, with 1,230 individuals only targeting one occupation among the 1,735 individuals who declared searching for a job; while 510 and 240 individuals declared searching for two and three different occupations, respectively. We use the responses on the preferred job to inform us about directions of job search, differentiating between targeted job transitions and realised ones. We focus on all male workers between 16 and 65 years of age and all-female workers between 16 and 60 years of age.[14]

The LFS is a quarterly household survey that provides the official employment and unemployment measures for the UK.[15] The cross-sectional LFS usually contains around 75,000 individuals and 36,000 households organized in 5 rotation groups or waves. Each wave denotes the quarters since the household first appeared in the survey and each household is followed for up to 5 quarters. This means that, at each quarter, one-fifth of the sample is replaced by a new group. The two-quarter longitudinal version of the LFS (2QLFS) comprises about 22,000 individuals. We use the 2QLFS to construct flows between states of economic activity, occupations and industries. This subset of the LFS focuses on the population of working age individuals. For this reason we restrict the cross-sectional and longitudinal samples

[13]The Job Search Module was also implemented in September 2021. However, we decided not to use this information as the evolution of the pandemic and the changes in the UK Government's policies renders this last wave less comparable to the other three.

[14]In the Job Search Module we also asked those individuals actively searching for a job about the use of search channels. Among those not searching for a job we asked about their reasons. Among employed individuals, the vast majority declared they were not actively searching as they were content with their current job/pay. Among the non-employed we found a significant proportion that were not searching due to health reasons or retirement.

[15]For details on how we define search activity, see Appendix D.1.

of LFS to the same age groups as used with the UKHLS. Although our main analysis focuses on the 2020-2021 period, we use the LFS from 2002Q2 in order to contrast the performance of the labour market during the Covid-19 pandemic to that of the Great Recession and its immediate aftermath.

During the LFS interviews, individuals are asked about their current employment status, if they are actively searching for a job and which search channels they use (e.g. job postings, networks, employment agencies, etc). The interviews also cover questions about the nature of their current job or their last job (if non-employed). Professional coders then use this information to classify occupations (and industries) into the existing SOC or SIC. During our period of analysis, there were two changes in the structure of SOC. In 2011, SOC 2000 was replaced by SOC 2010, and in 2020 the latter was replaced by SOC 2020. The Covid-19 pandemic also affected the response rates of the LFS. In order to address this issue new demographic characteristics were included in the survey's weighting procedure to further mitigate the impacts of sample representation.[16]

3.3 Aggregate Labour Market Shocks

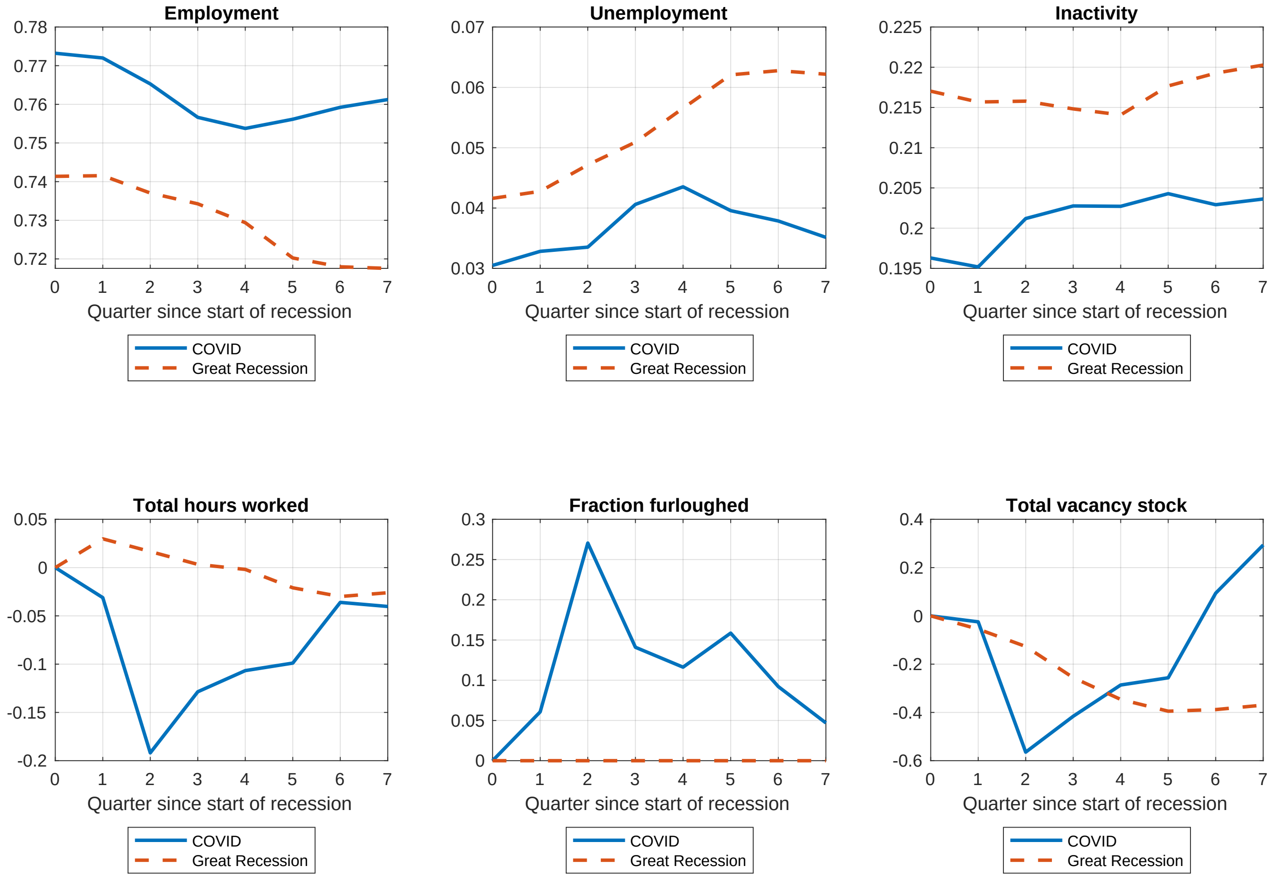
To set the context of the UK economy during the pandemic, we start by describing its impacts on labour market aggregates using LFS data. We use the Great Recession (GR) as a comparison to emphasise the unique features of the pandemic recession.

3.3.1 Stocks

Figure 11 depicts the behaviour of the stocks of employment, unemployment and inactivity as a proportion of the working age population, together with total hours worked, share of furloughed individuals and number of vacancies. These series are presented for the first seven quarters of the Great Recession (GR) and the Covid-19 pandemic in relation to their values observed during the quarter immediately preceding these events. The figure shows that the fall in employment during the pandemic has been similar to that observed in the GR, for the equivalent total number of quarters, despite a much larger GDP shock during the pandemic. This implies that while the Job Retention Scheme (JRS) implemented by the UK Government in April 2020 likely prevented a larger employment shock, it did not stop a very large fall in employment. The size of the GDP shock, combined with the

[16]The Office for National Statistics (ONS) provides information on the impact of Covid-19 on survey response and methodology changes in their Performance and Quality Monitoring Report, see Office for National Statistics (2021).

Figure 11 – Aggregate Labour Market Stocks during Covid-19 and the Great Recession



Note: Employment, unemployment, inactivity and hours worked series are computed from the LFS. The first three series are presented as a proportion of the working age population. The stock of vacancies is computed from the ONS vacancy survey. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic in relation to their values observed during the quarter immediately preceding these events. Start dates ($t = 1$) for the Great Recession and pandemic recession are 2008Q2 and 2020Q1 respectively. All series are seasonally adjusted with a stable seasonal filter.

presence of the JRS, is likely reflected in a fall in hours worked that has been much larger during the pandemic that in the GR; while at the same time we observe a rapid rise in the share of furloughed workers.

As has been documented elsewhere, the fall in employment during the pandemic was not accompanied by an equivalent rise in unemployment. This is in stark contrast with the experience during the GR. Instead the fall in employment initially manifested itself through a steep increase in the number of non-participants during 2020 Q2, as vacancies fell. The contribution of higher unemployment occurred during the second half of 2020, while the increase in inactivity slowed down and vacancies started to recover. The subsequent drop in unemployment then led to the recovery in the employment stock, albeit the increased number of non-participants tempered the employment recovery. This indicates that changes in the extensive margin of job search among non-employed individuals played an important part in shaping the aggregate labour market during the pandemic since the boundary between unemployment and inactivity is defined exactly by whether a non-employed worker is actively job searching or not.

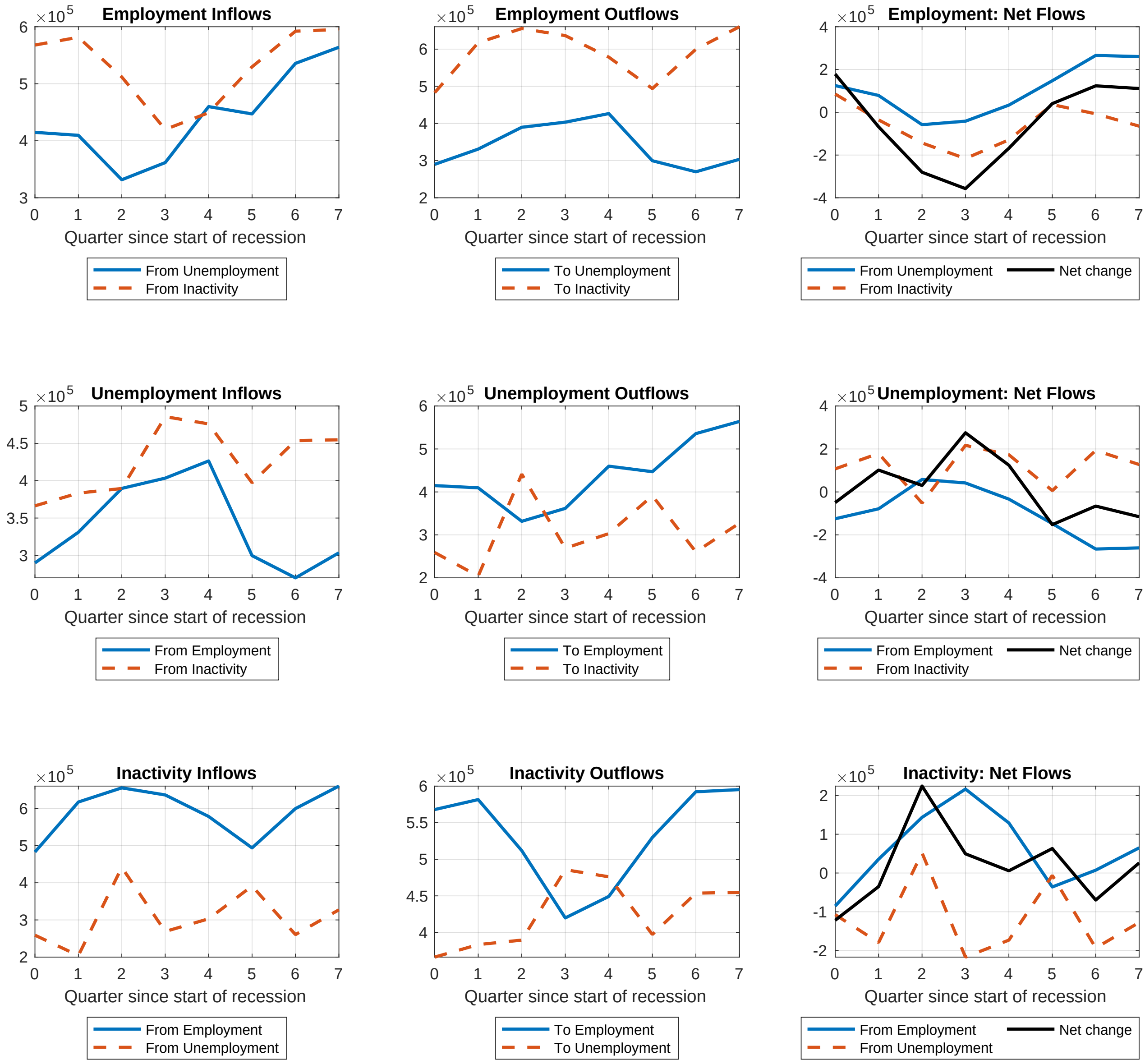
3.3.2 Worker Flows

To investigate the forces behind the changes in the stocks of employment, unemployment and inactivity, Figure 12 shows the absolute numbers of workers (in thousands) flowing between these different labour market states. In each row, the first graph depicts the total inflows to a given labour market state from the other two states. The second graph depicts the corresponding outflows and the third graph the net flows, which are defined as inflows minus outflows. Positive net flows therefore increase the stock of individuals in a given labour market state, while the negative net flows decrease this stock.

Taken together, these flows confirm and nuance the view of worker search activity suggested by the evolution of the stocks. Starting with employment, the top-centre panel of Figure 12 shows the flows from employment to inactivity and unemployment. During the initial two quarters of the crisis, outflows to inactivity increased by much more than outflows to unemployment. This implies that workers who lost their jobs during the early phase of the crisis mostly chose not to look for a new job, and were hence classified as inactive. The flow from employment to unemployment rises much more gradually, and during the second half of 2020 workers who transition out of employment were increasingly likely to transition into unemployment, and less likely to transition into inactivity. Hence, workers who lost their job later in the crisis were more likely to immediately search for a job, and hence be classified as unemployed. Combining these outflows with the inflows to employment gives the net flows to employment. Here we verify that early in the pandemic the increasing net outflow to inactivity is the main driver of the fall in employment, while later the increasing net outflow to unemployment played an important role. The outsized role of inactivity in this recession speaks to the importance of search dynamics.

Interesting dynamics are also at play between unemployment and inactivity, as can be inferred from the plots in the second and third rows. Early in the crisis there is a large inflow of workers from unemployment to inactivity, or in other words a large number of workers quitting active job search. This is shown by the spike in the dashed line in the bottom left panel at $t = 2$, which corresponds to flows between 2020 Q1 and 2020 Q2. Thus, the increased stock of inactive (i.e. non-searching) workers in 2020 Q2 corresponds both to recently unemployed workers

Figure 12 – Aggregate Labour Market Flows during Covid-19



Note: All flow series are computed from the two quarter LFS dataset. The left hand column shows the inflow into state X from state Y (where the state is employment, unemployment or inactivity) in period t, defined as the weighted number of employees in state X in quarter t who reported being in state Y in quarter $t-1$. Start dates ($t=1$) for the Great Recession and pandemic recession give the flow from 2008Q1 to 2008Q2 and 2019Q4 to 2020Q1, respectively. The middle column shows the outflow from state X to state Y in period t and is the weighted number of employees in state Y in quarter t who reported being in state X in quarter $t-1$. The right hand column shows net flows between state X to state Y, defined as the inflows to X from Y minus the outflows from X to Y. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic. All series are seasonally adjusted with a stable seasonal filter.

who choose not to search, *and to previously unemployed workers* who choose to stop searching and temporarily leave the labour force. Thus, the events of the first half of 2020 reduced worker search activity, even among those who had been previously searching. Importantly, this movement from unemployment to inactivity kept the unemployment rate lower in 2020 Q2, despite the non-trivial flows from employment to unemployment.

We then observe, within a single quarter, a reversal of this search decline, and flows away from inactivity and towards unemployment. In particular, in the bottom-centre panel we observe a jump in outflows from inactivity to unemployment starting in period 3, which corresponds to flows between 2020 Q2 and 2020 Q3. Combined with the increasing inflows from employment, this starts to finally raise the unemployment stock in 2020 Q3.

These flows paint a nuanced picture of worker search during the pandemic. Unemployment remained low early in the crisis both because workers who were fired early in the pandemic transitioned directly to inactivity, and many previously unemployed workers chose to temporarily stop searching and enter inactivity. Once this initial phase was over, and during the opening up of the economy and recovery of vacancies later in the year, workers began to transition to unemployment. Overall, worker search activity at the extensive margin appears very responsive to the state of the economy. Appendix D.2 further investigates how aggregate search activity evolved during the pandemic. The LFS allows us to do this directly as the survey asks both employed and non-employed workers whether they are actively searching for a job. In addition, those out of the labour force are asked whether they are willing to take up a job in the near future. These set of individuals are sometimes labelled as “marginally attached" workers who exert a low degree of search intensity relative, for example, to the unemployed. This analysis shows that aggregate search activity first decreased early on in the pandemic, rebounded during the second half of 2020, but then decreased to pre-pandemic levels by the end of 2021.

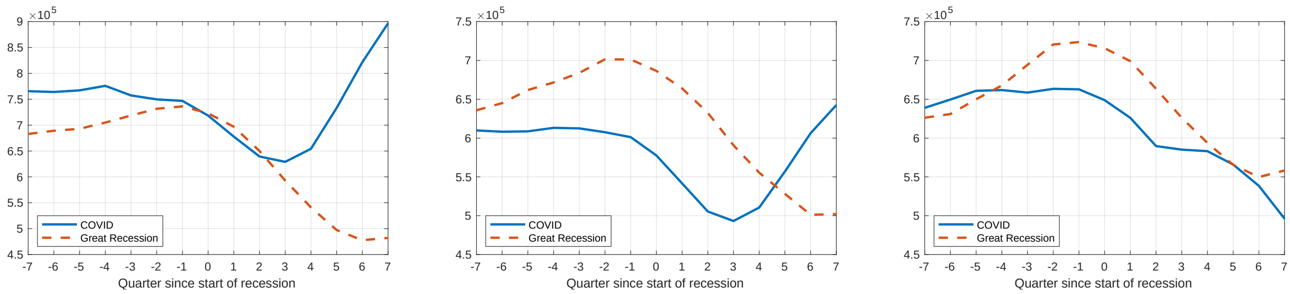
Of course, workers also flow between employers, industries and occupations as well between employment, unemployment and inactivity. Figure 13 shows quarterly job-to-job ($J2J$) flows, alongside gross flows between industries and occupations.[17] $J2J$ flows are the number of workers who are employed in two consecutive quarters and report a job tenure of less than three months with no spells of unemployment in the second quarter. Gross flows between one-digit industries (occupations) is the number of workers who change employer, either through a spell of non-employment

[17]All series are constructed from the two-quarter longitudinal LFS. We apply a five quarter moving average filter in order to smooth the mobility data. We apply the same smoothing to the J2J flow series for comparability.

or not, and reported an industry (occupation) in the new job that is different from the one reported in the last job held.[18] A large proportion of the gross occupation or industry mobility flows cancel each other and hence do not contribute to the changing size of occupations/industries. These "excess" mobility flows are typically interpreted as representing mobility due to workers' idiosyncratic career reasons, rather than mobility due to structural reallocation which we refer to as "net mobility" and discuss in the next section.

Figure 13 – Flows Between Employers, Occupations and Industries

a. Job-to-Job Mobility | b. Gross Mobility Flows: Industries | c. Gross Mobility Flows: Occupations



Note: All series are computed from the LFS. Industry and occupation classifications are based on the one-digit 2007 Standard Industrial Classification and one-digit 2010 Standard Occupational Classification, respectively. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic in relation to their values observed during the quarter immediately preceding these events. Start dates ($t = 1$) for the Great Recession and pandemic recession give the flow from 2008Q1 to 2008Q2 and 2019Q4 to 2020Q1, respectively. All series are seasonally adjusted with a five quarter moving average filter.

In the first panel of Figure 13 we observe that J2J flows fell during the initial phase of the pandemic, and then rapidly recovered as employment rebounded as we documented earlier with the other flows. Indeed, J2J flows (which imply reallocation of workers across firms) are now significantly above their pre-pandemic level. Gross reallocation across industries and occupations also fell during the beginning of the pandemic, meaning that workers transitioned across sectors less while the economy was weak. However, during the recovery, gross reallocation across industries and occupations behave differently: gross mobility across industries recovers, while gross mobility across occupations remains subdued. This suggests that larger numbers of workers might still be hesitant to change their occupation, despite the perceived strength of the labour market, while perceiving that changing industry presents less

[18]Several studies, notably Moscarini and Thomsson (2007) and Kambourov and Manovskii (2008), have emphasized measurement error in occupation and industry codes which create spurious mobility. Carrillo-Tudela and Visschers (2020), however, show that among employer movers correcting for coding errors when using a one-digit level of aggregation will decrease the observed gross occupational mobility rate by about 10 percentage points. In the case of industry mobility the decrease is of about 5 percentage points. This strongly suggests that the high levels of occupation and industry mobility we observe in the data will remain after correction.

of a risk. One possibility is that workers perceive that changing occupation requires forgoing occupation-specific human capital, while changing industry does not. We now turn to investigate patterns of occupational and industry reallocation in more detail, focusing on the occupations and industries job searchers targeted during the pandemic.

3.4 Labour Market Reallocation during Covid-19

In order to uncover the relative attractiveness of different sectors to individuals searching for jobs, we start by documenting the observed change in employment levels occupations/industries have experienced during 2020 and 2021 Q1, relative to their pre-recession trend. This provides a natural way to separate the declining occupations and industries from those that expanded during the pandemic. We define an industry or occupation as declining (expanding) according to whether the employment deviation for that industry/occupation in 2021 Q1 from the pre-recession trend was less (more) than the employment deviation for aggregate employment from its pre-recession trend. We compute the pre-recession trends from log-linear time trend based on 5 years of pre-recession data. We then examine whether job search behaviour reflects the observed patterns of employment changes by occupation and industry.

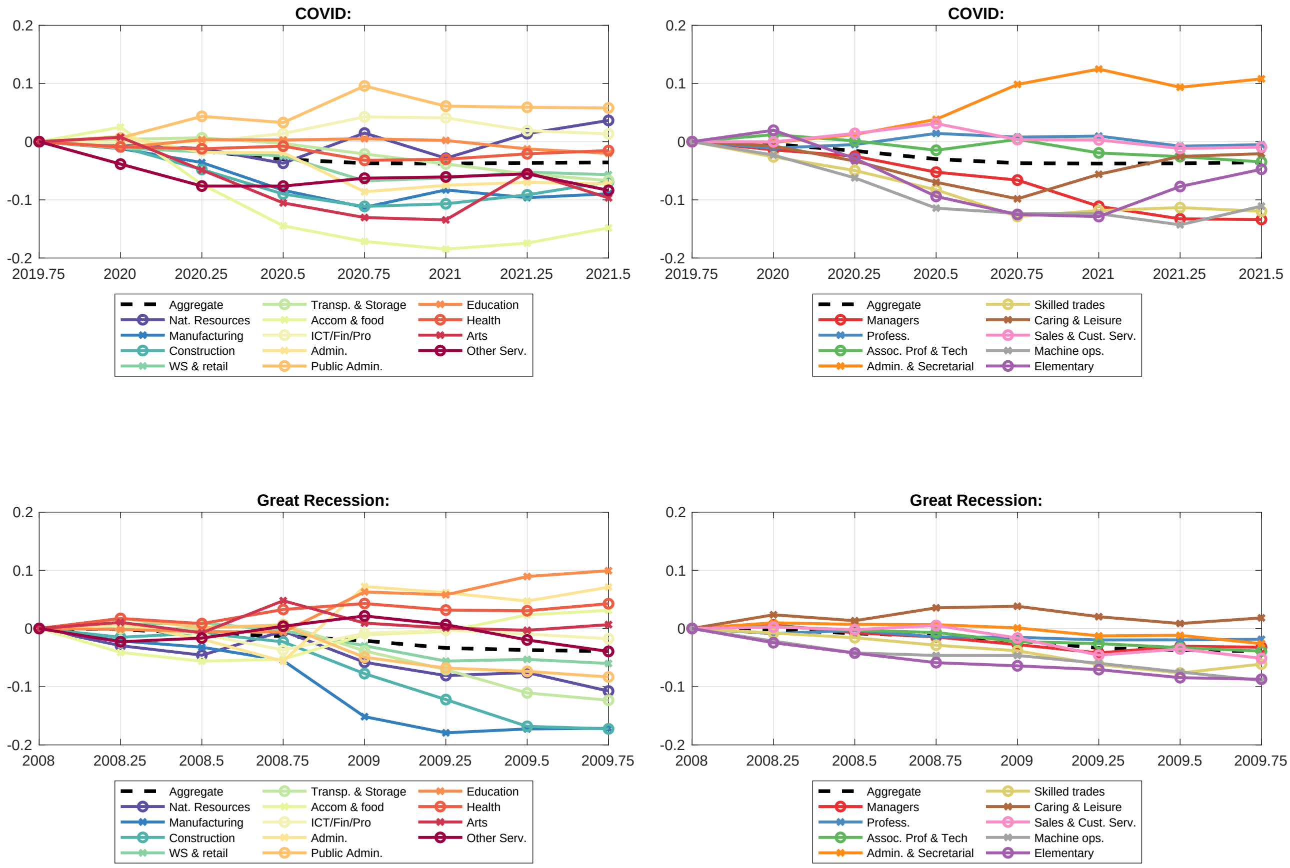
3.4.1 Changes in Employment by Industry and Occupation

The top row of Figure 14 depicts the change in employment relative to pre-pandemic levels experienced by one-digit industries and occupations (see Appendix D.1 for the complete classification labels). Given the lockdown measures applied in the UK, it is not surprising that the Accommodation and Food industry has been the worse performing industry, losing 20% of its employment by the first quarter of 2021. In contrast, Public Administration was the industry which experienced the largest increase, with about a 10% change in employment by 2021 Q1. In between these two we observe that the majority of the remaining industries lost employment, some of them by about 10%, while Education, Natural Resources and Technology/Financial/Professional Services related industries grew. A similar picture arises across occupations, with the majority of them declining and Elementary occupations (trade and services) being one of the worst affected, exhibiting about a 12% reduction by the end of 2020.

The bottom row of Figure 14 depicts the change in employment during the GR, relative to pre-recession levels for occupations and industries. The large heterogeneity in employment changes across occupations in the pandemic stands in contrast with that seen during the GR, where all occupations experienced smaller employment changes. This is evidenced by a much larger standard deviation of employment changes during the pandemic, 8.8%, relative to the one during the GR, 3.4%. Changes in employment among industrial sectors did display similar levels of heterogeneity across the two episodes. In this case, the standard deviation of employment changes during the pandemic and the GR are 7.5% and 7.0%, respectively. The nature of the GR, however, implies that the identity of the worst affected industries and occupations has been different.

Figure 14 – Employment during two recessions

a. Employment by Industry b. Employment by Ocupation



Note: All series are computed from the LFS. Industry and occupation classifications are based on the 2007 Standard Industrial Classification and 2010 Standard Occupational Classification respectively. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic in relation to their values observed during the quarter immediately preceding these events. Start dates for the Great Recession and pandemic recession are 2008Q2 and 2020Q1, respectively. All series are seasonally adjusted with a stable seasonal filter.

Figure 15 shows that the employment dynamics observed in the pandemic accelerated the longer term trends in the labour market by industry and occupation. Most of those occupations and industries that grew between 2002 Q1 to 2020 Q1, not only grew during the pandemic but experienced employment growth rates twice the size of their pre-pandemic growth rates. However, there are important exceptions.

For example, the Accommodation and Food industry was a long term growth sector with an average employment growth of 2%, but fared very badly during the pandemic.

Figure 16 shows that those industries and occupations that experienced employment losses in the pandemic also tend to be those that exhibit low average wages. It is therefore not surprising that workers with lower levels of educational attainment have seen outsize employment losses (not shown here), accompanied by large falls in labour force participation as documented in the previous section.

3.4.2 Nature of Shocks to Employment

One possible explanation behind the large changes in occupations' employment shares observed in the ongoing pandemic is that they are driven by underlying changes in employment shares by industry (or vice-versa). To investigate this possibility, we can decompose an occupation's percent change in employment, $\Delta e_o \equiv (e_{o,t} - e_{o,t-1})/e_{o,t-1}$, into a "between-industry" effect and a "within-industry" effect as shown below:

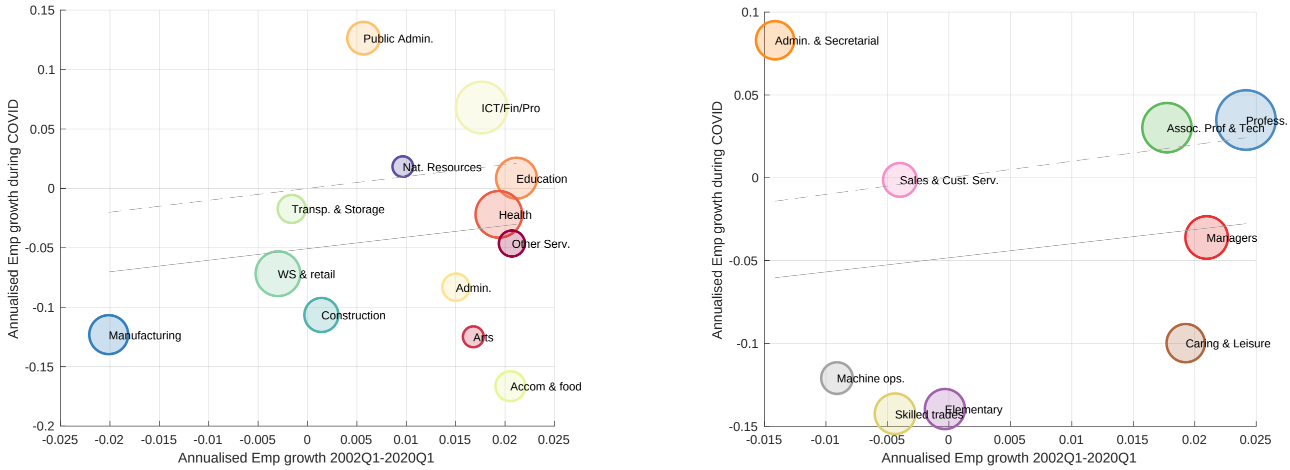
$$\Delta e_o = \sum_i \Delta e_{i,o} s_{i,o} = \underbrace{\sum_i \Delta e_i s_{i,o}}_{\text{industry effect}} + \underbrace{\sum_i (\Delta e_{i,o} - \Delta e_i) s_{i,o}}_{\text{occupation effect}} \tag{3.1}$$

where $s_{i,o} \equiv e_{i,o,t_0}/e_{o,t_0}$ is the employment share of industry i in total occupation o employment at time t_0, $\Delta e_i \equiv (e_{i,t_1} - e_{i,t_0})/e_{i,t_0}$ is industry employment growth between t_1 and t_0, and $\Delta e_{i,o} \equiv (e_{i,o,t_1} - e_{i,o,t_0})/e_{i,o,t_0}$ is joint industry-occupation employment growth. The first term in equation (3.1) calculates the predicted employment change if all industry-occupation bins in this occupation grew at the same rate as the overall industries. This is thus the industry effect. The second term captures the change in employment explained by occupation specific factors. That is, by industry-occupation pairs growing at a different rate from the industry averages.

The results are given in Table 9(a). The first column gives the employment fall during the pandemic for that occupation, up to the depth of the aggregate employment fall (2019Q4 to 2020Q4). The second column gives the industry effect from (3.1). For robustness, the third column gives the industry effect when the occupation's own employment is excluded from the industry employment changes.[19] The results clearly show that declining occupations have large occupation specific effects, since total employment fall for those occupations is much larger than the industry effects. This holds true for both measures of industry effects.

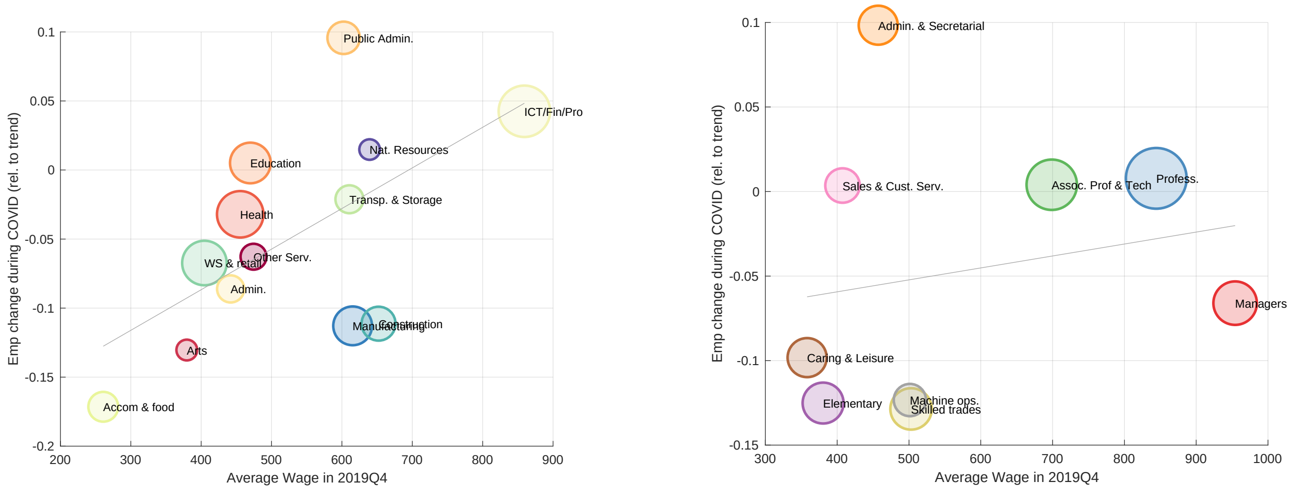
[19]That is, for each o we replace Δe_i in $\sum_i \Delta e_i s_{i,o}$ with $\Delta(e_i - e_{i,o})$. For industries where one occupation makes up a large share, this measure gives a more robust measure of the shock to the industry which excludes the shock to the occupation in question.

Figure 15 – Employment change from 2002 Q1 to 2020 Q1 vs. employment change Covid-19



Note: All series are computed from the LFS. Industry and occupation classifications are based on the 2007 Standard Industrial Classification and 2010 Standard Industrial Classification respectively. The size of the bubble indicates employment size in 2019 Q4. Employment growth during the Covid-19 pandemic is calculated from 2019Q4 to 2020Q4 using detrended employment.

Figure 16 – Employment change in Covid-19 vs. average wage



Note: All series are computed from the LFS. Industry and occupation classifications are based on the 2007 Standard Industrial Classification and 2010 Standard Occupational Classification respectively. The size of the bubble indicates employment size in 2019Q4. Employment growth during the Covid-19 pandemic is calculated from 2019Q4 to 2020Q4 using detrended employment.

Table 9 – Decomposing employment falls during Covid-19

Table 10 – Occupations

Occupation	Δe_o	Ind. effect	Ind. effect*
Admin & Secretarial	0.100	-0.016	-0.030
Professionals	0.009	-0.005	-0.010
Assoc Professionals	0.005	-0.013	-0.011
Sales & Cust Services	0.004	-0.055	-0.073
Managers	-0.068	-0.045	-0.041
Caring & Leisure	-0.099	-0.030	-0.010
Elementary	-0.123	-0.085	-0.073
Process Plant & Machine Op	-0.123	-0.060	-0.042
Skilled Trades	-0.129	-0.086	-0.071

Industries

Industry	Δe_i	Occ. effect	Occ. effect*
Public Admin	0.097	0.007	-0.002
ICT Finance & Profess	0.043	-0.001	-0.018
Natural Resources	0.015	-0.060	-0.063
Education	0.005	-0.021	-0.027
Transport & Storage	-0.020	-0.089	-0.099
Health	-0.033	-0.031	-0.029
Other Services	-0.062	-0.045	-0.043
Wholesale & Retail	-0.068	-0.033	-0.014
Admin & Support	-0.086	-0.060	-0.061
Construction	-0.112	-0.083	-0.075
Manufacturing	-0.113	-0.058	-0.054
Arts & Leisure	-0.131	-0.026	-0.019
Accom. & Food	-0.171	-0.092	-0.075

Note: All series are computed from the LFS. Industry and occupation classifications are based on the 2007 Standard Industrial Classification and 2010 Standard Occupational Classification respectively. In each table, the first column gives the detrended employment change of the industry or occupation from 2019Q4 to 2020Q4, and the second and third give the predicted employment fall given the joint industry-occupation makeup of the sector. See main text for definitions.

As an example consider Elementary occupations, which is the occupation group with the third largest decline. It is tempting to think its performance could be fully explained by the fall in employment in the Accommodation and Food industry. However, that industry only makes up 23% of the Elementary Services' employment. Hence the 17% fall in employment in the Accommodation and Food industry is not alone enough to explain why the Elementary occupation fell so much. Averaging across all industries still leaves a large proportion unexplained. Additionally, the best performing occupation, Administrative and Secretarial, is performing well for occupation specific reasons, since its industry effect is actually negative.

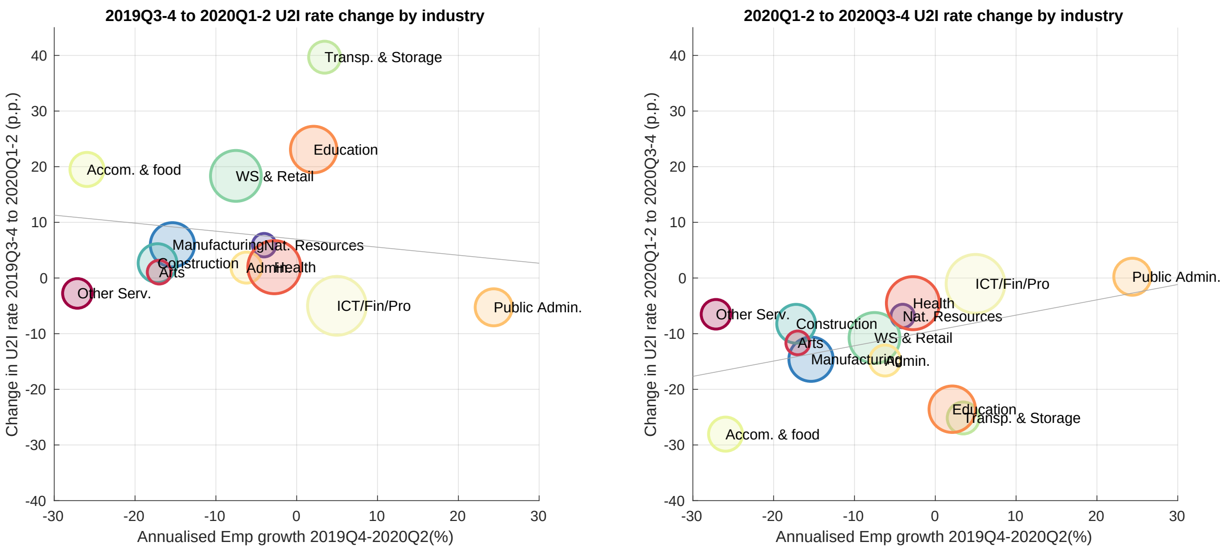
Table 9(b) shows the results from this same exercise, but now decomposing industry employment changes into equivalent components using $\Delta e_i = \sum_o \Delta e_o s_{o,i} + \sum_o (\Delta e_{o,i} - \Delta e_o) s_{o,i}$, with the first term giving the occupation effect. As with occupations, worst and best performing industries are hit by industry specific shocks.

3.4.3 Job Search at the Extensive Margin and Employment Shocks

The analysis of the previous section establishes that the pandemic has been characterised by significant variation in employment shocks by industry and occupation. A key question is whether and how this is reflected in workers' job search strategies. We now briefly consider this at the extensive margin – i.e. whether heterogeneity in employment shocks influence workers' decisions to search or not – before considering the intensive margin – i.e. the nature of jobs sought and how this varies by industry/occupation experience – in Section 3.4.4.

The left hand panel of Figure 17 plots the change in the rate of workers flowing from unemployment to inactivity – the change in the 'search-quit' rate of the unemployed – from 2019 Q3-Q4 to 2020 Q1-Q2 against annualised employment growth from 2019 Q3 to 2020 Q2. This is broken down according to the industry that unemployed workers previously worked in.[20] We observe larger increases in search-quit rates for the unemployed previously working in industries with larger falls in employment. For example, the Accommodation and Food industry had one of the largest falls in employment and a relatively large increase in search-quit rates. Conversely, the Public Administration sector saw increases in employment and a fall in search-quit rates. The right hand panel then shows how search-quit rates changed from 2020 Q1-2 to 2020 Q3-4, i.e. as the economy recovered in the second half of 2020. We see that unemployed workers who previously worked in initially harder hit industries saw decreases to their search quit rates on average. Both findings suggest that search activity at the extensive margin responds to heterogeneity in employment shocks in workers' previous industries.

Figure 17 – Changes in search quitting vs employment shocks



Note: All data comes from the LFS. The "U2I" rate shows the flow rate of individuals from unemployment to inactivity by industry previously worked using the SIC 2007 sector classification. The size of the bubble indicates employment size in 2019 Q3.

3.4.4 Jobs Sought by Occupation and Industry

In light of the importance of occupation and industry specific shocks, we now investigate whether individuals searching for jobs during the pandemic reacted by adjusting their search direction. A key innovation of the paper is that we collected

[20]Due to restricted data access we are unable to present the same figure according to the occupation that unemployed workers previously worked in.

information, through the Job Search Module of the UKHLS COVID-19 Study, on which occupations and industries job searchers were targeting during the second half of 2020 (June and September) and January 2021. As documented in Appendix B, focusing on this period is warranted by the observed rebound in the level of job search among employed and non-employed workers, which was accompanied by an increase in the proportion of individuals reporting that a major reason to engage in job search was to change occupation/sectors and subsequently by the recovery of gross occupation and industry mobility.

3.4.4.1 Distribution of targeted occupations and industries

Figure 18 documents the distribution of 1-digit occupations (left hand column) and industries (right hand column) associated with the first job choices declared by job searchers in the COVID-19 Study. We show this for all job searchers (top row), and then condition on whether the searcher is employed (middle row) or non-employed (bottom row). We further divide these targeted occupations and industries by whether they expanded or declined relative to aggregate employment during the pandemic as depicted in the top row of Figure 14. Crucially, we also show how the distribution of jobs targeted changes over time, which gives us a clear indication that workers adjusted their search patterns in response to the industry and occupation employment shocks experienced over the pandemic. As of June 2020, 55% of job searchers targeted occupations that were experiencing increases in their employment levels during the pandemic: this proportion increased to 71% by January 2021 (top left panel). The proportion of job searchers targeting expanding industries went from 38% to 46% over the same time period (top right panel). In that sense, the intensive margin of job search is responsive to occupation/industry-wide shocks. This qualitative pattern holds true for both employed and non-employed job searchers. However, in the bottom left hand panel we see that non-employed searchers are less responsive in the sense that the increased targeting of expanding occupations over time is less pronounced for non-employed searchers than the employed. In levels terms, we see a greater tendency of all searchers to target declining industries than occupations.[21] The tendency to target declining industries, in levels terms, is

[21]Note that the largest declining industry targeted by searchers is Other Services, which was only marginally declining during the pandemic and has expanded more than aggregate employment in the longer term (2002-2020). It is likely that respondents used this sector label inconsistently with its use by professional coders since around 23% of respondents say they are targeting the Other Services sector, but only around 3% are coded as having Other Services as their previous sector in 2019-2020. This discrepancy is larger than for any other sector. However, excluding those targeting the Other Services sector does not significantly change our key findings, with the exception of the probit analysis in Table 11. When we drop these respondents, non-employment still has a positive effect on the probability of targeting a declining industry but the effect is no longer statistically significant.

strongest among the non-employed (consistent with the probit analysis in Table 11) however both non-employed and employed workers change their search patterns over time in favour of expanding industries.

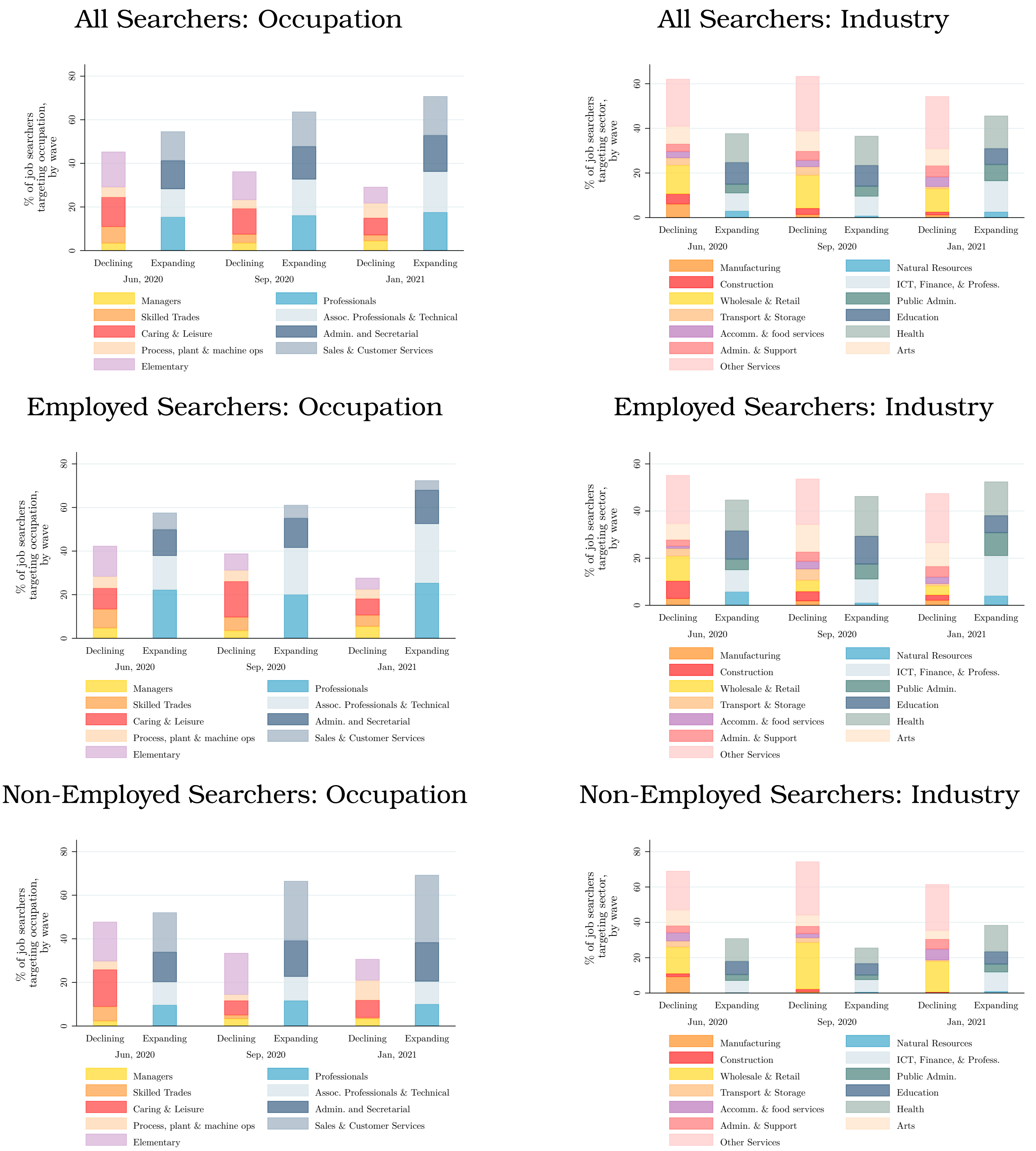
3.4.4.2 Probability of targeting declining occupations and industries

Figure 18 shows that, while workers adjust their job search in favour of expanding industries and occupations, a significant proportion target declining industries and occupations throughout the sample period. To investigate the latter we estimate the effects of demographic characteristics on the probability of targeting declining industries and occupations. Column 1 of Table 11 shows the marginal effects resulting from a Probit model where the dependent variable takes the value of one if the individual targeted a declining industry and zero if they target an expanding industry. Column 2 does the same for occupations. In both cases we control for whether the worker is female, young (16-34), has low education attainment (GCSE/other or less), is white, from London or not employed as well as for individual fixed effects computed from a Mincer wage equation as described in Section 2. We also control for whether the searcher comes from a declining industry or occupation, and investigate the correlation between targeting a declining industry and targeting a declining occupation (last two rows).

Workers with low education levels are significantly (at the 5% level) more likely to target a declining occupation but the impact of education on the probability of targeting a declining industry is insignificant. This suggests skill gaps may inhibit applications to growing occupations more so than growing industries. Row 6 confirms the analysis in Figure 18 showing that non-employed searchers are significantly more likely to target declining industries . Rows 7-8 clearly show the importance of attachment: coming from a declining industry (occupation) very significantly increases the probability of targeting a declining industry (occupation). Rows 8-9 investigate the extent to which targeting a declining industry is associated with targeting a declining occupation: there is no significant correlation. Overall, the key findings here are first that searchers' attachment to their previous jobs is a significant determinant of the probability of targeting a declining industry or occupation. Second, those with low education levels are more likely to target declining occupations , and this is not simply due to their increased likelihood from coming a declining occupation/sector, which is controlled for.[22]

[22]Note that we do not find that these effects are driven by composition of those looking to search, since the probability of job search is not significantly correlated with any of the covariates above with the exception of age, where we find the young are more likely to be job searching: see Appendix D.3

Figure 18 – Jobs sought over the pandemic



Note: Data from the COVID-19 Study, Job SearchModule of the UKHLS. Classification of industries or occupations as expanding or contracting is based on employment changes from the LFS, as detailed in the text.

Table 11 – Probability of targeting a declining industry or occupation

	(1) Declining Target Ind.	(2) Declining Target Occ.
Female	-0.103*	-0.0595
	(0.0535)	(0.0463)
Young (16-34)	0.0933	0.00962
	(0.0592)	(0.0539)
Low Educ. (GCSE or less)	0.0946	0.118**
	(0.0761)	(0.0573)
White	0.137	-0.0409
	(0.104)	(0.114)
London	0.135	-0.112
	(0.0829)	(0.0859)
Not Employed	0.120**	0.0506
	(0.0544)	(0.0517)
Declining Source Ind.	0.275***	0.114**
	(0.0484)	(0.0567)
Declining Source Occ.	0.0293	0.255***
	(0.0594)	(0.0438)
Declining Target Ind.		-0.0630
		(0.0620)
Declining Target Occ.	-0.0631	
	(0.0714)	
Individual Fixed Effects	-0.0168	0.0821*
	(0.0511)	(0.0472)
Observations	732	732

Standard errors in parentheses
Marginal effects ehown
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Data from the COVID-19 Study, Job Search Module of the UKHLS. Classification of sectors as expanding or contracting is based on employment changes from the LFS, as detailed in the text. Table shows the results (as marginal effects) of Probit estimations where the dependent variable takes the value of one if the individual targets a declining industry or occupation.

3.4.4.3 Probability of targeting an occupation change

The previous analysis shows the extent to which job searchers targeted expanding or declining occupations/industries. We now document the extent to which these individuals also target an occupational change. As mentioned earlier, we asked employed individuals whether they are searching for new employment in an occupation they are currently doing, have done in the past or have never done. Similarly, we asked non-employed individuals whether they are searching for new employment in an occupation they have done in the past or have never done before. The results are presented in Table 12. For each targeted occupation the first three columns show the proportion of employed individuals who currently have a job in such an occupation, have done a job in that occupation in the past or have never done a job in that occupation. The last two columns present the associate proportions for non-employed individuals.

Table 12 – Targeted occupational mobility (%)

	Employed			Non-employed	
	Current	Previous	Never	Previous	Never
Expanding					
Professional	38.83	21.95	39.22	57.96	42.04
Associate Profess. & Technical.	35.86	15.72	48.42	48.94	51.06
Admin. & Sec.	33.20	22.75	44.05	53.91	46.09
Sales & Customer Serv.	22.26	34.33	43.41	48.74	51.26
Declining					
Managers	43.33	19.56	37.11	90.85	9.15
Skilled Trade	62.72	18.84	18.44	100.00	0.00
Caring & Leisure	24.07	14.81	61.12	47.84	52.16
Process & Machine Op.	33.68	21.36	44.96	94.08	5.92
Elementary	42.33	16.27	41.39	65.79	34.21
Aggregate	39.05	17.31	43.63	68.99	31.01

Note: Data from the COVID-19 Study, Job Search Module of the UKHLS. Each row shows the characteristics of workers who report that they are looking for a job in that occupation. Of those workers, the first three columns show the proportion of employed individuals who currently have a job in such an occupation, have done a job in that occupation in the past or have never done a job in that occupation. The last two columns present the associate proportions for non-employed individuals. Classification of industries as expanding or contracting is based on employment changes from the LFS, as detailed in the text.

The table shows that employed individuals who actively engage in job search are largely looking for an occupational change. Across nearly all occupations we observe that less than half of these employed workers are searching for jobs in their current occupation. In contrast, non-employed individuals seem to prefer to go back to their previous occupations i.e. targeting occupations were they do have some experience.

In Appendix D.4 we show the results from a Probit regression where the dependent variable is an indicator variable taking the value of one if the respondent targets a new occupation i.e. one never previously performed, controlling for sex, race, age, education, whether they live in London, they come from a declining occupation or industry. We observe that non-employed respondents remain less likely to target a completely new occupation relative to employed workers. Young workers are significantly more likely to target a new occupation, consistent with occupation changing being more frequent in a worker's early career, as are female workers.

3.4.5 Targeted vs Realised Transitions

We now consider whether targeted job transitions are reflected in the realised job transitions observed in the data. We focus here on targeted and realised transitions that involve a change in occupation since these transitions are key to understanding worker reallocation over the pandemic.[23]

We focus on how frequently individuals target and realise transitions into expanding occupations. The blue bars of Figure 19 shows the fraction of individuals targeting switches into expanding occupations (taken from the COVID-19 study of the UKHLS) and the red bars show the fraction of individuals realising switches into expanding occupations (taken from the LFS). We additionally differentiate between individuals who come from expanding occupations — in the left hand of the figure — and declining occupations — in the right hand of the figure.

We observe that the proportion of workers actually making occupation switches into expanding occupations (blue bars in Figure 19) is substantially lower than the proportion of job seekers targeting a switch into an expanding occupation (red bars in Figure 19), particularly for those coming from declining occupations. The larger gap between desired and realised switches into expanding occupations for those starting in declining occupations suggests a pattern of segmentation, where it is harder to break into expanding occupations unless you start from one.

In Appendix D.5 we provide the full transition matrices underlying Figure 19. There we also compare the targeted transition matrices to the realised transition matrices computed for the period 2016-2019. We do this in order to investigate whether the gap between the targeted and realised transition matrices documented above arises because individuals were basing their search on past transition probabilities. This comparison suggests that there is some degree of past behaviour that could be driving a wedge between targeted and realised occupation transition matrices during the pandemic.

3.4.6 **Occupation and Industry Mobility Over Time**

The realised transition depicted in Figure 19 and detailed in the transitions matrices in Appendix D.5 provide a static measure of mobility over the course of 2020. A more dynamic measure comes from plotting the net flows from declining to expanding industries and occupations over time. This is done in Figure 20, with the right panel showing net flows from declining to expanding industries and the left

[23]We do not consider industry transitions here as targeted industries are not coded on a strictly consistent basis in the COVID-19 study of the UKHLS and in the UK LFS, where realised transition data is taken from.

panel showing net flows from declining to expanding occupations. Net flows (NF) are defined as

$$NF_{de,t} = I_{de,t} - O_{de,t},$$

where $I_{de,t}$ is the total inflow to expanding (e) from declining (d) occupations or industries, including through non-employment, and $O_{de,t}$ denote the total outflows from expanding to declining occupations or industries.

We observe higher levels of net flows from declining to expanding industries than from declining to expanding occupations, and a significantly steeper increase in industry net flows over the pandemic. This is again consistent with occupation mobility being more constrained, potentially by skill gaps or other demand side factors (i.e. experience requirements), than industry mobility.

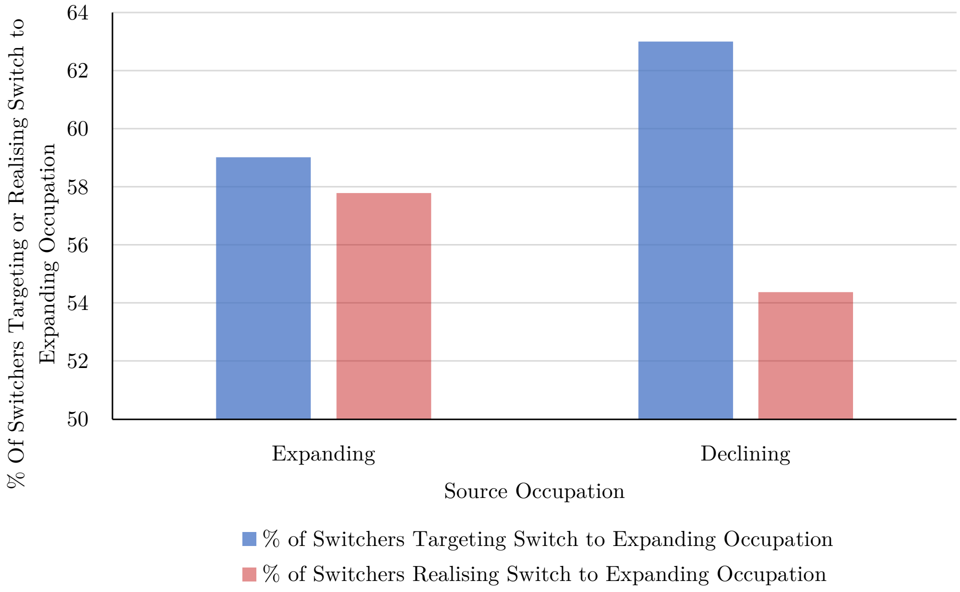
So far we have considered only mobility between two broad categories of industries and occupations: those that have declined and those that have expanded during the pandemic. It is also instructive to consider mobility across all industries and occupations. This broader measure of mobility captures the reallocation of individuals across occupations/industries such that their moves lead to the growth of some occupations/industries and the decline of others. One would expect this type of mobility to rise in the presence of large sectoral differences as individuals reallocate from poorly performing sectors to better performing ones. Given the evidence presented so far, one would expect net mobility to have increased during the Covid-19 pandemic. To investigate this conjecture we compute the aggregate net mobility flow using the following expression:

$$NM_{n,t} = \frac{1}{2} \sum_{n=1}^{N} |I_{n,t} - O_{n,t}| \omega_{n,t},$$

where $I_{n,t}$ and $O_{n,t}$ denote the total inflows and outflow to and from a given occupation or industry n at time t and $\omega_{n,t}$ denote the employment share of occupation or industry n at time t. The absolute values of the net flow to/from each sector are summed to make the total economy-wide flow. It is necessary to divide the summation by two in order to avoid double counting, as an inflow into one occupation/industry represents an outflow from another occupation/industry.

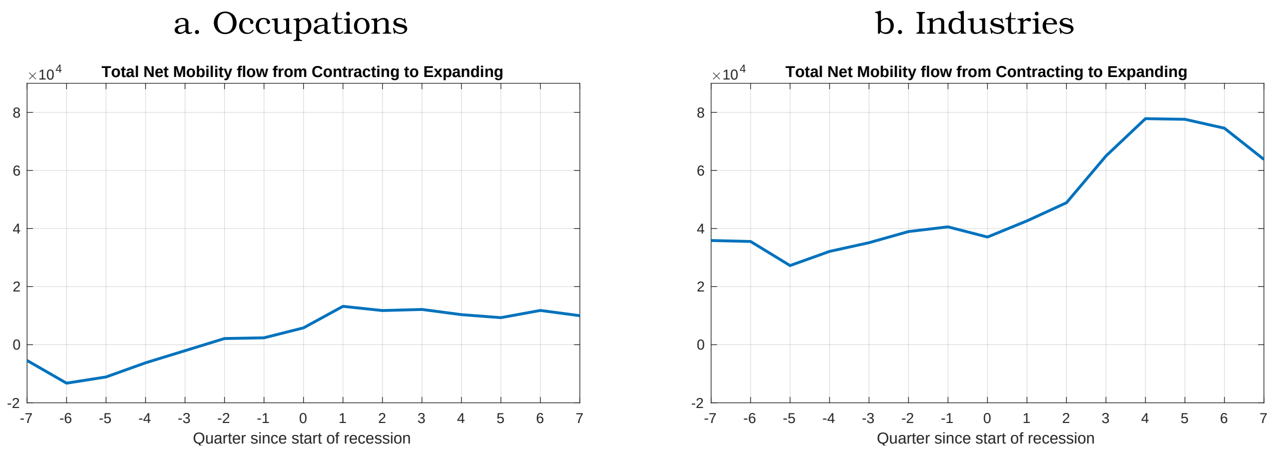
Figure 21 plots the aggregate net mobility flow, $NM_{n,t}$, across occupations and industries, comparing the pandemic with the pre-pandemic periods. While net mobility flows for occupations stay relatively flat, as per the Great Recession, there is a large increase in net mobility across industries. This increase is also much larger and persistent than the one observed during the Great Recession even though Figure 14 shows a similar dispersion in employment changes across industries during the two episodes. Thus, individuals appear to have reacted much more strongly to

Figure 19 – Targeted vs Realised Transitions



Note: Data on targeted transitions from the COVID-19 Study, Job Search Module of the UKHLS and data on realised transitions from the two-quarter LFS. Classification of occupations as expanding or contracting is based on employment changes from the LFS, as detailed in the text.

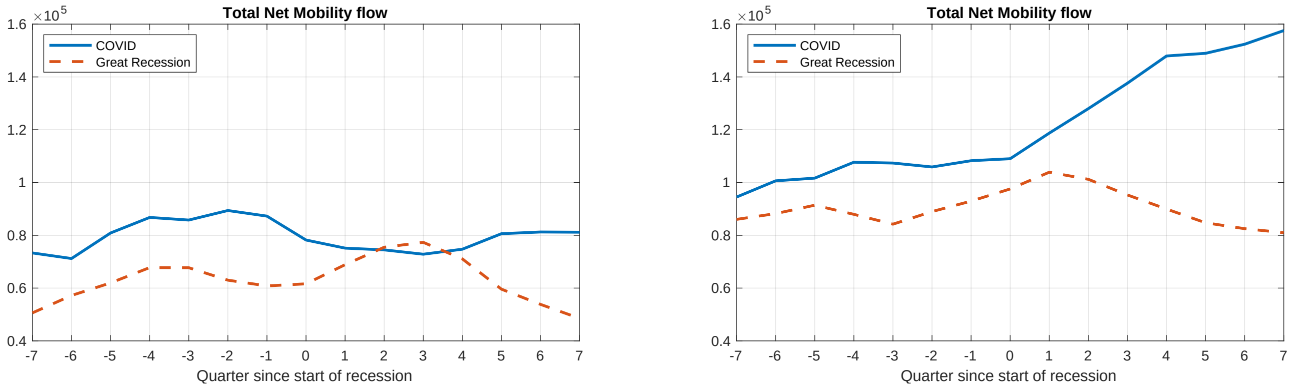
Figure 20 – Net Flows: Declining to Expanding Occupations and Industries



Note: All series are computed from the LFS. Net mobility flow from contracting to expanding sectors are defined in the text. The series are presented for the first seven quarters of the Covid-19 pandemic. Start date ($t = 1$) gives flows between 2019Q4 and 2020Q1. All series are seasonally adjusted with a five quarter moving average filter.

Figure 21 – Net mobility across occupations and industries

a. Occupations b. Industries



Note: All series are computed from the LFS. Net mobility flows are as defined in the text. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic. Start dates ($t = 1$) for the Great Recession and pandemic recession give flows from 2008Q1 to 2008Q2 and 2019Q4 to 2020Q1 respectively. All series are seasonally adjusted with a five quarter moving average filter.

industry differences during the pandemic that in the Great Recession. Note that the large discrepancy between the increase in industry net mobility and flat occupation net mobility occurs despite higher dispersion in employment changes for occupations than industry during the pandemic.

We have seen previously in Figure 13 that gross mobility across both occupations and industries fell more in the pandemic than in the Great Recession. The fact that gross mobility dropped even though net mobility stayed flat (occupations) or increased (industries) during the pandemic is a reflection that net mobility flows are much smaller than gross flows. That is, a large proportion of the occupation or industry mobility flows cancel each other and hence do not contribute to the changing size of occupations/industries. These "excess" mobility flows are typically interpreted as representing mobility due to workers' idiosyncratic reasons. The decrease in gross mobility then suggests that overall many individuals decided not to reallocate during the pandemic, perhaps waiting for the recovery to change careers and/or due to the effects of the JRS, which kept a significant part of the employment population in their jobs.

This highlights that for many individuals changing careers remains a difficult decision: do they wait for jobs to reappear in their previous industries/occupations, risking long periods of unemployment? Or do they accept available jobs, even if they lose their occupation/industry-specific skills which potentially means less job stability and lower earnings? The fall in gross reallocation suggests that the first motive has been more important for many individuals. Among those that reallocated, however, the rise in net mobility suggests that many did take into account industry (but less so occupation) differences when making mobility decisions. These patterns are consistent with Carrillo-Tudela and Visschers (2020) who link the fall in gross occupational mobility to the rise in unemployment using US data. Interestingly, in this paper we find that the recovery from the COVID recession is asymmetric between industries (which saw both gross and net mobility rise) and occupations (where gross mobility remains subdued, and net mobility remained flat).

3.5 Discussion

This paper has examined the importance of workers' search behaviour in driving labour market trends during the pandemic, as well as how search behaviour has reacted to labour market shocks. The relatively modest rise in unemployment during the pandemic has been accompanied by a more significant rise in inactivity. This suggests the margin between searching or not is important at an aggregate level. We also observe a tight link between changes to job search participation by employed and non-employed workers and changes to the vacancy stock, suggesting the extensive margin of job search responds to aggregate economic conditions. However despite increased outflows from inactivity to unemployment over 2021—that is increased numbers of previously inactive workers starting job-search—and a decline in workers flowing from unemployment to inactivity, labour market tightness (vacancies/unemployment) has still surged well above its pre-pandemic level as of 2021 Q3 due to the strength of vacancy creation and hiring as shown in Appendix D.6.

There has been considerable heterogeneity by sector as shown in Figure 37 in Appendix D.6. A key novelty of the paper is that it sheds light on the nature of the link between the direction of job search and labour market shocks at the level of industries and occupations. This is achieved by dis-aggregating search behaviour by workers' past and intended occupation and industry, using the COVID-19 Study of the UKHLS.

Of course the nature of the pandemic shocks on occupations and industries has been heavily mediated by policy interventions like the JRS scheme. Given the relatively robust job-to-job and unemployment-to-employment transition rates throughout the pandemic, and a strong rise in net mobility between industries, the balance of evidence suggests the JRS had a stronger impact in limiting job destruction than in holding back job creation or mobility. Indeed job-to-job mobility rates have recovered to a greater extent than the numbers of employees who report actively searching for a job. The fall in employees' job search is in contrast to the Great Recession and is consistent with the JRS limiting search effort. The fact that job-to-job mobility rates have recovered more robustly than the numbers of employees searching is, in turn, broadly consistent with Marinescu, Skandalis and Zhao (2021) who find that increases in unemployment benefits in the US decreased search effort but did not decrease job creation. These patterns also support the hypothesis that search congestion is likely to be particularly high during recessions meaning changes to search effort have a weaker impact on mobility rates, as predicted by job rationing models such as Michaillat (2012).

BIBLIOGRAPHY

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh (2020) "Furloughing*," *Fiscal Studies*, 41 (3), 591–622, https://doi.org/10.1111/1475-5890.12242.

——— (2022) "Perceived Returns to Job Search," *Labour Economics*, Forthcoming.

Albanesi, Stefania and Jiyeon Kim (2021) "Effects of the COVID-19 Recession on the US Labor Market: Occupation, Family, and Gender," *Journal of Economic Perspectives*, 35 (3), 3–24, 10.1257/jep.35.3.3.

Altonji, Joseph G. (1993) "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, 11 (1), 48–83, http://www.jstor.org/stable/2535184.

Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer (2016) "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success," *Journal of Labor Economics*, 34 (S1), S361–S401, 10.1086/682938.

Altuğ, Sumru and Robert A. Miller (1998) "The Effect of Work Experience on Female Wages and Labour Supply," *The Review of Economic Studies*, 65 (1), 45–85, http://www.jstor.org/stable/2567003.

Angrist, Joshua, Philip Oreopoulos, and Tyler Williams (2010) "When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards," Working Paper 16643, National Bureau of Economic Research, 10.3386/w16643.

Arcidiacono, Peter (2004) "Ability sorting and the returns to college major," *Journal of Econometrics*, 121 (1), 343–375, https://doi.org/10.1016/j.jeconom.2003.10.010, Higher education (Annals issue).

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom (2024) "College Attrition and the Dynamics of Information Revelation," *Journal of Political Economy*, 0 (ja), Forthcoming, 10.1086/732526.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo (2010) "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2 (4), 76–104, http://www.jstor.org/stable/25760233.

Arcidiacono, Peter and Robert A. Miller (2011) "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79 (6), 1823–1867, https://doi.org/10.3982/ECTA7743.

——— (2016) "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Mimeo*.

——— (2020) "Identifying dynamic discrete choice models off short panels," *Journal of Econometrics*, 215 (2), 473–485, https://EconPapers.repec.org/RePEc:eee:econom:v:215:y:2020:i:2:p:473-485.

Avery, Christopher and Caroline M Hoxby (2003) "Do and Should Financial Aid Packages Affect Students' College Choices?" Working Paper 9482, National Bureau of Economic Research, 10.3386/w9482.

Avery, Christopher and Thomas J. Kane (2004) "Student Perceptions of College Opportunities. The Boston COACH Program," in *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*, 355–394: National Bureau of Economic Research, Inc.

Avery, Christopher and Sarah Turner (2012) "Student Loans: Do College Students Borrow Too Much–Or Not Enough?" *Journal of Economic Perspectives*, 26 (1), 165–92, 10.1257/jep.26.1.165.

Baker, Rachel, Eric Bettinger, Brian Jacob, and Ioana Marinescu (2018) "The Effect of Labor Market Information on Community College Students' Major Choice," *Economics of Education Review*, 65, 18–30, https://doi.org/10.1016/j.econedurev.2018.05.005.

Balgová, Mária, Simon Trenkle, Christian Zimpelmann, and Nico Pestel (2022) "Job search during a pandemic recession: Survey evidence from the Netherlands," *Labour Economics*, 75, 102142, https://doi.org/10.1016/j.labeco.2022.102142.

Barahona, Nano, Caue Dobbin, and Sebastian Otero (2022) "The Equilibrium Effects of Subsidized Student Loans," working paper.

Bauer, Anja, Kristin Keveloh, Mariano Mamertino, and Enzo (2020) Weber (2020) "Competing for jobs: How COVID-19 changes search behaviour in the labour market," *IAB- Discussion Papers*, 33.

Baum, Sandy and S. Schwartz (2015) "Student aid, student behavior, and educational attainment," *Decision Making for Student Success: Behavioral Insights to Improve College Access and Persistence*, 38–60.

Becker, Gary S. (1965) "A Theory of the Allocation of Time," *The Economic Journal*, 75 (299), 493–517.

Belzil, Christian, Jorgen Hansen, and Xingfei Liu (2017) "Dynamic skill accumulation, comparative advantages, compulsory schooling, and earnings," *Quantitative Economics*, 8.

Belzil, Christian and Jörgen Hansen (2002) "Unobserved Ability and the Return to Schooling," *Econometrica*, 70 (5), 2075–2091, http://www.jstor.org/stable/3082032.

Benabou, Roland and Jean Tirole (2003) "Self-Knowledge and Self-Regulation: An Economic Approach," in *The Psychology of Economic Decisions: Rationality And Well-Being*: Oxford University Press, 10.1093/oso/9780199251063.003.0008.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen (2015) "Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education," *American Economic Journal: Economic Policy*, 7 (3), 86–125, 10.1257/pol.20130225.

Bertanha, Marinho (2020) "Regression discontinuity design with many thresholds," *Journal of Econometrics*, 218 (1), 216–241, https://doi.org/10.1016/j.jeconom.2019.09.010.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015) "GENDER IDENTITY AND RELATIVE INCOME WITHIN HOUSEHOLDS," *The Quarterly Journal of Economics*, 130 (2), 571–614, https://www.jstor.org/stable/26372609.

Betts, Julian R. (1996) "What Do Students Know about Wages? Evidence from a Survey of Undergraduates," *The Journal of Human Resources*, 31 (1), 27–56, http://www.jstor.org/stable/146042.

Blackwell, David (1953) "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24 (2), 265–272, http://www.jstor.org/stable/2236332.

Brasil (2017) "Manual de orientacoes RAIS,"Technical report.

Brooks, Benjamin, Alexander Frankel, and Emir Kamenica (2024) "Comparisons of Signals," *American Economic Review*, 114 (9), 2981–3006, 10.1257/aer.20230430.

Bulman, George, Robert Fairlie, Sarena Goodman, and Adam Isen (2021) "Parental Resources and College Attendance: Evidence from Lottery Wins," *American Economic Review*, 111 (4), 1201–40, 10.1257/aer.20171272.

Bárány, Zsófia L, Moshe Buchinsky, and Pauline Corblet (2023) "Late Bloomers: The Aggregate Implications of Getting Education Later in Life," Working Paper 31874, National Bureau of Economic Research, 10.3386/w31874.

Caetano, Gregorio, Miguel Palacios, and Harry Patrinos (2011) "Measuring Aversion to Debt: An Experiment Among Student Loan Candidates," *Journal of Family and Economic Issues*, 10.1007/s10834-018-9601-8.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2019) "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 23 (2), 192–210, 10.1093/ectj/utz022.

Cameron, Stephen V. and James J. Heckman (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106 (2), 262–333, http://www.jstor.org/stable/10.1086/250010.

Carneiro, Pedro, Italo López García, Kjell G. Salvanes, and Emma Tominey (2021) "Intergenerational Mobility and the Timing of Parental Income," *Journal of Political Economy*, 129 (3), 757–788, 10.1086/712443.

Carrillo-Tudela, Carlos, Bart Hobijn, Powen She, and Ludo Visschers (2016) "The extent and cyclicality of career changes: Evidence for the U.K," *European Economic Review*, 84 (C), 18–41, https://EconPapers.repec.org/RePEc:eee:eecrev:v:84:y:2016:i:c:p:18-41.

Carrillo-Tudela, Carlos and Ludo Visschers (2020) "Unemployment and Endogenous Reallocation over the Business Cycle," *CEPR*, DP14697, https://cepr.org/active/publications/discussion_papers/dp.php?dpno=14697.

Cattaneo, Matias D., Luke Keele, Rocio Titiunik, and Gonzalo Vazquez-Bare (2018) "Extrapolating Treatment Effects in Multi-Cutoff Regression Discontinuity Designs."

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011) " How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star *," *The Quarterly Journal of Economics*, 126 (4), 1593–1660, 10.1093/qje/qjr041.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan (2017) "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility," Working Papers 2017-059, Human Capital and Economic Opportunity Working Group.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan (2020) "Income Segregation and Intergenerational Mobility Across Colleges in the United States*," *The Quarterly Journal of Economics*, 135 (3), 1567–1633, 10.1093/qje/qjaa005.

Clark, Damon and Heather Royer (2013) "The Effect of Education on Adult Mortality and Health: Evidence from Britain," *The American Economic Review*, 103 (6), 2087–2120, http://www.jstor.org/stable/42920645.

Cohodes, Sarah R. and Joshua S. Goodman (2014) "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy," *American Economic Journal: Applied Economics*, 6 (4), 251–85, 10.1257/app.6.4.251.

Conti, Gabriella, James Heckman, and Sergio Urzua (2010) "The Education-Health Gradient," *American Economic Review*, 100 (2), 234–38, 10.1257/aer.100.2.234.

Crossley, Thomas F., Yifan Gong, Ralph Stinebrickner, and Todd Stinebrickner (2022) "The ex post accuracy of subjective beliefs: A new measure and decomposition," *Economics Letters*, 210, 110160, https://doi.org/10.1016/j.econlet.2021.110160.

Dale, Stacy B. and Alan B. Krueger (2014) "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data," *The Journal of Human Resources*, 49 (2), 323–358, http://www.jstor.org/stable/23799087.

De Groot, Morris (1970) *Optimal statistical decisions*: McGraw-Hill.

Deming, David J (2023) "Why Do Wages Grow Faster for Educated Workers?" Working Paper 31373, National Bureau of Economic Research, 10.3386/w31373.

Deming, David J. and Susan Dynarski (2010) *Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor*: University of Chicago Press.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977) "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1), 1–38, http://www.jstor.org/stable/2984875.

Dix-Carneiro, Rafael and Brian K. Kovak (2015) "Trade Liberalization and the Skill Premium: A Local Labor Markets Approach," *American Economic Review*, 105 (5), 551–57, 10.1257/aer.p20151052.

Dynarski, Susan (2005) "Building the Stock of College-Educated Labor," Working Paper 11604, National Bureau of Economic Research, 10.3386/w11604.

Eckstein, Zvi and Kenneth I. Wolpin (1999) "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67 (6), 1295–1339, http://www.jstor.org/stable/2999563.

Eng, Amanda and Jordan Matsudaira (2021) "Pell Grants and Student Success: Evidence from the Universe of Federal Aid Recipients," *Journal of Labor Economics*, 39 (S2), S413–S454, 10.1086/712556.

Epple, Dennis, Richard Romano, and Holger Sieg (2006) "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education," *Econometrica*, 74 (4), 885–928.

Faberman, Jason, Andreas Mueller, and Aysegul Sahin (2022) "Has the Willingness to Work Fallen during the Covid Pandemic?" *Mimeo*.

Faberman, Jason, Andreas Mueller, Aysegul Sahin, and Giorgio Topa (2020) "The Shadow Margins of Labor Market Slack," *Journal of Money, Credit, and Banking*, 52 (S2), 355–391.

Forsythe, Eliza, Lisa Kahn, Fabien Lange, and David Wiczer (2022) "Where have all the workers gone? Recalls, retirements, and reallocation in the COVID recovery," *Labour Economics*, forthcoming.

Francesconi, Marco (2002) "A Joint Dynamic Model of Fertility and Work of Married Women," *Journal of Labor Economics*, 20 (2), 336–380, 10.1086/338220.

Fryer, Jr., Roland G. and Steven D. Levitt (2013) "Testing for Racial Differences in the Mental Ability of Young Children," *American Economic Review*, 103 (2), 981–1005, 10.1257/aer.103.2.981.

Griliches, Zvi (1980) "Schooling Interruption, Work While in School and the Returns from Schooling," *The Scandinavian Journal of Economics*, 82 (2), 291–303, http://www.jstor.org/stable/3439860.

Hall, Bob and Marianna Kudlyak (2022) "The Unemployed with Jobs and without Jobs," *Labour Economics*, forthcoming.

Heckman, James and Pedro Carneiro (2003) "Human Capital Policy," Working Paper 9495, National Bureau of Economic Research, 10.3386/w9495.

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2008) "Earnings Functions and Rates of Return," *Journal of Human Capital*, 2 (1), 1–31, http://www.jstor.org/stable/10.1086/587037.

Hensvik, Lena, Thomas Le Barbanchon, and Roland Rathelot (2021) "Job search during the COVID-19 crisis," *Journal of Public Economics*, 194, 104349, https://doi.org/10.1016/j.jpubeco.2020.104349.

Herskovic, Bernard and Joao Ramos (2019) "Should College Admissions Favor Low-Income Applicants?", 10.2139/ssrn.2628303.

Hotz, V. Joseph and Robert A. Miller (1993) "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60 (3), 497–529, http://www.jstor.org/stable/2298122.

Hoxby, Caroline M. and Sarah Turner (2015) "What High-Achieving Low-Income Students Know about College," *American Economic Review*, 105 (5), 514–17, 10.1257/aer.p20151027.

Hoxby, Caroline and Sarah Turner (2015) "What High-Achieving Low-Income Students Know About College †," *American Economic Review*, 105, 514–517, 10.1257/aer.p20151027.

Institute For Government (2020) "The Coronavirus Job Retention Scheme: How has it been used and what will happen when it ends?" *IfG Insight*.

James, Jonathan (2011) "Ability Matching and Occupational Choice," *Working paper*.

Jamieson, Anne (2007) "Higher education study in later life: what is the point?" *Ageing and Society*, 27 (3), 363–384, 10.1017/S0144686X06005745.

Jones, Stephen R.G., Fabian Lange, W. Craig Riddell, and Casey Warman (2021) "Canadian Labour Market Dynamics During COVID-19," Working Paper 29098, National Bureau of Economic Research, 10.3386/w29098.

Kambourov, Gueorgui and Iourii Manovskii (2008) "Rising Occupational and Industry Mobility in the United States: 1968-97," *International Economic Review*, 49 (1), 41–79, http://www.jstor.org/stable/20486788.

Keane, Michael P. and Kenneth I. Wolpin (1997) "The Career Decisions of Young Men," *Journal of Political Economy*, 105 (3), 473–522, http://www.jstor.org/stable/10.1086/262080.

van der Klaauw, Wilbert (2002) "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43 (4), 1249–1287.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin (2023) "The Anatomy of Sorting—Evidence From Danish Data," *Econometrica*, 91 (6), 2409–2455, https://doi.org/10.3982/ECTA16425.

Liu, Kai, Kjell G. Salvanes, and Erik O. Sorensen (2016) "Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession," *European Economic Review*, 84, 3–17, https://doi.org/10.1016/j.euroecorev.2015.08.015, European Labor Market Issues.

Lochner, Lance and Enrico Moretti (2004) "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 94 (1), 155–189, http://www.jstor.org/stable/3592774.

Lucas, Robert E. (1988) "On the mechanics of economic development," *Journal of Monetary Economics*, 22 (1), 3 – 42, https://doi.org/10.1016/0304-3932(88)90168-7.

Lucas, Robert and Edward Prescott (1974) "Equilibrium Search and Unemployment," *Journal of Economic Theory*, 7, 188–209.

Malamud, Ofer (2011) "Discovering One's Talent: Learning from Academic Specialization," *ILR Review*, 64 (2), https://doi.org/10.1177/001979391106400209.

Manski, Charles F. (1989) "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education Review*, 8 (4), 305 – 312, https://doi.org/10.1016/0272-7757(89)90016-2.

Marinescu, Ioana Elena, Daphné Skandalis, and Daniel Zhao (2021) "The Impact of the Federal Pandemic Unemployment Compensation on Job Search and Vacancy Creation," SSRN Scholarly Paper ID 3801802, Social Science Research Network, Rochester, NY, 10.2139/ssrn.3801802.

Michaillat, Pascal (2012) "Do Matching Frictions Explain Unemployment? Not in Bad Times," *American Economic Review*, 102 (4), 1721–50, 10.1257/aer.102.4.1721.

Miller, Robert A. (1984) "Job Matching and Occupational Choice," *Journal of Political Economy*, 92 (6), 1086–1120, 10.1086/261276.

Modena, Francesca, Enrico Rettore, and Giulia Martina Tanzi (2020) "The Effect of Grants on University Dropout Rates: Evidence from the Italian Case," *Journal of Human Capital*, 14 (3), 343–370, 10.1086/710220.

Moretti, Enrico (2004) "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *The American Economic Review*, 94 (3), 656–690, http://www.jstor.org/stable/3592947.

Moscarini, Giuseppe and Kaj Thomsson (2007) "Occupational and Job Mobility in the US," *The Scandinavian Journal of Economics*, 109 (4), 807–836, http://www.jstor.org/stable/25195318.

Murphy, Kevin M. and Robert H. Topel (2016) "Human Capital Investment, Inequality, and Economic Growth," *Journal of Labor Economics*, 34 (S2), S99–S127, 10.1086/683779.

Office for National Statistics (2021) "Labour Force Survey performance and quality monitoring report: January to March 2021,"Technical report, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork.

Peterson, P.E., William Howell, P.J. Wolf, and D.E. Campbell (2003) "School vouchers: Results from randomized experiments," *The Economics of School Choice*, 107–144.

Pilossoph, Laura (2022) "Sectoral Shocks and Move Unemployment," *Mimeo*.

Pissarides, Cristopher (2001) *Equilibirum Unemployment*: MIT Press.

Pizzinelli, Carlo and Ippei Shibata (2022) "Has COVID-19 Induced Labor Market Mismatch? Evidence from the US and the UK," *Mimeo*.

Romer, Paul M. (1990) "Human capital and growth: Theory and evidence," *Carnegie-Rochester Conference Series on Public Policy*, 32, 251 – 286, https://doi.org/10.1016/0167-2231(90)90028-J.

Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999–1033, http://www.jstor.org/stable/1911259.

Sauer, Robert M. (2004) "Educational Financing and Lifetime Earnings," *The Review of Economic Studies*, 71 (4), 1189–1216, 10.1111/0034-6527.00319.

Stange, Kevin M. (2012) "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, 4 (1), 49–84, 10.1257/app.4.1.49.

Stinebrickner, Todd and Ralph Stinebrickner (2012) "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, 30 (4), 707–748, http://www.jstor.org/stable/10.1086/666525.

Sullivan, Paul (2010) "A dynamic analysis of educational attainment, occupational choices, and job search," *International Economic Review*, 51 (1), 289–317, https://doi.org/10.1111/j.1468-2354.2009.00580.x.

Vieira, Renato and Mary Arends-Kuenning (2019) "Affirmative action in Brazilian universities: Effects on the enrollment of targeted groups," *Economics of Education Review*, 73, 10.1016/j.econedurev.2019.101931.

Wiczer, David (2015) “Long-term Unemployment: Attached and Mismatched?” *Federal Reserve Bank of St. Louis Working Paper*, 2015-042A.

Appendix

APPENDIX A – DATA SUPPLEMENT FOR CHAPTER 1 AND 2

A.1 Data resources and access

Four different datasets were used to construct the final panel of individuals analysed in Chapter 1 and 2. These datasets, ACoHE, RAIS, ENEM and ENADE are administrative datasets that contain identifiers that allow for tracking individuals across datasets. For that reason, access to the data is controlled by the Brazilian Protected Data Access Service (SEDAP). To gain access to the identified microdata, the researcher must submit a project proposal through their website and, if selected, sign a data privacy agreement before they are able to access the data in a secure lab in Brazil[1].

A.2 Cohort Selection for Chapter 1

At any calendar year, I observe the stock of students in higher education and the stock of people in formal employment. These include people of all ages and education levels. As I focus on post-secondary education and work choices, I drop from the sample all individuals in formal employment who have not completed high school before 2010. I restrict the sample to individuals over 16 and below 65 years old. The ENEM and ACoHE datasets contain a few individuals without a person ID who are, therefore, non-identifiable. In consultation with the data provider, those observations usually account for individuals in prisons who are often unable to join the labour force or standard higher education status for a certain period of time. Due to their specific nature and the impossibility of matching observations in other datasets, I drop these observations from our sample. I also exclude from the sample individuals who have taken retirement at any point in time, from 2010 to 2018 and people who have enrolled in higher education for the first time after the age of 58.

In each of the original yearly cross-sectional datasets, a person's ID could appear more than one time if it was matched to more than one job or enrolled in more than one degree, for example. To deal with individuals observed more than once in RAIS I establish a series of filters for which employment register to keep.

[1] For more information on how to obtain access to the microdata, see: https://www.gov.br/pt-br/servicos/usar-o-servico-de-acesso-a-dados-protegidos-do-inep.

Table 13 – Number of individual observations in each dataset post cohort filtering

	RAIS	ENEM	Censo Sup	Enade
RAIS	91,502,590	7,319,023	6,062,213	1,154,854
ENEM		11,293,762	6,775,360	1,094,306
Censo Sup			13,553,813	1,717,936
Enade				2,664,223

The overruling register is the most recent employment register. In case there is more than one employment register dated December of the calendar year, then I keep the main one, the one with the most contract hours, then the highest pay, then the highest duration, in that order. In case of multiple-degree matches, I keep track of all enrolments to define graduates, stop-outs, dropouts and switchers. For details of how these variables are defined, see section A.4 in this Appendix. About 0.37% of individuals in our sample are enrolled in more than one program at the same point in time.

Next, I describe the algorithm used to select the cohort and match observations across datasets. As a last step, after the final panel is created, I further eliminate individuals with very unusual decision sequences in the data for purely modelling purposes. [2] which represent less than 0.1% of the sample for data privacy reasons. I later artificially add observations at very low counts to represent potential sequences of decisions not observed in the data. This is done purely for modelling purposes. Table 13 describes the number of observations of cohort members in each dataset after step 1 of the cohort creation algorithm, while Table 14 describes which dataset originates each one of the variables used to create the final dataset

A.2.1 The cohort creation algorithm

Due to the large number of personal records for every calendar year in each administrative dataset and intermittent missing variables, I use the following algorithm to create the cohort sample used in the study. The algorithm follows 3 steps:

1. I start by creating an intermediate dataset of all individual identifiers and all their time-invariant characteristics. I start from the dataset that contains the

[2] An unusual decision sequence is a sequence of decisions taken by $n \leq 3$ individuals across the cohort. There are $d^t = 1048576$ potential choice sequences, not all of the potential combinations are observed in the data.

Table 14 – Variable availability per dataset

	RAIS	ENEM	Censo Sup	Sisu	Enade
Sex	x	x	x		x
Ethnicity	x	x	x		x
Educaton Level(t)	x	x	x		
SE institution		x			
HE institution(t)			x	x	x
ENEM scores(t)		x		x	
Financial Aid(t)			x		
Degree Area(t)			x		x
Occupation(t)	x				
Industry(t)	x				
Wages(t)	x				
Employment status (t)	x		x		
Enrolment status (t)			x		
Employment duration (t)	x				
Region of residence(t)	x	x	x		
Parents' schooling		x			
Family income		x			
House owner		x			
Program applications				x	
Test score cutoff				x	
Rank in application				x	
Grades general					x
Grades specific					x
Grades total					x

largest collection of personal identifiers, RAIS, and then I loop over all datasets and all periods using the following steps:

- Collapse the dataset by person ID keeping the most-frequent or non-missing value in each variable generating an using dataset.
- Match these person ids to the ones from the using dataset from the previous iterations, updating the value of in case of missing variables
- save a cross-sectional dataset of all people and their time-invariant characteristics.

2. I use the dataset of IDs and time-invariant characteristics to select a list of personal identifiers of people who fulfil certain cohort characteristics. These are all individuals above 16 years old in 2009 who had not yet completed 50 years old in 2018.

3. With the list of personal IDs in hand, I search for each individual again in every dataset to generate panel variables. For each calendar year, I generate new variables that record the individual's current decision, state and ability signal in the final panel dataset.

A.3 Cohort Selection for Chapter 2

To construct the cohort in Chapter 2, I simply reduce the sample in Chapter 1 to everyone who takes the national entry exam, ENEM, in 2015. Note that individuals do not need to enrol in higher education in the following year, they need only to have taken the exam. In 2015, there were approximately 8 million candidates taking the exam. The following year saw about 8 million people enter higher education. Although numbers are close, these are not necessarily the same people, as some universities still hold their own entry exams for privately funded students.

A.4 Definition of variables

Education: The ACoHE dataset display students status using 6 categories: *enrolled, transferred, intermitted, graduated, evaded and deceased*. I use these to define the enrollment and non-enrolment decision. For this study, an individual is considered enrolled if it the student status variable (ST_ALUNO) at time t is equal to *enrolled*. If their status is *transferred*, they are recorded as a course switcher at time t or if they continued to be enrolled in $t+1$ but are now recorded in a different degree program. Similarly, a student is considered a graduate if at time t, if their status switches to *graduated* from $t-1$ to t. A stopout is a student whose status in t is *intermitted* or if they are not registered in a degree in the next period but are again registered as *enrolled* in the same program within the next 3 years. Finally, a student is considered a dropout if their registered status is equal to *evaded* in time t or if they are not observed as $enrolled$ in the same program at any point over the next 3 years[3].

Employment: The employment variables are sourced from RAIS. An individual is considered employed at time t if they possess at least one active *employer-employee* register in that year's RAIS. Note that the model's baseline category, non-employment, is simply a person ID from our cohort that does not possess an active job in that year. This procedure would code as non-employed people whose employment register was mistakenly not included in RAIS by the employer (estimated to be less than 3% of the formally employed population (Brasil, 2017)), or people who work informally[4].

Migration: If migrations are defined by municipality, 47.5% of the observations have changed municipalities by the time of the ENEM. If defined by the state level,

[3]Althoug there is attrition (people flowing in and out of the data) in the ACoHE panel data, only 0.34% are seem to be affected by it. Some of these are classified as stopouts or dropouts according to the program in which they are next observed in the Census

[4]It is true that Brazil is one of the few countries for which one can find data on informal labour. However, Brazilian informality statistics are derived from surveys with anonymous responses, precluding linkage to the microdata in this study

20.6% have migrated from their place of birth. By the time students enter higher education, 29.2% of the sample is observed as having changed states of residency since birth.

Graduation: Graduation is equal to 1 if an individual has obtained a degree within 6 years since time of enrolment and 0 otherwise.

Mature Student: A mature student is a student aged 21+ the year they enrolled for the first time in higher education.

A.5 Using ENADE scores as ability signals

Enade is a compulsory exam: failing to take ENADE, when drafted, hinders graduation. At the end of their university careers, the student must have received at least one ability signal from ENADE and potentially many. I assume that ENADE scores are a noisy signal of individual ability. One might worry about the use of ENADE scores as ability signals the assumptions on the timing of information revelation are wrong or if the student receives other signals about their ability (such as module grades or GPA)[5] while in Higher Education and is able to adjust their choices before they take ENADE.

[5] In certain aspects, ENADE may be a closer proxy of an individuals position in the ability distribution as it is an uniform exam applied nationally to measure cohorts and program quality.

APPENDIX B – LEARNING FROM WHAT

B.1 Description of the cohort

Table 15 presents a comparison of various characteristics of individuals in the final cohort. Summary statistics are broken down across four groups: non-enrollees, enrollees, graduates, and dropouts.The table shows that males are slightly more represented among dropouts that graduates and less likely to ever join higher education. As far as ethnicity goes, non-white individuals are more common among non-enrollees and dropouts. Graduates also tend to come from families with higher incomes and better-educated parents as expected from higher socioeconomic backgrounds than non-graduates. Also individuals who studied in public schools are both less likely to enrol and graduate from higher education than their private schooled peers. Graduates generally show stronger academic backgrounds, reflected in their better high school performance in ENEM scores across various subjects. Geographic differences are also evident, with people in education more concentrated in the Southeast and South[1]. Additionally, younger individuals and those who moved to different cities for higher education are more likely to complete their degrees. At the same time, financial programs like PROUNI and FIES, which target students in private institutions, seem to positively influence graduation rates.

Once enrolled, graduates and non-graduates seem to achieve similar grades on average. Average age at enrolment is slightly lower for non-graduates than for graduates. Average time to graduate is 3.3 years, aligned with the duration of most Bachelor degrees in the country. As expected, stopouts take about 1.2 years longer to graduate. Longer degrees and ever switching degree programs (Bachelors as oppose to License degrees " *Licenciaturas*" or technician degrees that last up to 2 years) are also more likely among those who stopout. The distribution of fields of study is rather homogeneous across graduates and non-graduates except for Engineering. Only 11% of graduates are engineers who make 13% of enrollees and 18% of stopouts.

Regarding individual's career history, descriptions show that around 40% of students have ever worked and studied during their careers. Students who have switched to education through inactivity are less representative among graduates (17%) than stop outs and dropouts (27%). Individuals who have transited from work to study make up about 42% of graduates versus 32% of dropouts. For those who

[1]The Brazilian Southeast and South are historically the richest and best developed regions in Brazil where manufacturing and service activities tend to be concentrated. For the same reason these are the regions with the highest ratio of Universities per inhabitant.

have worked within the duration of our panel, Table 15 reports their average earnings over time. Earnings of both graduates and individuals who never joined education seem to be in line (unconditional of age) with non-graduates and stopouts earnings being slightly less. Earnings growth however, is very homogeneous for all who ever participate in education (about 0.26% of real growth a year compared to 0.22% for those who never joined education. At the end of the period $t = 10$, the duration of employment in their last job ranges from 3.5 to 4.5 years - with the highest duration being attributed to those with a degree. Overall, the table highlights clear disparities in socioeconomic and academic backgrounds between graduates and those in other groups and the richness of the Brazilian administrative data.

I now look into descriptive statistics for two different types of students: young and mature students[2]. Mature students, with an average age of 30, have lower ENEM and ENADE scores across all subjects compared to Young students, who have an average age of 17.9 years at the time of enrolment and are more likely to be female (61% compared to 53%) . Younger students seem to come from higher socioeconomic backgrounds than mature students. As expected, most of mature student work (87%) as opposed to 54% of younger students. Mature students are also 9 p.p. more likely to have migrated and tend to be less concentrated in the Southeast.

At University, mature students are less likely to stop out (2% as compared to 12% for younger students and take less time to graduate (3.2 years versus 4 years for younger students). Mature students are also more likely to take up private loans than their younger counterparts. Regarding program preferences, mature students tend to enrol less in Engineering courses and more in Education programs relative to younger students.

Regarding career transition, only 14% of mature students transit from inactivity to studies while this is more common (60%) for younger students. The relationship is reversed when looking at transitions from work to study. As expected, younger student are more likely to be solely dedicated to studying (24% as opposed to 7% of mature students).

When considering average wage earnings over the period, Mature students earn higher amounts which may reflect higher experience and job duration.

B.2 Model derivation

In this section I detail the solution to the dynamic programming problem that leads to the estimating equation. Redefine equation (1.1) as the inclusive value of

[2]mature students are those aged 21+ at first-year enrolment.

Table 15 – Summary statistics of the cohort by education attainment

	(1) Non-enrolee		(2) Enrollee		(3) Graduate		(4) Stopout	
	mean	sd	mean	sd	mean	sd	mean	sd
Females	0.44	0.496	0.54	0.498	0.58	0.493	0.52	0.500
Not declared ethnicity	0.02	0.139	0.29	0.456	0.29	0.454	0.23	0.420
White	0.60	0.489	0.38	0.486	0.40	0.489	0.41	0.492
Non-White	0.38	0.485	0.32	0.468	0.31	0.464	0.36	0.481
People in hh	4.28	1.897	4.41	1.897	4.38	1.864	4.58	1.947
No child	0.67	0.471	0.86	0.342	0.85	0.357	0.86	0.350
Family Income*	2.96	3.173	3.28	3.252	3.23	3.164	3.36	3.409
Family Income per capita*	0.96	1.286	1.03	1.301	1.01	1.238	1.05	1.350
Family owns house	0.67	0.469	0.72	0.451	0.72	0.449	0.73	0.443
Works	0.91	0.293	0.79	0.404	0.82	0.382	0.75	0.395
Has repeated an year in high school	0.35	0.478	0.29	0.455	0.27	0.443	0.32	0.465
Has repeated an year in school	0.44	0.496	0.36	0.479	0.33	0.471	0.38	0.486
Looking for first job	0.05	0.222	0.11	0.314	0.09	0.289	0.11	0.307
Ever worked and studied	0.70	0.458	0.67	0.470	0.68	0.466	0.68	0.466
Public schooled	0.37	0.482	0.31	0.463	0.29	0.453	0.33	0.471
Father Not alfabetized	0.08	0.269	0.04	0.206	0.05	0.218	0.04	0.186
Father Primary School	0.66	0.472	0.56	0.497	0.58	0.494	0.55	0.498
Father Higher Education	0.20	0.402	0.28	0.449	0.26	0.441	0.30	0.456
Mother Not alfabetized	0.07	0.263	0.04	0.195	0.04	0.206	0.03	0.178
Mother Primary School	0.62	0.486	0.49	0.500	0.52	0.500	0.48	0.500
Mother Higher Education	0.23	0.423	0.31	0.463	0.30	0.458	0.33	0.470
Moved city	0.36	0.481	0.34	0.475	0.35	0.478	0.37	0.482
Moved city for HE	0.23	0.423	0.26	0.439	0.26	0.438	0.29	0.452
Region of residence:								
Central-west	0.08	0.266	0.06	0.238	0.06	0.241	0.06	0.229
Northeast	0.16	0.370	0.13	0.341	0.14	0.344	0.11	0.311
North	0.10	0.302	0.09	0.286	0.09	0.282	0.10	0.293
Southeast	0.56	0.496	0.64	0.481	0.63	0.483	0.65	0.478
South	0.09	0.291	0.08	0.270	0.08	0.278	0.09	0.291
ENEM score:								
Natural Sciences ENEM Score	488.96	72.243	502.39	80.598	501.58	78.819	496.20	75.983
Human Sciences ENEM Score	543.70	81.869	552.44	85.938	551.64	84.110	540.61	83.895
Languages ENEM Score	522.22	68.363	532.74	72.714	533.55	70.260	525.74	69.335
Maths ENEM Score	513.30	102.255	528.30	108.680	527.87	106.268	526.36	104.157
Writing ENEM Score	388.06	260.876	415.79	263.337	424.37	261.486	405.53	256.611
Total ENEM Score	516.34	67.304	528.22	74.436	527.98	72.144	521.39	70.354
ENADE score:								
General Knowlegde ENADE Score	-	-	49.08	17.775	49.08	17.456	49.74	17.562
Subject Specific ENADE Score	-	-	40.15	15.654	40.40	15.495	39.33	15.141
Total ENADE Score	-	-	42.39	14.155	42.58	13.924	41.94	13.734
Enrolment:								
Age at enrolment	-	-	29.21	7.889	29.67	7.933	28.38	7.762
Years to graduate	-	-	3.31	1.563	3.25	1.540	4.55	1.525
Public HE institution	-	-	0.15	0.353	0.16	0.340	0.15	0.357
Bachelor degree	-	-	0.59	0.493	0.55	0.497	0.68	0.465
Quota	-	-	0.03	0.173	0.03	0.163	0.03	0.179
Loan	-	-	0.44	0.496	0.48	0.500	0.49	0.500
FIES	-	-	0.34	0.474	0.36	0.479	0.44	0.496
PROUNI	-	-	0.18	0.384	0.21	0.406	0.17	0.376
Switched courses	-	-	0.17	0.371	0.14	0.345	0.23	0.419
Program of study:								
Education	-	-	0.17	0.375	0.19	0.390	0.14	0.350
Arts and Humanities	-	-	0.02	0.145	0.02	0.141	0.02	0.126
Social Sc., Business, Law	-	-	0.44	0.496	0.45	0.498	0.43	0.495
Maths and hard sciences	-	-	0.07	0.263	0.06	0.242	0.07	0.257
Engineering	-	-	0.13	0.338	0.11	0.313	0.18	0.384
Agriculture and Vet	-	-	0.01	0.103	0.01	0.096	0.01	0.102
Health	-	-	0.12	0.325	0.12	0.329	0.13	0.336
Transitions paths:								
I-S	-	-	0.36	0.479	0.17	0.378	0.27	0.444
W-S	-	-	0.67	0.470	0.42	0.359	0.32	0.475
Ever worked and studied	-	-	0.83	0.372	0.39	0.488	0.42	0.493
S-B	-	-	0.26	0.440	0.14	0.346	0.14	0.350
Work:								
RAIS earnings**	3303.67	4251.792	3141.32	2825.536	3231.52	2774.014	3031.65	2652.291
% Change in yearly earnings	0.22	0.325	0.26	0.381	0.26	0.330	0.26	0.336
Job duration (years)	3.78	2.117	3.48	2.086	3.67	2.099	3.42	2.087
Observations	33,542,881		3,275,772		1,469,443		60,764	

Note: *Family income is measured in minimum wages measured at $t = 0$. **Earnings refer to wages reported in RAIS averaged over time (in 2017 prices).

Table 16 – Summary statistics of the student cohort by age at enrolment

	(1) Mature Student		(2) Young Student	
	mean	sd	mean	sd
Females	0.53	0.499	0.61	0.488
Not declared ethnicity	0.30	0.460	0.28	0.450
White	0.37	0.483	0.47	0.499
Non-White	0.33	0.468	0.24	0.430
People in hh	4.30	1.874	4.74	1.835
No child	0.81	0.395	0.99	0.115
Family Income	3.19	3.227	3.38	3.031
Family Income per capita	1.02	1.284	1.01	1.340
Family owns house	0.69	0.464	0.81	0.391
Works	0.87	0.339	0.54	0.499
Has repeated an year in high school	0.29	0.454	0.29	0.454
Has repeated an year in school	0.36	0.480	0.33	0.468
Father Not alfabetized	0.05	0.224	0.01	0.119
Father Primary School	0.60	0.489	0.42	0.493
Father Higher Education	0.26	0.437	0.37	0.482
Mother Not alfabetized	0.05	0.212	0.01	0.101
Mother Primary School	0.54	0.498	0.34	0.473
Mother Higher Education	0.29	0.455	0.39	0.488
Moved city	0.36	0.480	0.27	0.445
Moved city for HE	0.24	0.429	0.28	0.451
Region of residence:				
Central-west	0.06	0.245	0.02	0.150
Northeast	0.14	0.350	0.04	0.204
North	0.09	0.292	0.02	0.125
Southeast	0.61	0.487	0.90	0.298
South	0.09	0.282	0.02	0.126
ENEM score:				
Natural Sciences ENEM Score	493.88	75.572	515.72	87.515
Human Sciences ENEM Score	547.85	83.409	547.99	91.056
Languages ENEM Score	526.03	70.111	543.89	75.992
Maths ENEM Score	518.28	104.974	550.37	110.035
Writing ENEM Score	395.29	261.884	488.44	254.554
Total ENEM Score	520.78	70.187	538.74	80.063
ENADE score:				
General Knowlegde ENADE Score	48.30	17.544	54.19	17.698
Subject Specific ENADE Score	39.76	15.503	41.97	15.823
Total ENADE Score	41.91	13.966	45.04	14.314
Enrolment:				
Age at enrolment	30.32	7.586	17.91	0.295
Stopped out	0.02	0.139	0.12	0.152
Time to graduate	3.22	1.539	4.03	1.359
Public HE institution	0.13	0.331	0.19	0.393
Bachelor degree	0.56	0.496	0.76	0.425
Quota	0.03	0.164	0.03	0.183
Loan	0.41	0.492	0.49	0.500
FIES	0.35	0.475	0.26	0.440
PROUNI	0.17	0.374	0.29	0.454
Switched courses	0.14	0.352	0.10	0.294
Program of study:				
Education	0.17	0.379	0.09	0.288
Arts and Humanities	0.02	0.138	0.03	0.166
Social Sc., Business, Law	0.45	0.498	0.45	0.498
Maths and hard sciences	0.07	0.257	0.10	0.295
Engineering	0.13	0.334	0.17	0.380
Agriculture and Vet	0.01	0.094	0.02	0.131
Health	0.12	0.319	0.12	0.320
Transitions:				
I-S	0.14	0.345	0.60	0.490
W-S	0.82	0.451	0.11	0.318
S-WS	0.10	0.305	0.34	0.474
Mainly studied	0.07	0.247	0.24	0.426
Work				
RAIS earnings	3133.02	2814.817	2872.60	2464.988
Change in yearly earnings	0.25	0.347	0.31	0.464
Choice tenure	3.51	3.151	2.26	2.179
Work experience RAIS	4.20	4.159	2.65	2.922
Job duration	3.59	2.102	2.67	1.747
Observations	1,985,416		1,174,918	

Note: *Family income is measured in minimum wages measured at $t = 0$. **Earnings refer to wages reported in RAIS averaged over time (in 2017 prices).

state Ω, in sector s at time t and let $T(.)$ represents a transition function on state variables:

$$D_t(\Omega, s) = \sum_{s'} [u_t(s, s', \Omega)/\rho + \beta E_t V_{t+1}(T(s, s', \Omega), s')/\rho] \tag{B.1}$$

Then, as shown in Rust (1987), and the GEV(1) assumption on idyossincratic shocks, the integrated value function once aggregate shocks are realized can be written as:

$$V_t(s, \Omega) = \gamma\rho - \rho \log\left(D_t(\Omega, s)\right) \tag{B.2}$$

Where ρ is the GEV variance scale parameter and γ is the Euler constant. Define $\tilde{V}$ as $V/\rho - \gamma/(1-\beta)$ and plugging in (B.1) gives the following Bellman equation:

$$\tilde{V}_t(\Omega, s) = \left(w(\Omega, s) + \delta(\omega, s) - \tau(\omega, s)\right)\rho + \beta \tilde{E} V_{t+1}\left(T(\Omega, s', \Omega), s'\right) \tag{B.3}$$

Equation (B.3) could be solved via Bellman Function Iteration for any path of wages and grades, the learning process and a process on forecasting aggregate shocks. Transition rates take the following logit form from Rust (1987) and the GEV assumption:

$$\log \pi_t(\Omega, s, s') = u_t(s, s', \Omega) + \beta E_t V_{t+1}\left(T(s, s', \Omega), s'\right) - \log D_t(\Omega, s) \tag{B.4}$$

(B.1) and (B.2) combined yields:

$$\tilde{V}_t(\Omega, s) = \gamma + u_t(, s', \Omega)/\rho + \beta E_t \tilde{V}_{t+1}\left(T(s, s', \Omega), s'\right) - \log \pi_t(\Omega, s, s') \tag{B.5}$$

Combining (B.5) and (B.4) one can iterate it for finite number of periods to arrive at the equation in the main text, equation (1.1). Iterating forward yields,

$$\tilde{V}_t(\Omega, s) = \frac{u_t(s, s', \Omega)}{\rho} + \beta E_t \left(\frac{u_{t+1}(s', s'', \Omega)}{\rho} + \beta E_{t+1} \tilde{V}_{t+2}\left(T(\Omega', s', s''), s''\right) - \log \pi_t(\Omega', s', s'') \right)$$
$$- \log \pi_t(\Omega, o, s')$$

Rearranging so that the transition probabilities are in the RHS:

$$\log \pi_t(\Omega, S, s') + \beta E_t \log \pi_{t+1}(\Omega, s', s'') = \frac{u_t(s, s', \Omega)}{\rho} + \beta E_t u_{t+1}(s', s'') + \beta^2 E_t V_{t+2}\left(T(\Omega', s', s''), s''\right)$$

$$-\tilde{V}_t(\Omega, s)$$

Finally, we can substitute $E_t \log \pi_{t+1}(\Omega, s', s'')$ with the realization of the transition rate and an expectational error ζ . For rational expectations, the forecast error will be uncorrelated with current variables. This leads to:

$$\log \pi_t(\Omega, s, s') + \beta \log \pi_{t+1}(\Omega, s', s'') = \frac{u_t(o, s', \Omega)}{\rho} + \beta E_t u_{t+1}(s', s'') + \beta^2 E_t V_{t+2}\left(T(s', s'', \Omega), s''\right)$$

$$-\tilde{V}_t(\Omega, s) + \zeta_{s', s'', t+1} \qquad \text{B.6}$$

From the equation above, I can apply the finite dependence results (Arcidiacono and Miller, 2016, 2020) to difference out the remaining expectation terms. As discussed in the main text, finite dependence uses the idea of a *renewal action*. Consider individual 1 and 2, that are identical in every way, except in accumulated human capital, and their most recent occupation. If both switch to a new occupation in the final period, s'' is such that $s'' \neq s_1$ and $s'' \neq s_2$ then we have:

$$T(\Omega_1, s_1, s'') = T(\Omega_2, s_2, s'')$$

Which means that[3]:

$$\beta^2 E_t V_{t+2}(T(\Omega_1, s_1, s''), s'') = \beta^2 E_t V_{t+2}(T(\Omega_2, s_2, s'), s'')$$

Next I subtract (B.5) for individual 2 from (B.5) evaluated for individual 1:

$$\log \frac{\pi_t(\Omega, s_1, s_2)}{\pi_t(\Omega, s_1, s_1)} + \beta \log \frac{\pi_{t+1}(\Omega, s_2, s'')}{\pi_{t+1}(\Omega_1, s_1, s'')} =$$

$$\frac{[u_t(s_1, s_2, \Omega) + \beta E_t u_{t+1}(s_2, s'', \Omega_2)] - [u_t(s_1, s_1, \Omega) + \beta E_t u_{t+1}(s_1, s', \Omega_1)]}{\rho}$$

$$+\zeta_{s_2, s'', t+1} - \zeta_{s_1, s'', t+1}$$

This leaves only utility parameters and forecast errors on the RHS of the equation. First taking care of the difference in u_t:

$$u_t(s_1, s_2, \Omega) - u_t(s_1, s_1, \Omega) = w(s_2) - w(s_1) + \delta(s_2) - \delta(s_1) - (\tau(s_2) - \tau(s_1) + \frac{1}{\rho}(E_t(A_2) - E_t(A_1))$$

And taking it to the next period:

[3]Check the main text for an illustration of the intuition behind the renewal action idea

$$E_t(u_{t+1}(s_1, s_2, \Omega) - u_{t+1}(s_1, s_1, \Omega)) = E_t\left(w(s'') + \delta(s'') - \tau(s'')E_\zeta(T(\Omega_2, s_2, s''), s'')\right)$$
$$-w(s'') - \delta(s'') + \tau(s'') - E_\zeta(T(\Omega_1, s_1, s''), s'')$$
$$= E_t(A_2(\Omega_2, s_2)) - E_t(A_1(\Omega_1, s_1))$$

Since switching is a renewal action, and the workers are otherwise identical, the only difference in the flow utility terms is the difference in switching costs for the workers. This is a constant, and so the expectation term can be dropped.

Combining these into the first equation one has the final estimating equation:

$$\log\frac{\pi_t(\Omega_1, s_1, s_2)}{\pi_t(\Omega_2, s_1, s_1)} + \beta\log\frac{\pi_{t+1}(\Omega, s_2, s'')}{\pi_{t+1}(\Omega_1, s_1, s'')} = \frac{w(\Omega_2, s_2, s'') - w(\Omega_1, s_1, s'')}{\rho}$$
$$+\frac{\delta(\Omega_2, s_2) - \tau(\Omega_2, s_2) - \delta(\Omega_1, s_1) + \tau(\Omega_1, s_1)}{\rho} + \frac{1}{\rho}\left(E_\zeta(A(\Omega, s_2, \zeta)) - E_\zeta(A(\Omega, s_1, \zeta))\right)$$
$$+\zeta_{s_2,s'',t+1} - \zeta_{s_1,s'',t+1} + m_{s_2,s_1,s'',t,t+1} \qquad \text{B.4}$$

where the new term, $m_{s_2,s_1,s'',t,t+1}$, reflects measurement error in the $\log\pi$ terms.

B.3 The Likelihood Equation

Because the idyossincratic shocks are assumed mutually and serially uncorrelated, a version of the model without unobserved heterogeneity can be estimated sequentially.[4] For simplicity, let's first considered a version of the model without any time-invariant unobserved heterogeneity (always known to the individual). The timing assumptions make it possible to integrate out the initially unknown heterogeneity parameters from the individual likelihood contribution. The likelihood contribution can then be written as an integral where $F(A)$ is the probability density function of the initially unobserved ability term. Let $G(.)$ and $W(.)$ represent the outcome functions of grades and wages previous defined.

$$L(si1, ..., siT, G_{i1}, ..., G_{iT}, W_{i1}, ..., W_{iT}) = \int L(si1, ..., siT, G_{i1}, ..., G_{iT}, W_{i1}, ..., W_{iT}|A)F(A)dA \qquad \text{(B.1)}$$

Because choices depend on ability A solely through the observed signals we can separate this expression between choices and outcomes using the law of successive conditioning:

[4] As explained in the main text, once the parameter from the outcome equations (wages and grades) and the CCPs are obtained in a first stage, the flow payoff parameters can be obtained by assuming the likelihood contributions of outcomes and choices to be separable.

$$L(si1, ..., siT, G_{i1}, ..., G_{iT}, W_{i1}, ..., W_{iT}) = L_{i,s} \times L_{i,GW} \tag{B.2}$$

The contribution of the sequence of decisions L_{id} is given by the product of the decision probabilities obtained in a discrete choice problem with Generalized extreme value distribution over time and, in a similar manner, the likelihood contribution from the sequence of outcomes - grades G and wages w - can be written as:

$$\mathrm{L}_{i,s} = L(si1) \times L(si2|si1, G_{i1}, W_{i1}) \times ... \times L(siT|siT-1, ..., si1, G_{iT-1}, ..., G_{i1}, W_{iT-1}, ..., W_{i1}) \tag{B.3}$$

In a similar manner, the likelihood contribution from the sequence of outcomes-over time is given by:

$$L_{i,G,W} = L(G_{i1}|si1A)...L(G_{iT}|si1, ...siT, A) \times L(_{i1}|si1A)...L(W_{iT}|si1, ...siT, A)F(A)d(A) \tag{B.4}$$

Where the normal conditional pdfs for each outcome is given by $L(Y_{it}|si, ..., sit, A)$. With the above likelihood structure in mind, it turns out that the the choice component and the outcome component of log-likelihood are additively separable. Having established these facts, I now turn to adding time-invariant unobserved heterogeneity to the model.

Likelihood with unobserved heterogeneity

Keeping the notation in the main text, suppose an individual's unobservable type is represented by m. I assume that m is drawn from a finite distribution, $\mathbb{Q}_m$ and define $M = |\mathbb{Q}|$. Let X_{it} represent the vector of observed covariates for individual i at time t and $d_it = s_it$, their current decision. The full likelihood becomes:

$$\begin{gathered}
L_i = \sum_{m=1}^{M} \pi(m) \times f(W_{iT}|s_{iT}, X_{iT}, m)\pi(s_{iT}|s_{iT-1}, X_{iT}, m) \\
\times \cdots \times f(W_{i1}|s_{i1}, X_{i1}, m)\pi((s_{i1}|s_{i0}, X_{i0}, m)\pi(s_{i0}|X_{i0}, m) \\
\times f(G_{iT}|s_{iT}, X_{iT}, m)\pi(s_{iT}|s_{iT-1}, X_{iT}, m) \\
\times \cdots \times f(G_{i1}|s_{i1}, X_{i1}, m)\pi((s_{i1}|s_{i0}, X_{i0}, m)\pi(s_{i0}|X_{i0}, m)
\end{gathered}$$

This is no longer additively separable, but the EM algorithm can be employed in order to estimate the model parameters (Arcidiacono and Miller, 2011). Before specifying things further I introduce some notation. Define: H_{it} to be the history of data for worker i up to time t; $q_{im} = P(m|H_{iT})$, the probability that worker i is of type m given their full history; Ξ to be the set of all parameters to be estimated. In this case, the objective function becomes:

$$
\begin{aligned}
J = \max \frac{1}{NT} \sum_{i=1}^{N} \sum_{m=1}^{m} q_{im} \times \Bigg(& \log \pi(s_{i0}|X_{i0}, m; \Xi) + \log \pi(X_{i0}|m; \Xi) \\
& + \sum_{t=1}^{T} [\log f(W_{it}|_{it}, s_{it}, m; \Xi) + \log \pi(W_{it}, s_{it}|H_{it-1}, m; \Xi)] \\
& + \sum_{t=1}^{T} [\log f(G_{it}|s_{it}, m; \Xi) + \log \pi(G_{it}, s_{it}|H_{it-1}, m; \Xi)] \Bigg)
\end{aligned}
$$

Where the first two terms in the summation account for the initial state of the worker and the sum term reflects the log likelihood of a particular history given a type. Using the EM algorithm one can maximize this likelihood by iterating over a guess of q_{im} and the parameters of the model using the steps described in Section 1.4.

To optimize the above equation without having to solve the full model, I approximate the likelihood by leveraging the observed transition matrix across states Ω. This enables me to estimate the π, the CCPs. Specifically, I maximize the following objective:

$$
\begin{aligned}
\hat{J} = \max \frac{1}{NT} \sum_{i=1}^{N} \sum_{m=1}^{M} q_{im} \times (& \log \hat{\pi}(s_{i0}|X_{i0}, m; \Xi) + \log \hat{\pi}(X_{i0}|m; \Xi) \\
& + \sum_{t=1}^{T} \log f(W_{it}, s_{it}, m; \Xi) + \log \hat{\pi}(W_{it}, s_{it}|H_{it-1}, m; \Xi) \\
& + \sum_{t=1}^{T} \log f(G_{it}|, s_{it}, m; \Xi) + \log \hat{\pi}(G_{it}, s_{it}|H_{it-1}, m; \Xi))
\end{aligned}
$$

In this case, $\hat{\pi}$ is estimated directly from the data. Given the values of q_{im}, solving for the parameters in the income equation becomes straightforward. The model follows a weighted least squares regression structure, meaning we need to stack the outcome and covariates vectors M times and include $M-1$ dummy variables for each type. The q_{iM} serves as regression weights in this setup. For any given guess of q_{iM}, solving for $\hat{\pi}$ is essentially equivalent to calculating a weighted average of transitions. This concludes the first stage likelihood set up.

Before moving to the regressions in the second stage, let's first explore how the left-hand and right-hand side variables are constructed. The left-hand side (LHS) in the estimation is expressed as:

$$\text{LHS} = \log \frac{\pi(s'|s, X)}{\pi(s|s, X)} + \beta \log \frac{\pi(s''|s', X')}{\pi(s|s, X')}$$

This equation indicates that the LHS represents the discounted log difference in observing a career path moving from $s \to s' \to s''$ and from $s \to s \to s''$, all while controlling for the same initial set of covariates, X. The notation X' reflects changes in covariates such as age and tenure, as these factors evolve with a worker's career trajectory. Note that this formulation assumes that the unobserved worker type is known, which is why the EM algorithm is applied in the first stage: it allows us to derive π as the solution to a weighted least squares regression. The weights correspond to the probability of an individual belonging to a specific type. Given a particular value of β, the LHS can then be constructed using parameters already estimated in the first stage.

The right-hand side variables represent the flow utilities tied to different career choices as specified in the main text, I won't go into detail about that here. As with the left-hand side, the agent's type is assumed to be known when considering these outcomes. The wages and grades for specific types and the type distribution are derived from the first stage. Consequently, in the second stage, the estimated flow utility for agents in a particular state is based on the grade and wage equation results from the first stage.

For the flow utility estimation happening in the second stage, I leverage the difference out equations in (B.5). In order to do that I:

1. I use all observed career transitions to create hypothetical one-step deviation paths. For instance, if I observe transitions from occupation A to B, from A to C, and from B to C, I then generate counterfactual probabilities for transitions like A to B to C and A to A to C.

2. After defining the possible paths in step 1, step 2 involves creating a grid across individual states by Selecting a subset of ages, starting at 28 and increasing in increments of 5 up to age 53 andepeating this process for all possible skill groups and worker types.

This results in a manageable size of data points in terms of memory usage.

APPENDIX C – THE ROLE OF FINANCIAL-AID FOR LOW-INCOME LOW-ACHIEVERS

C.1 Descriptive Statistics of the cohort

Table 17 displays summary statistics for those students who have enrolled but received any financial aid, those who enrolled with private loans, those who enrolled with subdized loans through FIES and those who enrolled with government-sponsored grants (PROUNI).

Program participation seems well-balanced in terms of gender and ethnicity and even parents' level of education. Average family income per capita (measured in minimum wages), is lower for those who benefit from government programs. This is expected since low income is a requirement for award eligibility. Those who have received an award are also more likely to have moved cities just before enrolment. Regarding National High School Exam performance, grant receivers seem to have the highest average test scores in all of the 5 test components. Then, people with private loans, then those with subsidized loans and finally, those with the lowest average scores, those who never took any financial aid. This flags out why controlling for different cutoffs may be necessary. Not only may the cutoffs highlight different abilities, but they may also highlight differences in willingness to enrol in higher education. After enrolment, the pattern repeats: grant receivers achieve higher ENADE scores than private loan takers, and very close to these are the subsidized loan receivers (even though their score distribution is more dispersed), and finally, those who did not receive any financial aid. The age profile of students at the time of enrolment is similar for financial aid takers and non-takers, with grant receivers being slightly younger on average. Another interesting result is that subsidized loan takers are more likely to pursue degrees associated with higher returns, such as Engineering and Health (medical degrees included).

As far as average differences in career paths go, financial aid takers and non-takers seem quite similar, with those who were awarded state-sponsored financial aid being slightly more likely to ever transition from solely studying to working and studying - a fact that may reflect their stricter financial constraints. At the end of the 10-year period, those who have received state-sponsored financial aid still receive lower wages than private loan receivers and non-and receivers even though they have similar work experience and job duration profiles.

Determinants of ENEM performance

In Table 18 I investigate the socioeconomic determinants of ENEM performance. Males tend to score 4.2% lower marks than females in the general score. In fact, men perform worse than women in all exam subjects, including maths. Private-schooled students tend to achieve scores 3.1% higher than public high school students, while those from rural areas perform worse than students in urban areas in all subjects, but especially in language-related topics[1]. Once we condition for all other characteristics, non- white students tend to perform better than white students in natural sciences, history and languages - but worse in maths and in the final essay. The level of education of the student's mother displays a strictly increasing relationship with ENEM performance in all subjects. Relative to being non-alphabetized, having a mother who has completed primary school increase ENEM scores by 2.9%. This difference becomes 5.9% for completing high school and reaches 10.8% for completing higher education. The magnitude of this impact is particularly higher for student's performance in the maths component. There are also regional differences in ENEM performance with the South region showing the highest relative performance and the North region the lowest.

C.2 Alternative specifications and Robustness Checks

When estimating a Regression Discontinuity Design (RDD), several robustness checks are essential to ensure the validity of the results. First, balance tests should be conducted to confirm that covariates are similar on either side of the cutoff, ensuring that the assignment is as good as random. Second, placebo tests involve estimating treatment effects at alternative cutoffs where no treatment should occur, checking for spurious results. Third, manipulation tests (like the McCrary test) assess whether individuals manipulate the running variable around the cutoff, which would invalidate the design. Additionally, varying the bandwidth size and using different functional forms (e.g., linear vs. quadratic) for the running variable should yield consistent estimates. Finally, implementing a falsification test by applying the RD approach to outcomes not expected to be affected by the treatment can also provide evidence of robustness.

For expositional purposes, the robustness checks in this appendix are illustrated employing the normalized multi-cutoff RD. This means that we normalize all

[1]ENEM components include a Portuguese language multiple choice test and an essay.

Table 17 – Summary statistics of the 2015 cohort by type of financial aid

	(1) No aid		(2) Private loan		(3) Subsisidized Loan		(4) Grant	
	mean	sd	mean	sd	mean	sd	mean	sd
Female	0.55	0.498	0.54	0.498	0.54	0.498	0.53	0.499
Not declared ethnicity	0.33	0.470	0.29	0.454	0.25	0.432	0.25	0.432
White	0.38	0.485	0.37	0.483	0.36	0.481	0.37	0.482
Non-White	0.29	0.455	0.34	0.474	0.39	0.487	0.39	0.487
Public schooled	0.37	0.482	0.30	0.457	0.25	0.435	0.19	0.388
Father Not alfabetized	0.04	0.203	0.04	0.202	0.05	0.208	0.04	0.193
Father Primary School	0.55	0.497	0.59	0.493	0.59	0.492	0.61	0.488
Father Higher Education	0.28	0.451	0.28	0.449	0.28	0.448	0.28	0.450
Mother Not alfabetized	0.04	0.194	0.04	0.193	0.04	0.197	0.04	0.186
Mother Primary School	0.49	0.500	0.52	0.500	0.53	0.499	0.54	0.498
Mother Higher Education	0.31	0.464	0.31	0.464	0.31	0.463	0.32	0.466
Family Income*	3.24	3.174	3.12	3.020	3.05	3.041	3.10	2.839
Family Income per capita*	1.03	1.273	0.97	1.111	0.92	1.054	0.96	0.993
Owns house	0.71	0.452	0.70	0.458	0.70	0.460	0.69	0.461
Has repeated an year in school	0.13	0.333	0.13	0.339	0.14	0.347	0.13	0.336
Moved city	0.33	0.471	0.34	0.475	0.36	0.479	0.34	0.474
Moved city for HE	0.24	0.426	0.24	0.430	0.25	0.434	0.24	0.430
ENEM score:								
Natural Sciences ENEM Score	485.20	73.065	494.96	71.045	480.91	68.671	516.75	65.274
Human Sciences ENEM Score	531.81	80.321	546.78	78.179	532.53	75.577	573.59	70.467
Languages ENEM Score	516.17	69.697	527.84	66.564	513.62	64.912	550.98	57.549
Maths ENEM Score	505.72	100.821	520.40	98.163	505.65	97.894	546.95	89.675
Writing ENEM Score	369.67	264.160	423.43	249.432	383.38	241.817	527.83	201.551
Total ENEM Score	509.13	66.628	521.87	65.031	507.62	62.618	546.68	57.406
ENADE score:								
General Knowlegde ENADE Score	47.58	17.455	49.42	17.381	48.19	17.003	55.85	16.221
Subject Specific ENADE Score	38.70	15.039	40.00	15.133	38.65	14.552	45.23	15.048
Total ENADE Score	40.93	13.587	42.37	13.597	41.04	12.932	47.90	13.159
Enrolment:								
Age at enrolment	28.95	7.880	28.44	7.604	28.28	7.161	27.10	7.429
Stopped out	0.02	0.152	0.03	0.162	0.03	0.180	0.02	0.156
Time to graduate	3.27	1.528	3.50	1.484	3.89	1.399	3.42	1.389
Bachelor degree	0.64	0.480	0.70	0.457	0.81	0.389	0.69	0.462
Switched courses	0.19	0.394	0.21	0.405	0.22	0.415	0.23	0.423
Program of Study:								
Education	0.14	0.350	0.12	0.328	0.09	0.284	0.12	0.330
Arts and Humanities	0.02	0.125	0.01	0.117	0.01	0.094	0.02	0.129
Social Sc., Business, Law	0.47	0.499	0.44	0.496	0.38	0.486	0.45	0.498
Maths and hard sciences	0.07	0.256	0.07	0.251	0.06	0.228	0.08	0.277
Engineering	0.13	0.332	0.16	0.370	0.23	0.423	0.15	0.355
Agriculture and Vet	0.01	0.088	0.01	0.103	0.02	0.124	0.01	0.101
Health	0.13	0.340	0.15	0.361	0.20	0.397	0.13	0.339
Region of residence:								
Central-west	0.06	0.246	0.06	0.242	0.07	0.250	0.05	0.213
Northeast	0.12	0.320	0.12	0.330	0.15	0.355	0.11	0.309
North	0.08	0.264	0.07	0.262	0.07	0.255	0.08	0.268
Southeast	0.68	0.467	0.68	0.467	0.67	0.470	0.69	0.461
South	0.06	0.245	0.06	0.238	0.05	0.209	0.07	0.260
Transitions:								
I-S	0.36	0.480	0.37	0.482	0.36	0.479	0.41	0.492
W-S	0.68	0.467	0.67	0.470	0.68	0.466	0.63	0.482
Ever worked and studied	0.85	0.360	0.85	0.358	0.85	0.359	0.86	0.350
Mainly studied	0.15	0.360	0.15	0.358	0.15	0.359	0.14	0.350
S-WS	0.27	0.443	0.31	0.462	0.35	0.478	0.35	0.478
Work								
RAIS earnings	3132.69	2808.920	2797.53	2389.292	2409.75	1855.160	2556.32	1714.613
Change in yearly earnings	0.26	0.363	0.26	0.354	0.25	0.342	0.27	0.350
Choice tenure	1.95	1.764	1.81	1.703	1.57	1.591	1.85	1.687
Work experience RAIS	3.96	4.049	3.66	3.930	3.13	3.728	3.82	3.840
Job duration	3.48	2.078	3.31	2.055	3.02	1.995	3.35	2.037
Observations	249,356		498,952		169,761		89,738	

Note: *Family income is measured in minimum wages measured at t = 0. **Earnings refer to wages reported in RAIS averaged over time (in 2017 prices).

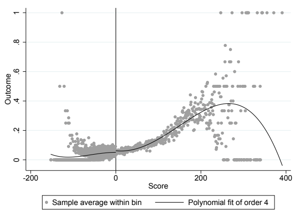
Table 18 – Estimates of ENEM performance by exam component

	(1) Nat. Sciences	(2) History	(3) Languages	(4) Maths	(5) Essay	(6) Total
Male	-0.0451***	-0.0413***	-0.00291***	-0.0806***	0.0266***	-0.0422***
	(0.000332)	(0.000349)	(0.000305)	(0.000436)	(0.000922)	(0.000288)
Private schooled	0.0342***	0.0294***	0.0270***	0.0346***	0.0514***	0.0312***
	(0.000372)	(0.000391)	(0.000342)	(0.000489)	(0.00103)	(0.000323)
Rural area	-0.0115***	-0.0159***	-0.0213***	-0.0157***	-0.0289***	-0.0160***
	(0.000659)	(0.000691)	(0.000604)	(0.000863)	(0.00183)	(0.000572)
Ethnicity:						
Brown	0.0210***	0.0149***	0.0178***	0.0310***	0.0340***	0.0210***
	(0.000450)	(0.000473)	(0.000414)	(0.000591)	(0.00125)	(0.000391)
Black	0.00123	0.0140***	0.00790***	-0.00235**	-0.0150***	0.00553***
	(0.000687)	(0.000721)	(0.000630)	(0.000900)	(0.00190)	(0.000596)
Asian	0.00528***	0.00893***	0.00571***	0.00876***	-0.0109***	0.00718***
	(0.000464)	(0.000487)	(0.000426)	(0.000609)	(0.00129)	(0.000403)
Native	0.0216***	0.0146***	0.0136***	0.0315***	-0.00110	0.0202***
	(0.00118)	(0.00124)	(0.00108)	(0.00155)	(0.00328)	(0.00103)
Other Ethnicity	0.00246	0.00821***	0.00347	-0.00446	-0.0222***	0.00236
	(0.00219)	(0.00230)	(0.00201)	(0.00287)	(0.00608)	(0.00190)
Mother's Education:						
Mother's educ = Low	0.0232***	0.0154***	0.0303***	0.0553***	0.00381	0.0299***
	(0.000867)	(0.000910)	(0.000797)	(0.00114)	(0.00242)	(0.000752)
Mother's educ = Med	0.0525***	0.0305***	0.0590***	0.105***	0.0266***	0.0598***
	(0.000892)	(0.000936)	(0.000820)	(0.00117)	(0.00249)	(0.000774)
Mother's educ = High	0.105***	0.0666***	0.0974***	0.171***	0.0864***	0.108***
	(0.000958)	(0.00101)	(0.000880)	(0.00126)	(0.00267)	(0.000831)
Region:						
Northeast	0.000932	0.0128***	0.00953***	-0.00138	0.0231***	0.00602***
	(0.000854)	(0.000896)	(0.000784)	(0.00112)	(0.00237)	(0.000741)
North	-0.0190***	-0.0122***	-0.0155***	-0.0288***	-0.00318	-0.0185***
	(0.000889)	(0.000933)	(0.000817)	(0.00117)	(0.00247)	(0.000772)
Southeast	0.0227***	0.00787***	0.0259***	0.0457***	0.0595***	0.0249***
	(0.000763)	(0.000801)	(0.000701)	(0.00100)	(0.00212)	(0.000663)
South	0.0315***	0.0335***	0.0311***	0.0456***	0.0563***	0.0351***
	(0.00102)	(0.00107)	(0.000932)	(0.00133)	(0.00282)	(0.000882)
Observations	769,728	769,749	754,060	754,024	742,026	771,719
R^2	0.096	0.049	0.069	0.133	0.020	0.116

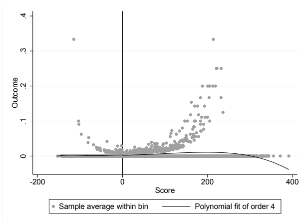
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Regression includes a constant. The baseline category for ethnicity is white, for mother's education is non-alphabetized, and for region is Central-west.

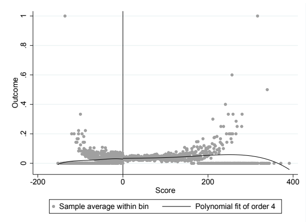
Figure 22 – Loan eligibility and ENEM scores



A.Enrolment



B.Graduation



C.Drop Out

Note: RD plots of selected outcomes on ENEM scores. Vertical line represents the policy cutoff. Each dot represent a mimicking variance quantile spaced bin (Cattaneo et al 2019).

cutoffs to zero and the running variable is treated as distance from cutoff. For each type of financial aid, our main robustness check specification follows a Fuzzy RDD:

$$Y_i = \beta + \alpha \mathbb{1}(X_i \geq c) + k(Si) + u_i \tag{C.1}$$

Running variable X_i is the distance from the institution/degree specific cutoff ($C = 0$).To get a sense for enrolment selection on post-enrolment outcomes, I estimate LATE from a subsample of retakers.

*nLoans

Figure 22 displays the relationship between loan availability and the normalized running variable across the entire support of the data[2]. The running variable (ENEM score) shows distinct behaviour at the cutoff compared to away from it. For enrolment (Panel A), there is an increase in the probability of enrolment just above the cutoff, indicating a significant effect of loan access, while further away from the cutoff, the relationship becomes less pronounced. In graduation (Panel B) and dropout (Panel C), the running variable shows minimal change at the cutoff, with little to no jump, suggesting that loans have a weaker impact on these outcomes compared to enrolment.

Actual regression results from the specified regression in equation (C.1) are

[2]As typical in RDD, for the actual regression, only observations within a short optimal bandwidth are kept for estimation of treatment effects.

Table 19 – Regression Results of a Normalize Multi RD for Loan eligibility

	Enrolment		Graduation		Drop Out	
	First Stage	Reduced Form	First Stage	Reduced Form	First Stage	Reduced Form
	(1)	(2)	(3)	(4)	(5)	(6)
. Unrestricted						
1(S>0)	.24362***	.36948***	.25458***	.24281***	.25484***	.0608**
	(.00991)	(.04238)	(.00879)	(.02682)	(.0084)	(.00663)
Bandwidth (h)	36.474	36.474	13.167	13.167	14.185	14.185
Observations	9488	9488	11734	11734	12626	12626
. If Retaker						
1(Si>0)	-	-	.09028***	1.6324***	-.10036***	-.32374**
	-	-	(.062075)	(.62075)	(.02566)	(.03024)
Bandwidth (h)	-	-	8.361	8.361	15.45	16.45
Observations	-	-	217	217	432	432

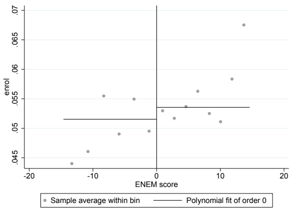
Note: Columns 1,3,5 show estimated first-stage model for the effect of passing 450-point threshold at any attempt (A) and second attempt(B). Remaining columns show estimated reduced-form models. Robust standard errors in parentheses. *: p-value< .1; **: p-value< .05; ***: p-value< .01

displayed in Table 19 and Figure 23. Being eligible for a loan increases enrolment and graduation, However these effects are much larger than the reported in the main text. Such an specification as in (C.1) tends to be biased in our case - pooling all cutoffs together assumes there are no differences in treatment effects across different values of the exam score distribution. When we look at results of the same regression for subsample of ENEM retakers, fo which selection at enrolment may play less of role in retention we see results more in line with the direction and size of coefficients from the main specification.

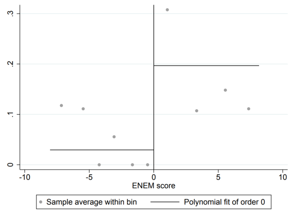
Next, I evaluate general evidence of sorting around the cutoff through the “Donut Hole Radius” test.

Based on Figure 24, the plot for enrolment shows that as the "Donut Hole Radius" increases (i.e., as more observations are excluded around the cutoff), the RD estimate remains relatively stable. The red dots, which represent the point estimates, consistently hover around a value close to 0.5. The vertical lines, representing the 95% confidence intervals, widen as the radius increases. This widening indicates increased uncertainty as more observations are excluded. For graduation, shown in panel B, the RD estimates are closer to zero across the different radii, with the point estimates declining slightly as more observations are excluded. The 95% confidence intervals also widen as the radius increases, suggesting a loss of precision as more

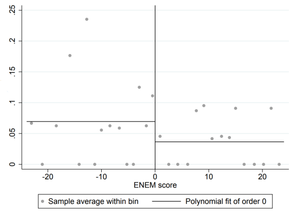
Figure 23 – Regression Results of a Normalize Multi RD for Loan eligibility



A.Enrolment



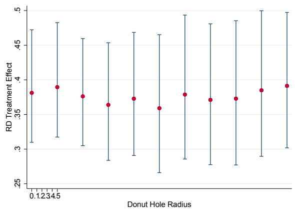
B.Graduation



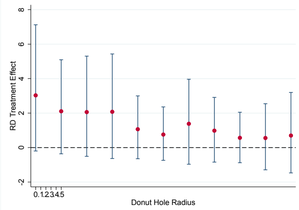
C.Drop Out

Note: RD plots of selected outcomes on ENEM scores. Vertical line represents the minimal policy cutoff. Each dot represent a mimicking variance quantile spaced bin (Cattaneo et al 2019).

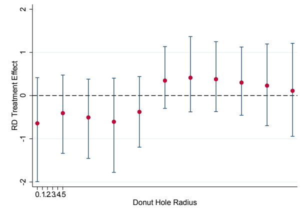
Figure 24 – Donut Hole radius: Loans



A.Enrolment



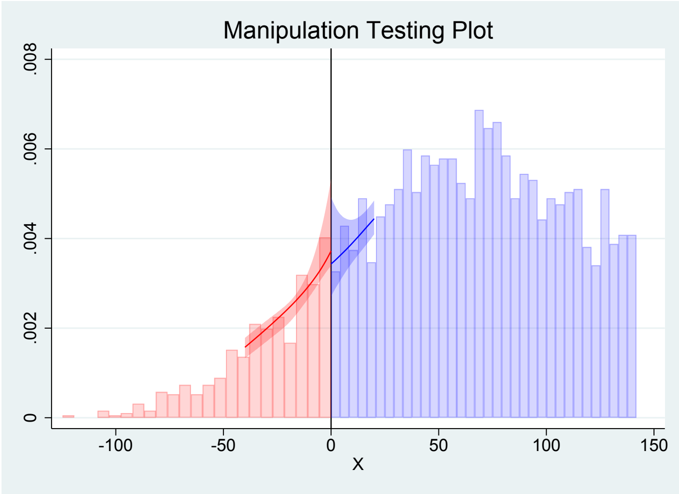
B.Graduation



C.Drop Out

Note: RD estimator sensitivity to excluding observation in an interval (x-axys) around the cutoff. Red dots represent the point estimates and vertical lines represent respective 95% confidence intervals.

Figure 25 – Manipulation Test: Loans

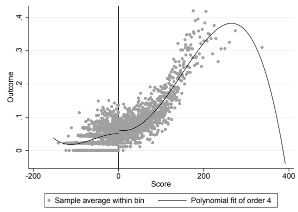


Note: Continuity in density test approach, exhibiting a histogram of the data and the actual density estimate with shaded 95% confidence intervals. The density estimates for treated and control groups at the cutoff 0 are close to each other, and the confidence intervals (shaded areas) overlap.Fail to reject null hypothesis that the density of the running variable is continuous at the cutoff.

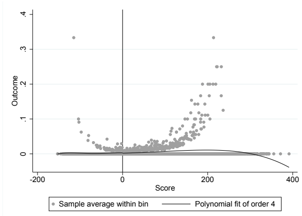
data is excluded. Notably, the estimates remain negative but are not statistically significant, as the confidence intervals consistently include zero. This implies that the RD estimator does not detect a strong effect of the treatment on graduation rates when observations close to the cutoff are excluded. The dropout panel C shows a similar pattern to graduation, with point estimates starting closer to zero and remaining negative across increasing donut hole radii. The estimates for dropout also show more variability and wider confidence intervals compared to enrolment, especially as the radius increases. Like graduation, these estimates are not statistically significant, as the confidence intervals consistently overlap with zero. This suggests that there is no strong evidence of a treatment effect on dropout rates when varying the exclusion of observations around the cutoff.

Figure 25 tests Density Continuity at the Cutoff. The density estimates on either side of this cutoff are fairly close to each other. This suggests that there is no significant discontinuity in the density of the running variable at the cutoff, as the densities before and after the cutoff are quite smooth and continuous. The shaded areas for both the red and blue density curves overlap at the cutoff, indicating that the density estimates for both the treatment (above the cutoff) and control (below the cutoff) groups are statistically similar. The test fails to reject the null hypothesis that the density of the running variable is continuous at the cutoff. This means there is no statistically significant evidence of manipulation around the cutoff, which strengthens the credibility of the RD analysis.

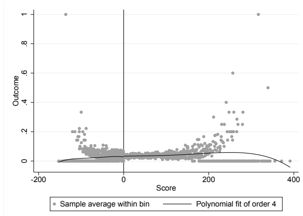
Figure 26 – Grant eligibility and ENEM scores



A.Enrolment



B.Graduation



C.Drop Out

Note: RD plots of selected outcomes on ENEM scores. Vertical line represents the minimal policy cutoff. Each dot represent a mimicking variance quantile spaced bin (Cataneo et al 2019).

Grants

Figure 26 presents the relationship between of grant outcomes for enrolment, graduation, and dropout over the support of ENEM scores. In Panel A, enrolment shows a sharp increase immediately after the cutoff, indicating a significant positive effect of the policy on enrolment rates. The polynomial fit reveals a steep rise, suggesting that students eligible for the grant are much more likely to enroll. Panel B, however, shows no significant jump in graduation rates at the cutoff, with a flat trend indicating that the policy does not appear to impact graduation outcomes. In Panel C, dropout rates slightly decrease after the cutoff, implying a modest reduction in dropouts for students receiving the grant, although the effect is weaker and less certain than enrolment.

The reduced-form results from the regression Table 20 and Figure 27 show a significant negative impact of passing the cutoff on enrolment and a small, non-significant positive effect on dropout , suggesting that loan access slightly reduces enrolment without strongly influencing dropout. For graduation, passing the cutoff significantly increases the likelihood of graduating, indicating a positive effect on this outcome. However, for retakers, the results are more mixed, with a strong positive effect on dropout but a smaller, albeit significant, positive effect on enrolment, suggesting differential impacts depending on the student's status as a retaker.

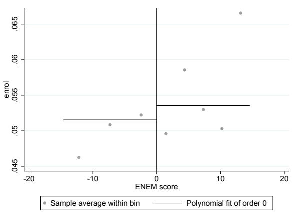
Figure 28 presents RD estimator sensitivity to excluding observations around the cutoff for grants in three outcomes: Enrolment, Graduation, and Dropout. In Panel A (Enrolment), the estimates remain relatively stable across different exclusion

Table 20 – Results from Normalized Multi RD of Grant Eligibility

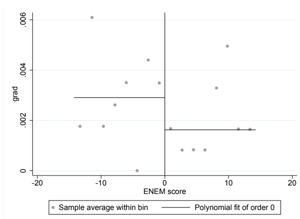
	Enrolment		Graduation		Drop Out	
	First Stage	Reduced Form	First Stage	Reduced Form	First Stage	Reduced Form
	(1)	(2)	(3)	(4)	(5)	(6)
A. Unrestricted						
1(S>0)	-.1332***	-.09693*	.06253***	-.47522	.0138***	.14343
	(.00173)	(.05166)	(.02119)	(.46131)	(.0018)	(.18138)
Intercept			16.45	16.45		
Bandwidth (h)	12.931	12.931	13.523	13.523	13.134	13.134
Observations	18810	18810	18232	18232	17351	17351
B. If Retaker						
1(Si>0)	-	-	.06827***	-.47432	.11473***	1.3241**
			(.02209)	(.41261)	(.03356)	(.56798)
Bandwidth (h)	-	-	16.45	16.45	8.341	8.341
Observations	-	-	432	432	217	217

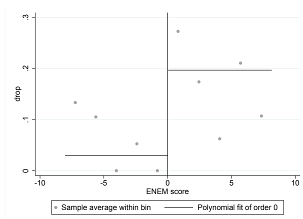
Note: Columns 1,3,5 show estimated first-stage model for the effect of passing 450-point threshold at any attempt (A) and second attempt(B). Remaining columns show estimated reduced-form models. Robust standard errors in parentheses. *: p-value< .1; **: p-value< .05; ***: p-value< .01

Figure 27 – Results from a Normalized Multi RD for grant Eligibility



A.Enrolment

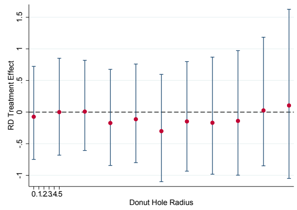




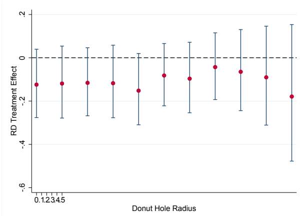
B.Graduation C.Drop Out

Note: RD plots of selected outcomes on ENEM scores. Vertical line represents the minimal policy cutoff. Each dot represent a mimicking variance quantile spaced bin (Cattaneo et al 2019).

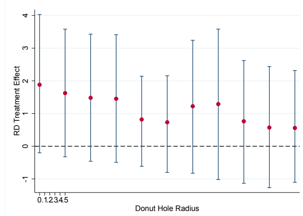
Figure 28 – Donut Hole Radius - Grants



A.Enrolment



B.Graduation



C.Drop Out

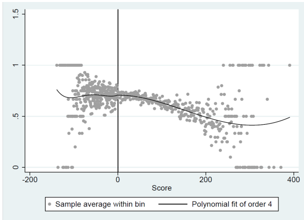
Note: RD estimator sensitivity to excluding observation in an interval (x-axys) around the cutoff. Red dots represent the point estimates and vertical lines represent respective 95% confidence intervals.

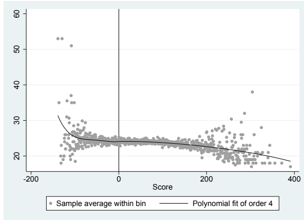
intervals, though the confidence intervals widen slightly as more observations near the cutoff are excluded, indicating some robustness in the enrolment effect. Panel B (Graduation) shows less consistency, with fluctuating point estimates and wide confidence intervals, suggesting the graduation outcome is more sensitive to the exclusion of observations around the cutoff. Panel C (Dropout) shows a fairly stable trend with minor fluctuations in point estimates, though the confidence intervals widen, indicating a weaker, less conclusive effect on dropout rates. Overall, the enrolment outcome appears the most robust to changes in the exclusion of data near the cutoff, while graduation and dropout show greater variability.

Covariate Balance

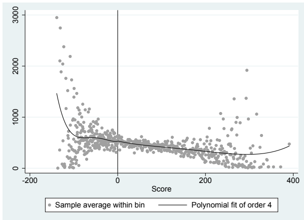
Figure 29 checks the behaviour of covariates around the cutoff. Scatterplot A indicates a slight imbalance in the sex variable as the scores move away from the cutoff. Particularly, on the right-hand side of the cutoff, there seems to be a slight decrease in the share of male students. However, the trend is relatively subtle, indicating that sex might not exhibit a significant jump or imbalance at the cutoff, which suggests that the allocation of students near the cutoff is fairly gender-balanced. The plot for age at enrolment in B shows some variation in age as the ENEM scores deviate from the cutoff. The scatterplot exhibits a more noticeable downward trend in age as scores increase. There seems to be a larger number of younger individuals scoring above the cutoff, as seen by the downward slope in the polynomial fit on the right side of the vertical line. This might suggest some

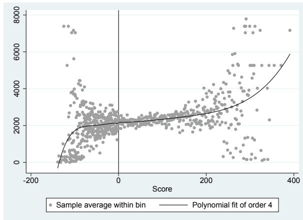
Figure 29 – Covariates' Balance: Loans





makebox[4cm]A.Sex B.Age at enrolment





C. Distance from University D. Population density of University

Note: RD plots of selected covariates on ENEM scores. Vertical line represents the minimal policy cutoff. Each dot represent a mimicking variance quantile spaced bin (Cattaneo et al 2019). Covariates

imbalance in terms of age distribution, with younger individuals more likely to be located above the cutoff threshold. However, the curve smooths out near the cutoff, meaning the jump is not sharp right at the threshold, so any imbalance is modest. Scatterplot C suggests a clear negative relationship between ENEM scores and distance from the university: individuals who score higher (to the right of the cutoff) tend to live closer to the university. There is a clear decreasing trend in the polynomial fit, which may indicate that those just above the cutoff live in closer proximity to the institution than those just below. This suggests potential imbalance in this covariate, which might require further controls in the RD model to ensure the distance factor doesn't confound the treatment effect. In the final panel the polynomial fit suggests that individuals with higher ENEM scores (above the cutoff) are associated with universities in more densely populated areas. The right-hand side of the plot (above the cutoff) shows a steeper upward trend, indicating that students who are eligible for the policy tend to be in regions with higher population density. This imbalance may indicate that geographical and urbanization factors could be playing a role in the cutoff, and these need to be accounted for in the regression analysis.

APPENDIX D – SEARCH AND REALLOCATION IN THE COVID-19 PANDEMIC

D.1 Data

This study utilises data from the UK Household Longitudinal Study (UKHLS) COVID-19 Study and the UK Labour Force Survey (LFS). This appendix describes how we define search activity, homogenise occupation classifications and other relevant variables across the two datasets.

Search activity: From the LFS questionnaire, we can quantify search activity from all three states of economic activity: Employment, Unemployment and Inactivity. By definition, all unemployed workers are looking for a job. We call an inactive worker "job searcher" if they self-declare as seeking work, but unavailable because of being a student, looking after family, temporarily sick or injured, long-term ill or disabled or due to other reasons or no reasons given. The LFS also asks employed workers whether they were searching for a replacement or additional job. If the answer is positive, these are on-the-job searchers.

Career changes: As in Carrillo-Tudela, Hobijn, She and Visschers (2016), we compute a career transition when a worker has changed employer, through a spell of non-employment or not, and reported an occupation or industry in the new job that is different from the one reported in the last job held. Because we use aggregate levels of occupation and industry classifications, the career transitions in this paper capture a substantial change in the nature of a worker's job. These transitions can occur from different states of the labour force. If a worker transitioned from a state of non-employment, our datasets inform the occupation or industry of their last job (if their previous job ended within the past eight years). A job-to-job change occurs when a worker is employed in two consecutive quarters and reports a job tenure of less than three months with no spells of unemployment in the second quarter.

Industry classifications: Both LFS and UKHLS use the Standard Industrial Classification (SIC) to code industries. Both datasets provide homogenised industry information for workers for the entire sample period based on the SIC2007. We use the industry section level from SIC2007, with 21 categories (ranging from A to U), to build our own industry code that portrays industry flows within 13

Table 21 – Industry section aggregation from SIC2007

Aggregated industry	Category	SIC 2007 Section	Category
Natural Resources	1	Section A: Agriculture, Forestry and Fishing	1
		Section B: Mining and Quarrying	2
		Section D: Electricity, Gas, Steam and Air Conditioning Supply	4
		Section E: Water Supply; Sewerage, Waste Management etc.	5
Manufacturing	2	Section C: Manufacturing	3
Construction	3	Section F: Construction	6
Wholesale and Retail	4	Section G: Wholesale and Retail Trade; Repair Of Motor Vehicles	7
Transportation and Storage	5	Section H: Transportation nd Storage	8
Accomodation and Food Services	6	Section I: Accommodation and Food Service Activities	9
ICT, Finance, and Professional Services	7	Section J: Information and Communication	10
		Section K: Financial and Insurance Activities	11
		Section M: Professional, Scientific and Technical Activities	13
Administration and Support	8	Section N: Administration and Support Services	14
Public Administration	9	Section O: Public Administration, Defence, Social Security	15
Education	10	Section P: Education	16
Health	11	Section Q: Human Health and Social Work	17
Arts	12	Section R: Arts, Entertainment and Recreation	18
Other Services	13	Section L: Real Estate Activities	12
		Section S: Other Service Activities	19
		Section T: Activities Of Households As Employers; Other Househols act.	20
(Excluded)	.	Section U: Activities of Extraterritorial Organisations And Bodies	21

categories. Industry sections from SIC2007 are aggregated by similarities in nature and employment growth patterns. Table 21 describes the SIC codes from which our definitions were constructed. Our sample excludes Section U: Activities Of Extraterritorial Organisations And Bodies.

Occupation classifications: Both the UK LFS and UKHLS use the Standard Occupational Classification (SOC) to code occupations. This study employs data from the first quarter of 2008 to the second quarter of 2021 using the SOC2010 occupational coding system introduced in the first quarter of 2011. Before 2011, occupations in the LFS were coded using SOC2000. To provide homogeneous occupations throughout the analysis period, we use a proportional mapping procedure to map SOC2000 4 and 3-digit occupations into 1-digit SOC2010. From 2011 to 2020, the LFS provides individuals current occupations coded in both SOC2010 and SOC2010. We use the observed mapping proportions to extrapolate the SOC. We focus on mobility across the 9 categories of major occupational groups.

Our proportional mapping procedure consists of, first, obtaining the proportion of each 4-digit SOC2000 category mapped into 1-digit SOC2010 by the LFS in each quarter from 2011Q1 to 2020Q4. We then get the normalised average of these

Table 22 – Occupation classification according to SOC2010

Abbreviated Occupation	Category	SOC2010 Group	Category
Managers	1	Managers, Directors and Senior Officials	1
Professional	2	Professional Occupations	2
A. Professional & Technical	3	Associate Professional and Technical Occupations	3
Admin	4	Administrative and Secretarial Occupations	4
Skilled Trades	5	Skilled Trades Occupations	5
Caring PS	6	Caring, Leisure and Other Service Occupations	6
Sales & CS	7	Sales and Customer Service Occupations	7
Machine Op	8	Process, Plant and Machine Operatives	8
Elementary	9	Elementary Occupations	9

ratios across periods. The person weights for the new occupations are obtained by multiplying the original person weights of each observation by the calculated proportions. We replicate this procedure to map the SOC2000 for the non-employed. In which we transform the SOC2000 3-digit codes for occupation in the previous job from the period before 2011 and also adjust SOC2020 3-digit codes occupation in the last job for individuals observed in 2021.

Throughout the paper, we refer to occupations using our own denomination. Those are short forms of the actual 1-digit SOC categories described in Table 22.

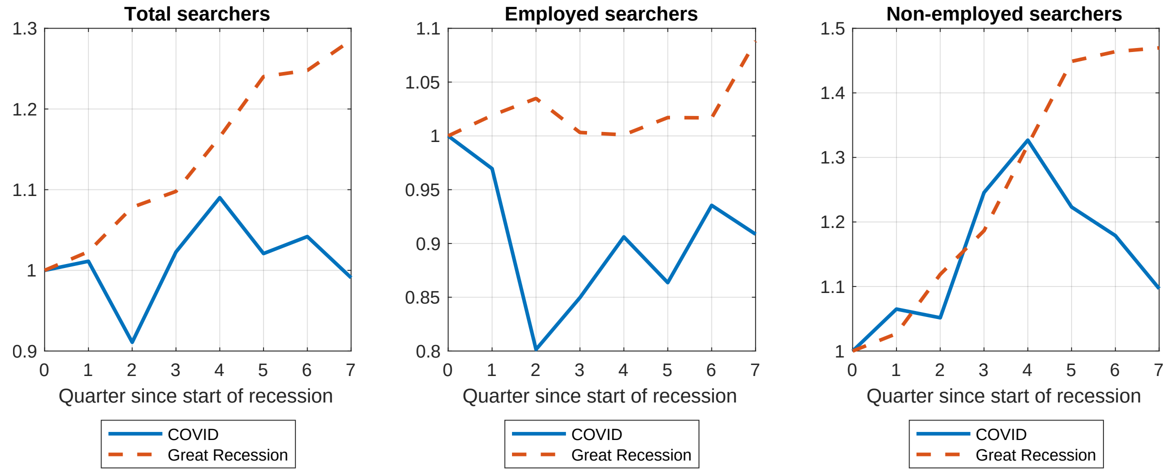
Skill levels: Low-skilled workers are defined as those with educational attainment below O-levels or GCSE grade C and equivalents. The medium-skilled range from those who achieved an O-level or GCSE grade A-C to those with an A-level qualification. The high skilled group includes all workers with post-school degrees from teaching qualifications to graduate studies.

D.2 Job Search: The Details

The LFS allows us to examine aggregate job search activity directly as it asks both employed and non-employed workers whether they are actively searching for a job. Figure 30 shows the change in the number of job search relative to the start of the Covid-19 pandemic. It presents these changes separately for the employed and non-employed (unemployed and marginally attached) as well as a comparison with the same data during the GR.

We observe that job search for the employed initially decreased in the pandemic, in contrast to the rise seen in the GR. Although to a lesser extent, this is also

Figure 30 – Change in Numbers Searching

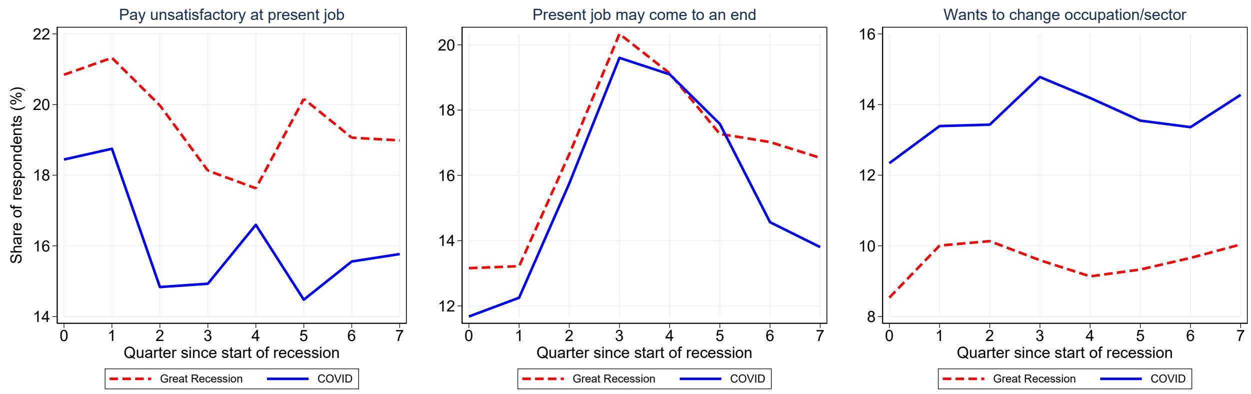


Note: All series are computed from the LFS. The LFS asks employed workers whether they were searching for a replacement or additional job. We define employed searchers as those who answer 'yes' to this question. Non-employed searchers are the sum of unemployed and inactive searchers. By definition, all unemployed workers are looking for a job. We define an inactive worker as a 'job searcher' if they self-declare as out-of-the-labour-force and unavailable to work currently, but are seeking work in the near future. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic in relation to their values observed during the quarter immediately preceding these events. Start dates ($t = 1$) for the Great Recession and pandemic recession are 2008 Q2 and 2020 Q1 respectively. All series are seasonally adjusted with a stable seasonal filter.

true of non-employed searchers as inactivity rose. For this latter group the initial fall was followed by a strong increase, such that the series converges with the one seen in the GR. The change in the number of non-employed searchers is principally due to the rise in the number of unemployed (all of whom by definition search). In contrast, the change in the numbers of employed searchers is principally due to a changes in the fraction of employed that search. Note that the recovery in the numbers of employed searchers occurs at the same time as the recovery in aggregate vacancies (see Figure 11) suggesting search behaviour responded to aggregate demand. The small rise in aggregate search activity at the extensive margin over the pandemic, shown in Figure 30, is also matched by a small rise in search activity at the intensive margin, as measured by the average number of search channels used by job searchers (see Figure 33).

Figure 31 plots the reasons stated for job search stated by employees looking for an alternative job, and highlights an important feature of the Covid-19 pandemic. First we observe that looking to move jobs due to dissatisfaction with current employment pay decreases and does not show much sign of recovery. This is consistent with the observed persistent rise in the share of searchers reporting they are searching due to fear of job loss. The increased fraction of searchers looking to move occupation or industry suggests that individuals have been responsive to the large differential experiences across occupations and industries observed during the pandemic. This

Figure 31 – Top 3 Reasons for Job Search Among Employees



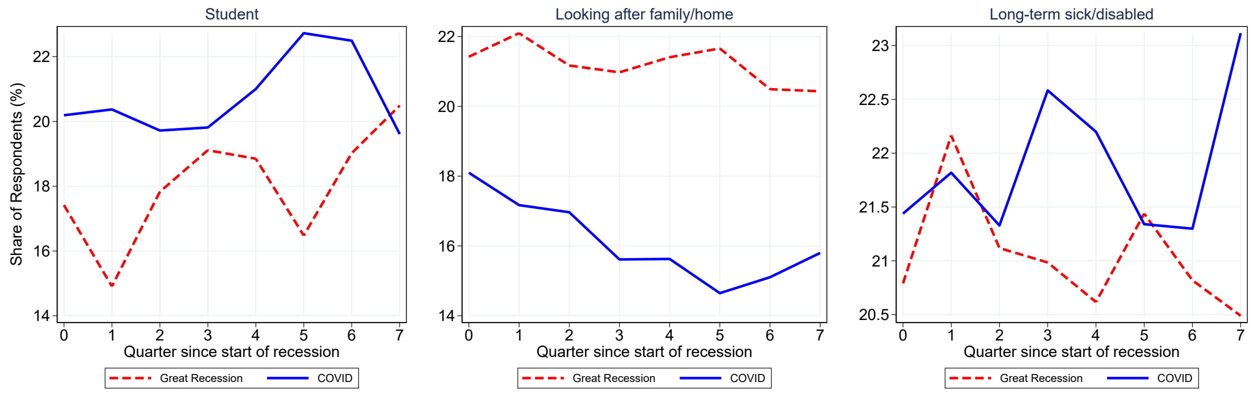
Note: All series are computed from the LFS. The LFS asks employed workers who report searching for an additional or replacement job why they are searching. We report the three most popular answers given as proportion of all responses. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic. Start dates ($t = 1$) for the Great Recession and pandemic recession are 2008 Q2 and 2020 Q1 respectively. All series are seasonally adjusted with a stable seasonal filter.

is important as the direction of job search is a crucial determinant of reallocation in the economy, which in turn has an important bearing on the recovery of the labour market and aggregate productivity.

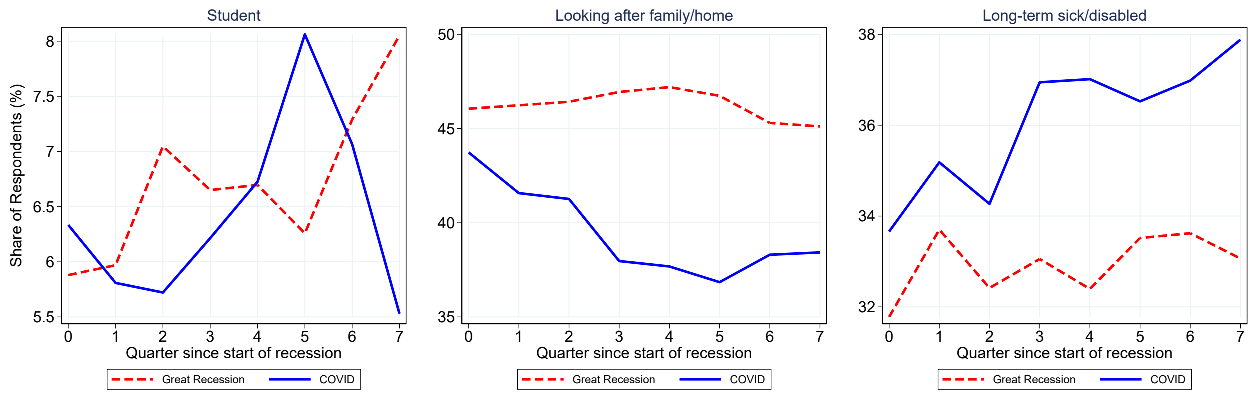
Just as we can look at the reasons for job search the LFS also asks inactive workers why they are not searching. Figure 32 shows the top 3 reasons why individuals state they are not looking for a job. There is a marked increase in those giving long-term sickness/disability and studying as a reason for not searching. These results continue to hold when looking just at prime-age workers (aged 25-55), as shown in the bottom row of Figure 32. Perhaps, surprisingly the numbers stating they are inactive due to looking after family/home decrease during the pandemic despite school closures.

Search intensity, as measured by the average number of search channels used by job searchers, increased both during the current downturn and in the GR albeit more mildly. However, in the GR this was driven by increased search intensity by unemployed workers whereas employees have increased their search intensity more in the current downturn. This may be a compositional effect i.e. we have seen that the numbers of employed searchers decreases in the pandemic while the numbers of unemployed searchers increase: if the marginal searcher searches less intensely, then we would expect the patterns above.

Figure 32 – Top 3 Reasons for Not Job Searching



Aged 16-65



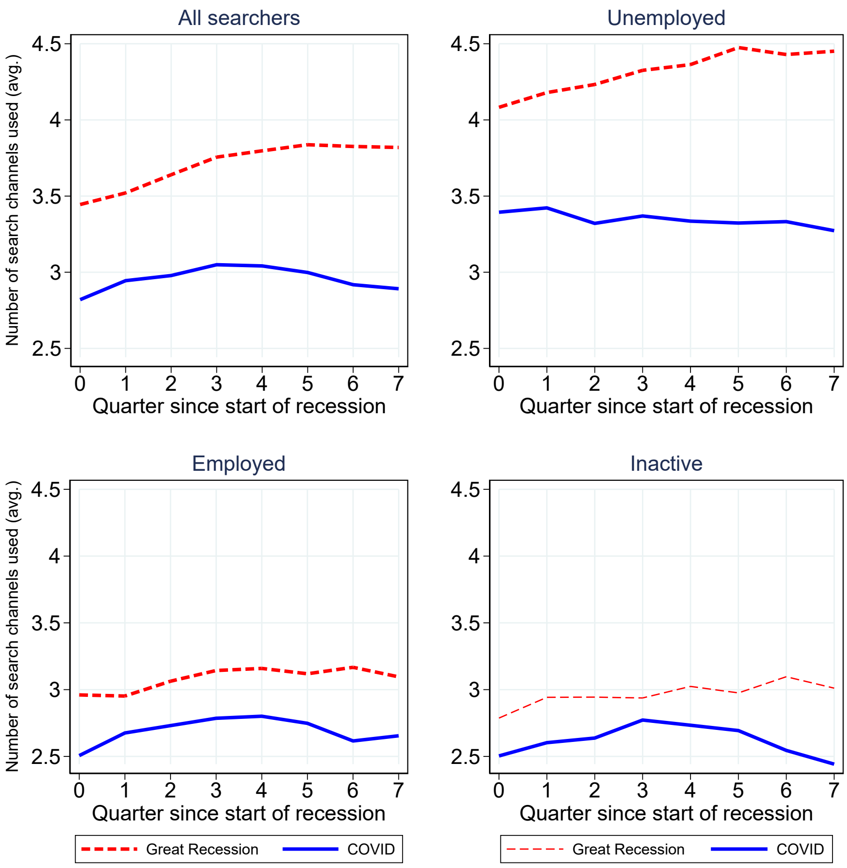
Aged 25-55

Note: All series are computed from the LFS. The LFS asks employed workers who report searching for an additional or replacement job why they are searching. We report the three most popular answers given as proportion of all responses. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic. Start dates ($t = 1$) for the Great Recession and pandemic recession are 2008 Q2 and 2020 Q1 respectively. All series are seasonally adjusted with a stable seasonal filter.

D.3 Job Search: Whs's Looking?

The COVID-19 study of the UKHLS asks employed and non-employed workers whether they have looked for a new job in the last 4 weeks. This allows us to examine whether any of the demographic or employment characteristics that influence the industry or occupation of job sought (see Table 11) are driven by the composition of those searching. We look at this in a probit regression with results reported in Table 23. The dependent variable is whether the individual searched in the last four weeks and the independent variables include dummy variables for whether the individual

Figure 33 – Search Intensity: Covid-19 vs the Great Recession



Note: All series are computed from the LFS. The LFS asks workers who report searching for a job what search channels they are using, and we define search intensity as the average number of channels used by each searching worker. The series are presented for the first seven quarters of the Great Recession and the Covid-19 pandemic. Start dates ($t = 1$) for the Great Recession and pandemic recession are 2008 Q2 and 2020 Q1 respectively. All series are seasonally adjusted with a stable seasonal filter.

is female, young (aged 16-34), has low education (maximum attainment of GCSE or less), is white, from London, was working in a declining industry or occupation in 2019 (see main text for definition of an industry/occupation that declined during the pandemic), and individual fixed effects from a Mincer wage regression. We do separate regressions for the employed (left hand column) and non-employed (right hand column). We find that young respondents are significantly (at the 5% level) more likely to search, when employed and non-employed. Overall, there are not strong demographic selection effects into job searching, with the exception of age, suggesting the impacts documented in Table 11 are not driven by the composition of those selecting into search.

Table 23 – Probability of Searching

	(1) Looking for Job: Employed	(2) Looking for Job: Non-employed
Female	-0.00539	-0.0309
	(0.00779)	(0.0603)
Young (16-34)	0.0429***	0.0441
	(0.00957)	(0.0569)
Low Educ. (GCSE or less)	-0.0209**	-0.0776
	(0.0105)	(0.0717)
White	-0.00974	-0.0779
	(0.0114)	(0.0897)
London	0.0186*	0.106
	(0.0107)	(0.0713)
Declining Source Ind.	0.0165**	0.165***
	(0.00720)	(0.0506)
Declining Source Occ.	0.00672	-0.0402
	(0.00731)	(0.0493)
Individual Fixed Effects	-0.0123	0.0322
	(0.00815)	(0.0475)
Observations	12410	1008

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Data from the COVID-19 Study, Job Search Module of the UKHLS. Classification of sectors as expanding or contracting is based on employment changes from the LFS, as detailed in the text. Table shows the results (as marginal effects) of Probit estimations where the dependent variable takes the value of one if the individual searches for work in the last four weeks.

D.4 Job Search: Targeting New Occupations

The COVID-19 study of the UKHLS asks workers searching for a new job if they are targeting an occupation they have ever previously performed. Table 24 shows results from a Probit regression where the dependent variable is an indicator variable taking the value of one if the respondent targets a new occupation i.e. one never previously performed. We see that non-employed respondents are significantly less likely to target a completely new occupation. This is also true of young workers, consistent with occupation changing being more frequent in a worker's early career.[1]

D.5 Transition Matrices by Occupation and Industry

D.5.1 Targeted occupational transition matrix

A novelty of our data is that it allows us to construct a "targeted" transition matrix, relating the occupations performed by individuals in 2019 to these individuals' targeted occupations during the Covid-19 pandemic. This helps analyse the degree of targeted attachment to an occupation and contrast it with the realised

[1]There may also be a mechanical effect present, even with random occupation choice.

Table 24 – Probability of Targetting A New Occupation

	(1) Targets New Occupation
Female	0.0962*
	(0.0511)
Young (16-34)	0.238***
	(0.0516)
Low Educ. (GCSE or less)	-0.0640
	(0.0585)
White	0.136*
	(0.0772)
London	0.119*
	(0.0690)
Not Employed	-0.186***
	(0.0553)
Declining Source Ind.	-0.0163
	(0.0523)
Declining Source Occ.	0.0354
	(0.0508)
Observations	810

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Data from the COVID-19 Study, Job Search Module of the UKHLS. Classification of sectors as expanding or contracting is based on employment changes from the LFS, as detailed in the text. Table shows the results (as marginal effects) of Probit estimations where the dependent variable takes the value of one if the individual targets an occupation they have never previously worked in.

transition patterns. The top panel of Table 25 presents the targeted transition matrix. It shows that those individuals who in 2019 were employed in the declining occupations during the pandemic, exhibited a lower degree of attachment relative to those individuals that in 2019 were employed in the expanding occupations. In particular, we observe a degree of attachment (defined as the share of those from a given occupation in 2019 saying they are targeting a job move in the same occupation) that ranges between 17.4% and 49.4% among those in declining occupations and one that ranges between 34.2% and 65.5% among those in expanding occupations.

The middle panel of Table 25 presents the observed occupational transition matrix during 2020 using LFS data. Although not composed by the same sample of individuals used to construct the targeted transition matrix (based on the UKHLS), it provides an estimate of the extent to which targeting an occupation translates into employment in such an occupation. By subtracting both matrices we can observe that, in the majority of cases, the proportion of searchers who targeted those occupations they performed in 2019 is very similar to the proportion of actual occupational stayers during 2020. However, it is among those who targeted a different occupation that we can observe the larger differences between the proportion of

Table 25 – Targeted and realised occupation transition matrices

Targeted occupation transition matrices, UKHLS (%)

	Expanding:				**Declining:**						
Targeted Occ. in 2020 Occ. in 2019	Professional	Assoc. Profess. & Technical	Admin. & Sec.	Sales & Cust. Serv.	Managers	Skilled Trade	Caring & Leisure	Process & Machine Op.	Elementary	Expanding	Declining
Expanding:											
Professional	62.96	15.29	5.32	3.38	7.91	2.64	0.68	0.52	1.31	86.94	13.06
Associate Profess. & Technical.	19.29	45.88	10.42	3.98	8.17	1.51	4.90	1.67	4.18	79.57	20.43
Admin. & Sec.	11.54	9.11	65.45	2.57	2.64	0.00	5.71	2.79	0.18	88.68	11.32
Sales & Customer Serv.	6.69	17.88	5.43	34.15	1.40	2.22	24.31	5.93	1.98	64.15	35.85
Declining:											
Managers	15.46	12.49	14.22	5.00	17.40	27.93	2.72	2.30	2.48	47.18	52.82
Skilled Trade	3.67	14.90	0.00	16.05	1.43	38.05	13.38	5.95	6.57	34.61	65.39
Caring & Leisure	11.96	5.05	4.33	6.36	0.53	0.00	49.41	1.88	20.48	27.70	72.30
Process & Machine Op.	13.92	4.72	1.68	3.17	5.06	8.01	0.93	35.50	27.01	23.49	76.51
Elementary	6.40	20.88	9.71	23.98	0.00	0.88	5.02	2.34	30.79	60.97	39.03
Expanding	26.22	25.11	19.57	9.41	5.64	1.62	7.79	2.44	2.20	80.31	19.69
Declining	10.48	11.91	6.95	11.63	4.33	10.61	17.76	6.99	19.35	40.97	59.03

Realised occupation transition matrices, UKLFS 2020 (%)

	Expanding:				**Declining:**						
Targeted Occ. in 2020 Occ. in 2019	Professional	Assoc. Profess. & Technical	Admin. & Sec.	Sales & Cust. Serv.	Managers	Skilled Trade	Caring & Leisure	Process & Machine Op.	Elementary	Expanding	Declining
Expanding:											
Professional	58.64	12.70	5.34	8.97	4.38	1.05	4.20	1.55	3.18	85.65	14.35
Associate Profess. & Technical.	9.37	60.21	0.00	2.35	6.63	0.00	2.50	5.32	13.62	71.93	28.07
Admin. & Sec.	8.76	22.42	48.61	5.85	0.00	8.11	4.12	0.00	2.14	85.64	14.36
Sales & Customer Serv.	2.49	8.51	18.66	31.59	4.10	3.02	15.67	6.51	9.45	61.26	38.74
Declining:											
Managers	12.58	33.77	5.06	0.00	35.52	0.00	0.00	2.21	10.86	51.41	48.59
Skilled Trade	20.19	0.00	0.00	8.89	0.00	21.54	9.81	7.05	32.53	29.07	70.93
Caring & Leisure	8.17	1.30	2.64	5.93	0.00	0.00	69.09	0.00	12.87	18.04	81.96
Process & Machine Op.	2.40	0.00	0.00	8.17	3.59	0.00	10.90	74.93	0.00	10.57	89.43
Elementary	3.72	9.22	9.20	10.85	1.85	7.23	18.66	1.99	37.27	32.99	67.01
Expanding	24.69	24.22	15.94	12.10	3.93	2.69	6.46	3.22	6.74	76.96	23.04
Declining	7.48	9.77	4.93	7.28	7.55	4.51	24.65	12.37	21.47	29.46	70.54

Realised occupation transition matrices, UKLFS 2016-19 (%)

	Expanding:				**Declining:**						
Targeted Occ. in 2020 Occ. in 2019	Professional	Assoc. Profess. & Technical	Admin. & Sec.	Sales & Cust. Serv.	Managers	Skilled Trade	Caring & Leisure	Process & Machine Op.	Elementary	Expanding	Declining
Expanding:											
Professional	61.84	15.68	4.00	2.20	6.26	0.99	5.42	0.78	2.83	83.72	16.28
Associate Profess. & Technical.	17.27	49.64	7.08	7.69	8.63	2.67	4.05	0.91	2.07	81.67	18.33
Admin. & Sec.	6.36	14.45	48.25	7.00	7.03	1.66	8.27	1.18	5.81	76.06	23.94
Sales & Customer Serv.	4.01	14.22	14.75	30.97	3.21	3.17	7.18	3.18	19.30	63.96	36.04
Declining:											
Managers	12.10	20.85	7.48	1.60	45.28	2.79	3.49	1.61	4.81	42.02	57.98
Skilled Trade	2.49	8.31	3.46	3.04	3.15	63.46	1.56	8.36	6.17	17.30	82.70
Caring & Leisure	8.38	6.87	10.05	13.64	1.05	2.44	46.36	4.84	6.37	38.93	61.07
Process & Machine Op.	2.46	3.45	5.65	4.19	4.07	11.66	4.05	41.86	22.59	15.76	84.24
Elementary	2.65	7.08	11.19	14.11	2.84	3.15	6.56	6.99	45.43	35.03	64.97
Expanding	25.12	23.94	16.41	11.45	6.31	2.09	6.06	1.47	7.15	76.92	23.08
Declining	5.28	8.97	8.74	9.49	9.32	11.88	13.39	9.71	23.22	32.47	67.53

Targeted versus realised occupational transitions during the Covid-19 pandemic. Data for targeted transitions are from the COVID-19 Study, Job Search Module of the UKHLS, and for realised are from the LFS.

individuals targeting certain occupations and the proportion of actual transitions.

In particular, we see that about 24% of those individuals who performed Elementary occupations in 2019 targeted Sales & Customer Services jobs. The realised transition matrix shows that less than half of this proportion actually found jobs in Sales & Customer Services and instead 18.7% found employment in Caring and Leisure occupations. We also highlighted that 20.9% of Elementary workers in 2019 targeted Associate Professionals jobs, but we observe that the realised transition in this direction only achieves 9.2%. Thus our evidence suggests that those in the worse performing occupations that targeted the better performing ones were not able to access them.

To investigate whether the gap between the targeted and realised transition matrices arises because individuals were basing their search on past transition probabilities, the bottom panel of Table 25 presents the transition matrix for the 2016-2019 period also obtained from the LFS. The average absolute difference between the targeted and realised 2020 matrices is 7.02 percentage points and between the targeted and the 2016-2019 matrices is 5.14 percentage points. This comparison suggests that there is some degree of past behaviour that could be driving a wedge between targeted and realised occupational transition matrices during the pandemic.

D.5.2 Targeted industry transition matrix

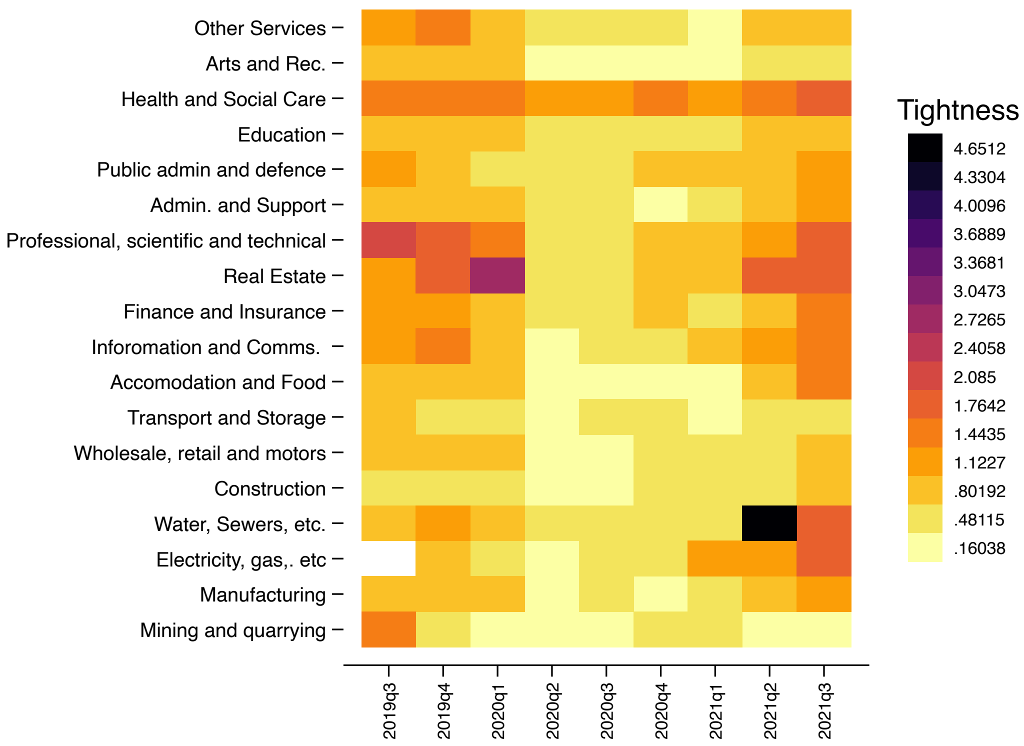
Table 26 shows the industry attachment of individuals during Covid-19 recessions in the same way as done with occupations. As before, we include the targeted transition matrix from the COVID-19 study of the UKHLS and the realised transition matrices from the LFS in 2020 and in 2016-19. However, direct comparison between the targeted and realised transition matrices is not possible as there is likely a substantial discrepancy between how industries were coded in the COVID-19 study of the UKHLS (self-chosen by respondent) and in the LFS (coded by professionals).

The bottom panel of Table 26 presents the realised industry transition matrix for the period 2016-2019. It shows that the proportion of individuals who did not switch industries after changing employers increased during the Covid-19 pandemic.

D.6 Labour Market Tightness

Labour market tightness (vacancies/unemployment) has still surged well above its pre-pandemic level as of 2021 Q3 due to the strength of vacancy creation and hiring. Figure 37 shows that there has been considerable heterogeneity by sector.

Figure 37 – Labour Market Tightness by Sector



Note: Vacancy data by sector is taken from ONS vacancy survey, unemployment data by sector is taken from the LFS. Tightness is defined as the ratio of vacancies to unemployment.

Table 26 – Targeted and realised industry transition matrices

Targeted industry transition matrices, UKHLS (%)

	Expanding:					**Declining:**									
Targeted Ind. in 2020 Ind. in 2019	Natural Resources	ICT, Fin. & Profess.	Public Admin.	Education	Health	Manufact.	Construc.	Wholesale & Retail	Transport & Storage	Accom. & Food	Admin.& Support	Arts	Other Services	Expanding	Declining
Expanding:															
Natural Resources	30.81	4.71	3.19	0.00	0.00	0.00	37.86	0.00	4.20	2.22	15.94	0.00	1.06	38.71	61.29
ICT, Finance, & Profess.	1.21	35.78	2.88	4.03	6.46	0.54	1.43	4.51	0.55	0.00	7.36	11.94	23.32	50.35	49.65
Public Admin.	0.00	10.91	23.28	9.36	14.66	0.00	0.00	0.00	5.73	0.00	2.36	11.15	22.56	58.20	41.80
Education	0.74	6.75	1.21	43.12	11.92	0.14	3.11	4.02	0.00	0.41	3.35	10.32	14.91	63.74	36.26
Health	0.00	4.85	7.84	12.77	45.99	0.00	0.00	0.35	0.52	0.53	6.19	0.43	20.54	71.44	28.56
Declining:															
Manufacturing	11.05	12.33	7.41	1.44	8.81	15.71	0.42	10.89	6.22	3.74	5.44	1.77	14.77	41.03	58.97
Construction	0.00	10.48	8.54	0.00	3.78	13.42	39.52	4.86	0.00	2.33	1.43	0.00	15.64	22.80	77.20
Wholesale & Retail	0.55	5.16	5.67	2.75	21.69	3.17	2.82	15.67	6.40	5.15	2.06	9.06	19.83	35.83	64.17
Transport & Storage	0.12	3.30	1.31	1.55	16.92	5.96	1.70	11.32	16.51	4.87	6.62	2.40	27.43	23.19	76.81
Accomm. & Food	0.00	8.61	1.22	3.31	7.70	0.00	2.36	9.94	3.19	11.94	5.55	12.72	33.46	20.83	79.17
Admin. & Support	0.00	3.60	18.49	3.64	3.24	0.00	5.21	17.08	8.89	0.00	8.15	0.72	30.99	28.96	71.04
Arts	15.03	0.26	1.85	4.39	4.70	0.00	0.00	13.22	0.57	3.50	1.59	22.46	32.42	26.23	73.77
Other Services	2.95	11.13	0.00	6.13	13.93	0.00	0.00	6.24	0.00	0.00	10.14	9.30	40.19	34.13	65.87
Expanding	1.49	15.17	5.47	17.12	21.99	0.20	2.36	2.51	0.83	0.36	5.89	7.09	19.51	61.24	38.76
Declining	4.03	6.00	5.10	2.91	11.91	4.03	3.97	12.51	5.44	4.98	4.18	9.19	25.75	29.96	70.04

Realised industry transition matrices, UKLFS 2020 (%)

	Expanding:					**Declining:**									
Destination Ind. Source Ind.	Natural Resources	ICT, Fin. & Profess.	Public Admin.	Education	Health	Manufact.	Construc.	Wholesale & Retail	Transport & Storage	Accom. & Food	Admin.& Support	Arts	Other Services	Expanding	Declining
Expanding:															
Natural Resources	44.87	12.44	9.85	0.00	7.49	0.00	13.72	5.84	0.00	0.00	0.00	5.78	0.00	74.66	25.34
ICT, Finance, & Profess.	1.34	67.87	6.66	1.91	4.49	0.65	2.06	3.84	3.92	1.29	3.93	0.74	1.30	82.27	17.73
Public Admin.	0.00	8.81	42.54	2.44	6.04	0.00	3.76	4.43	2.20	17.80	11.98	0.00	0.00	59.83	40.17
Education	0.00	7.09	3.77	59.01	9.09	2.05	0.00	9.32	0.00	0.00	0.72	6.52	2.44	78.96	21.04
Health	0.00	2.35	3.26	8.27	65.29	0.52	0.87	7.90	1.83	3.73	0.00	3.42	2.56	79.17	20.83
Declining:															
Manufacturing	0.58	4.74	2.17	6.73	7.08	51.11	2.09	8.36	6.92	2.71	3.00	4.05	0.47	21.30	78.70
Construction	5.35	8.31	0.00	1.64	4.13	3.14	73.15	2.23	0.00	0.00	0.00	2.05	0.00	19.43	80.57
Wholesale & Retail	1.71	12.64	3.96	2.52	9.77	7.91	2.32	41.43	2.54	3.38	6.31	2.25	3.26	30.60	69.40
Transport & Storage	5.70	5.65	3.90	6.63	6.43	6.40	1.03	4.49	46.94	0.00	7.68	5.16	0.00	28.31	71.69
Accomm. & food services	2.25	9.75	0.17	7.01	8.29	0.90	2.28	14.54	3.82	47.48	1.26	0.92	1.35	27.47	72.53
Admin. & Support	0.00	8.81	4.85	0.41	4.33	8.16	10.62	4.95	6.71	7.73	39.77	0.00	3.66	18.40	81.60
Arts	0.00	6.51	11.40	0.00	2.31	8.45	0.00	13.07	7.08	0.00	0.00	51.18	0.00	20.22	79.78
Other Services	5.10	5.44	0.00	6.59	9.02	0.00	0.00	27.15	0.00	0.00	6.62	0.00	40.08	26.16	73.84
Expanding	2.89	30.16	8.47	18.18	18.71	0.90	2.03	6.28	2.10	2.83	2.76	3.02	1.68	78.40	21.60
Declining	2.36	8.84	2.91	4.18	7.31	10.31	8.35	18.64	8.49	11.64	8.02	4.69	4.27	25.59	74.41

Realised industry transition matrices, UKLFS 2016-19 (%)

	Expanding:					**Declining:**									
Destination Ind. Source Ind.	Natural Resources	ICT, Fin. & Profess.	Public Admin.	Education	Health	Manufact.	Construc.	Wholesale & Retail	Transport & Storage	Accom. & Food	Admin.& Support	Arts	Other Services	Expanding	Declining
Expanding:															
Natural Resources	41.62	7.80	3.72	2.33	4.10	14.73	1.63	7.35	3.82	2.42	7.50	0.79	2.17	59.58	40.42
ICT, Finance, & Profess.	0.94	61.43	3.50	4.32	4.17	3.75	2.01	5.80	0.99	1.41	5.53	1.88	4.27	74.36	25.64
Public Admin.	0.38	13.85	41.36	3.60	7.62	2.19	4.00	7.88	2.54	2.43	7.60	2.10	4.46	66.81	33.19
Education	0.59	4.95	2.94	63.73	6.77	2.22	1.00	5.09	1.52	2.96	3.81	2.28	2.14	78.98	21.02
Health	0.39	3.50	3.86	6.81	63.57	1.60	1.44	8.98	0.76	2.83	4.52	0.16	1.57	78.14	21.86
Declining:															
Manufacturing	1.42	8.20	0.94	3.14	4.86	45.09	4.94	12.50	6.87	4.77	3.82	2.34	1.11	18.56	81.44
Construction	3.00	4.39	0.77	2.18	2.02	6.74	60.64	6.50	3.19	3.23	5.11	1.37	0.85	12.36	87.64
Wholesale & Retail	1.28	8.90	3.44	5.54	6.47	4.92	2.89	42.88	4.12	9.73	6.17	1.67	2.01	25.62	74.38
Transport & Storage	0.73	5.29	3.05	2.46	2.13	12.23	4.28	11.25	42.78	4.79	8.88	0.91	1.24	13.65	86.35
Accomm. & food services	1.62	6.40	2.35	3.88	9.66	3.77	1.76	14.09	3.13	44.12	5.12	2.35	1.76	23.90	76.10
Admin. & Support	0.26	7.99	3.38	4.23	4.37	6.19	3.25	11.11	4.84	5.66	43.64	1.17	3.90	20.23	79.77
Arts	2.09	8.26	1.86	5.53	5.88	5.45	3.97	14.59	3.85	7.60	1.77	35.98	3.16	23.62	76.38
Other Services	1.03	9.26	3.69	7.35	7.43	5.25	1.46	9.65	2.70	5.77	8.20	4.02	34.20	28.75	71.25
Expanding	2.65	26.31	7.19	18.20	20.58	3.23	1.81	6.75	1.34	2.28	5.18	1.49	3.00	74.93	25.07
Declining	1.40	7.58	2.52	4.34	5.94	11.72	8.02	20.58	7.00	13.67	9.35	3.96	3.93	21.77	78.23

Targeted versus realised industry transitions during the Covid-19 pandemic. Data for targeted transitions are from the COVID-19 Study, Job Search Module of the UKHLS, and for realised are from the LFS.