University of Essex

Doctoral Thesis

Three Essays on Migration Economics

Author:
Julián Costas-Fernández

Supervisor:
Dr. Ben Etheridge
Prof. Matthias Parey

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Department of Economics
University of Essex

November 2019

Declaration of Authorship

I, Julián Costas-Fernández, declare that this thesis titled, "Three Essays on Migration Economics" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Abstract

This thesis examines the selection of immigrants and their impact on the receiving economy.

After an introductory first chapter, I present an analysis of Borjas model of selection extended to multiple locations. In this extension, selection is determined by earnings' dispersion alone only for the most and least disperse locations. Therefore, it highlights that, what determines selection, is the ranking of locations rather than the relative dispersion of earnings between home and destination. Using interstate US migration, I provide stochastic dominance relations on pre-migration earnings that support the implications of the model. These give a stronger test for selection than selection on means.

The third chapter presents evidence on the effect of immigrants on aggregate labour productivity in the UK. I exploit variation on past settlement of natives across industries and regions to estimate the effect of immigrants on labour productivity. My estimates show that increasing the relative supply of immigrant labour has a positive effect on labour productivity. I show that part of this effect works through accumulation of capital stocks and provide evidence suggesting that immigrants trigger the development of technologies that complement them. Thus, I show that altering the labour mix produces effects that go beyond simple differences in marginal products and affects the accumulation of other inputs and technologies.

In the fourth chapter, Greta Morando and I exploit cross-cohort variation within majors and universities to estimate the effect of foreign peers on native students. We show that increasing the share of EU students lowers the probability of entering a university major. But, conditional on university and major, foreign peer effects on educational and early labour market outcomes are mild. Therefore, our research shows that there are no large foreign peer effects in higher education. However, it also suggests that there is scope for foreign students affecting natives' outcomes by shaping the universities and majors natives attend.

Acknowledgements

To my parents and siblings.

I would like to thank my supervisors Ben Etheridge and Matthias Parey for their continuous support and advice. Without them the research contained in this thesis wouldn't have been possible. I would also like to express my deep gratitude to Myra Mohnen, Michel Serafinelli, Tim Hatton, Giovanni Mastrobuoni and Gordon Kemp. Throughout my Ph.D they were always up for a chat and my work has greatly improved thanks to them.

I have been fortunate to share my Ph.D experience with wonderful friends. I would like to thank Patri Abal-Pino, Andrea Albertazzi, Alex Alonso-Campo, Patryk Bronka, Emil Kostadinov, Simon Lodato, Greta Morando, Ipek Mumcu and Federico Vaccari. A very special thanks goes to Cristina Corredor-Bejar for her continuous support while navigating through my studies.

The research here included has been funded by the ESRC. Moreover, a substantial part of the research contained in the second chapter of this thesis was carried while working as a Ph.D intern at the Migration Advisory Committee. As such, I am deeply grateful to the MAC. Particularly, I would like to thank Alan Manning and Maria del Castillo from whom my work has greatly benefited. Finally, the research contained in the fourth chapter of this thesis is the result of my collaboration with Greta Morando. Each of us has contributed a 50% to this project.

Contents

List of Figures

Chapter 1

Introduction

An increasingly large number of individuals live and work outside their country of birth. This has led, over the last twenty years, to the immigrant population share in developed countries doubling and reaching 14% of the total population by 2019.[1] Unsurprisingly, immigration and its consequences have become a central issue for academics, policymakers and the general public.[2]

The economic literature has recognised the importance of migration, at least, as early as Ravenstein (1885). In the modern literature, one of the most influential papers on migration economics is Borjas (1987) work on selection of migrants. Borjas uses the occupational framework introduced by Roy (1951) and models migration decisions as an earnings maximization problem. The main implication of his model is that earnings dispersion alone determines the direction of selection. High-skilled workers move to disperse locations that give them a better chance to increase their earnings and low-skilled workers choose more equal locations to insure themselves against extreme bad earnings draws (Parey et al. 2017).

Since Borjas (1987), immigrant self-selection has received widespread attention. For example, migrant selection from Mexico to the US has been the subject of intense debate (Ambrosini and Peri 2012; Chiquiar and Hanson 2005; Fernandez-Huertas Moraga 2011; Kaestner and Malamud 2014; Orrenius and Zavodny 2005). A common element to most of the migrant selection literature that followed Borjas (1987), is that it uses a two-location model. This is true even when studying migration phenomena where workers face multiple destinations (e.g. Belot and Hatton 2012; Borjas, Kauppinen, et al. 2018; Feliciano 2005; Fernandez-Huertas Moraga 2013; Parey et al. 2017). Borjas, Bronars, et al. (1992) is, to the best of my knowledge, the only exception. In their study of interstate migration, Borjas, Bronars, et al. (1992) present a Roy model of selection with perfect transferability of skills. This perfect transferability of skills simplifies the model, but it is a restriction regarding Borjas original

[1] According to United Nations data.

[2] As a public issue, immigration has gained salience across the political spectrum (Dancygier and Margalit 2019) and this, ultimately, translates into public policy (Facchini et al. 2016). For example, in the UK, immigration has played a key role in one of the UK biggest policy changes, leaving the European Union (Goodwin and Milazzo 2017). This growing interest in immigration gives a compelling argument to study its causes and consequences.

model that one may not want to impose. Particularly when studying international migration.[3]

In the second chapter of this thesis, I present an analysis of Borjas' model extended to multiple locations and general transferability of skills. I show that this extension of the model predicts positive (negative) selection to the most (least) disperse location. This highlights something that gets overlooked in the two-location model, what matters for selection is the ranking of locations in terms of earnings dispersion. This is in contrast with the most common interpretation in the literature that interprets the selection mechanism as given by relative earnings dispersion between home and destination. As a result, this interpretation of the selection mechanism guides, in some way, most of the analysis framed within the Borjas model. For example, Belot and Hatton (2012) regress measures of worker quality on differences in returns to education at home and destination. And Abramitzky and Boustan (2017), in their review of immigrant selection, note that some literature (Feliciano 2005; Grogger and Hanson 2011; Jasso et al. 2004; Kennedy et al. 2015) finds positive immigrant selection to the US independently of differences between home and destination.

That the home location plays no role in selection through differences in dispersion has the following intuition. In most models of migration (e.g. Dahl 2002; Kennan and Walker 2011) the home location is defined as a preference, or location amenity, shifter. This is, the home location does not enter the earnings equation or, in any case, only displaces the location of the net earnings distribution. This is also true in Borjas' model. Then, given that selection is driven by relative returns to skill and these are not affected by the home location, the home location must play no role in selection.

Although home location plays no special role on whether there is positive or negative selection, it affects the intensity of selection. However, this is intuitive and results from using earnings obtained at home to measure skill. To see the intuition, let me give an extreme example. Imagine that all workers are ex-ante equal, then the dispersion of earnings in the home location is zero, everyone earns the same wage, and there can be no-selection.[4] If workers have heterogeneous skills, then there can be selection and the dispersion of earnings at home will be non-zero. In addition, the more heterogeneous the skills of workers are, the greater the degree of selection can be. Moreover, in the model, if one substitutes earnings at home for some other measure of skill, the effect of earnings dispersion at home on the intensity of selection disappears. This shows that the effect of home location earnings dispersion is given by how one measures skill, not the mechanism of selection.

When exploring the implications of my multiple locations Borjas' model, I follow Borjas, Kauppinen, et al. (2018) and study stochastic dominance relations between earnings distributions. This has two advantages when compared with selection

[3]See the extensive literature on immigrant downgrading following Chiswick (1978). See also the review by Dustmann and Görlach (2015).

[4]This is inside the one price Roy model. In other set-ups it is possible to generate wage dispersion with ex-ante identical workers (e.g. Postel-Vinay and Robin 2002).

on mean earnings. First, stochastic dominance relations produce simpler expressions when dealing with multiple locations. Second, stochastic dominance produces a stronger test for the implications of the model. This is because, first order, stochastic dominance implies ordering of means. However, the opposite is generally not true.

To test the implications of the model, I use US interstate migration. This migration phenomenon gives me the best testing laboratory possible. This is because, when moving across states, workers do not face policy constraints. This is in contrast with most international migration where skill-based admission may contaminate the underlying pattern of selection.[5] Moreover, interstate migrants stay within the scope of the Survey of Income and Program Participation, from where I get the data for my analysis. This avoids, or at least limits, representativeness concerns that are, otherwise, an issue on international migration (Ibarraran and Lubotsky 2007) and allows me to observe characteristics and earnings of movers before they move. Finally, earnings differentials across states are significant and long-lasting (Dahl 2002). Thus, there is scope for the selection mechanism described by Borjas' model taking place.

Using the Survey of Income and Program Participation and the Current Population Survey I classify states according to their earnings dispersion and characterize immigrant selection using pre-migration earnings. The empirical evidence I provide matches quite well the selection pattern described by the model. Selection increases with earnings dispersion at the destination location. Furthermore, the earnings distribution of those moving to the least disperse location never dominates the earnings distribution of other migrants and the earnings distribution of those moving to the most disperse locations is never dominated. This pattern holds for observed, predicted and residual pre-migration earnings across sex and age groups.

Alongside migrant selection, a large branch within migration economics asks what are the consequences of immigration for receiving economies.[6] Earlier literature focused on the effects in the receiving labour market (e.g. Borjas and Tienda 1987; Card 1990; Grossman 1982; Hunt 1992; Johnson 1980; LaLonde and Topel 1991).[7] Specifically, the earlier literature focused on the effect of immigrants on natives' wages (see also Altonji and Card 1991; Borjas, Freeman, et al. 1992; Butcher and Card 1991).

Since the early 1990s, the literature has spread in terms of outcomes studied. For example, there are recent studies on the effect of immigrants on crime (Bell et al. 2013; Mastrobuoni and Pinotti 2015), attitudes towards migration (Dustmann and Preston 2007; Viskanic 2017), public services (Preston 2014; Wadsworth 2013), real-estate (Sá 2014) and education (Hunt 2017; Machin and Murphy 2017). Nonetheless, the effects of immigration on the receiving labour market are, still, the central issue (e.g. Dustmann, Fabbri, et al. 2005; Dustmann, Frattini, et al. 2013; Dustmann,

[5]There is some evidence that immigration policy has been shifting towards promoting high-skilled immigration (see Boucher and Cerna 2014, and other papers in the same special issue.).

[6]Spengler (1958) may be one of the earliest examples of this literature.

[7]An, earlier, exception is Reder (1963). He discusses the effects of post-war immigration on investment, output per capita, labour supply and income distribution.

Schönberg, et al. 2017; Hatton and Tani 2005; Manacorda et al. 2012; Ottaviano and Peri 2012).[8]

In the third and fourth chapters of this thesis, I contribute to this literature by estimating the effect of immigrants in two, relatively, unexplored outcomes. In my third chapter, I investigate the effect of immigrants on labour productivity in the UK. Immigration and productivity are growing concerns for British policymakers (see ministerial white papers BEIS 2017; Home Office 2018). However, there is little evidence of the effect of immigration on productivity. Aleksynska and Tritah (2009), Boubtane et al. (2016), and Ortega and Peri (2014) provide cross-country evidence showing that immigration has a positive effect on labour productivity. A similar conclusion is reached by Peri (2012) using data from the US and by Ottaviano, Peri, and Wright (2018) using service sector data for the UK. However, Paserman (2013) finds negative effects on Israeli manufacturing firms and Kangasniemi et al. (2012) provide mixed evidence for Spain and the UK.

My third chapter contributes to the literature by estimating the average effect of immigrants on labour productivity in the UK. Using variation on past-settlement of immigrants (Altonji and Card 1991; Card 2001) across regions and industries, I estimate the effect of altering the relative supply of immigrant labour on labour productivity. My estimates show that increasing the relative supply of immigrant labour by ten percentage points rises output per worker by 6.7%. Evaluated at the average, this is a £5,610 increase in annual output per worker.

To better understand the reduced-form estimate, I introduce a Constant Elasticity of Substitution production function which I use to decompose the reduced form estimate. I show that the reduced-form effect is composed of differences in marginal returns between immigrant and native labour and the changes immigrants induce on other inputs. Thus, a key insight from my decomposition is that reduced-form estimates of the effect of immigration on output and labour productivity depends on the mix of inputs. Therefore, it warns against extrapolating estimates across time, country, sectors or levels of aggregation as the input mix varies across these dimensions.[9]

To gauge the importance of effects on other inputs on the reduced form effect, I estimate the effect of immigration on capital stocks, native labour, and native skill mix. My estimates show that immigrants have a positive effect on capital stocks, native labour and native skill mix. However, the positive effect of immigration on labour productivity goes beyond effects on these inputs. Controlling for capital and native skill mix, I estimate that immigrants have a positive and significant effect on labour productivity. Increasing the relative supply of immigrant labour by ten percentage points increases labour productivity by 5%.

Given the persistence of the positive effect of immigration on labour productivity,

[8] See also chapters 3 to 6 in Borjas (2014) and the review by Card and Peri (2016).

[9] See figure 3.1, table 3.1 and evidence provided by Dustmann, Schönberg, et al. (2016).

I explore whether it works through differences on relative marginal returns or if immigrants may trigger the development of technologies that complement them. Using my decomposition of the reduced form effect I show that, for a wide range of parameter values, the evidence I provide is consistent with immigrants encouraging the development of technologies that complement them. This positive effect is consistent with Acemoglu (1998, 2002) model of endogenous technical change. In Acemoglu's model, as long as immigrant and native labour are gross substitutes, increasing the relative supply of immigrant labour increases the supply of technologies that complement immigrants. To provide further evidence about whether this mechanism is at place, I compare estimates from two production function estimators with the reduced form effect controlling for other inputs. Assuming that factor augmenting technologies are constant, my production function estimate produces an immigrant marginal effect that is only positive when I allow for effects through capital stocks. This is in contradiction with the reduced form effect and further suggest that immigrants have an effect on technologies that complement them. Moreover, when I compare this production function estimate with a polynomial approximation, that is consistent with endogenous technical change, I find that the approximation produces a larger and positive marginal effect. This suggests, again, that immigrants trigger the development of migrant labour augmenting technologies.

My work on the effect of immigration on labour productivity, therefore, shows that changing the labour mix produces effects that go beyond differences in marginal products and into effects through other inputs and, possibly, technologies. This fits well within the existing literature. Clemens et al. (2018) produce suggestive evidence pointing to endogenous technological change as a result of immigrant restrictions. Hornbeck and Naidu (2014) show that reductions in the availability of cheap labour lead to the introduction of technology to substitute it. Moreover, Hornung (2014) and Hunt and Gauthier-Loiselle (2010) show that immigrants can boost the diffusion and development of innovations. Finally, Hunt (2017), Llull (2017), and Ransom and Winters (2016) show that immigrants shape natives' educational decisions.

The fourth chapter of this thesis goes, precisely, into that last effect. Do immigrants change educational outcomes of natives? However, differently from Hunt (2017), Llull (2017), and Ransom and Winters (2016) that study the effect of changes on the aggregate stock of immigrants; Greta Morando and I ask what is the effect of sharing university and major with foreign peers on educational and early labour market outcomes of native students. Our research is motivated by an ever-growing number of international students in higher education that makes the effects of international students in the receiving country and, particularly, on native students key for policy (see MAC 2018). This is because higher education is a key stage for the future labour market performance of workers. At this stage, the universities and colleges that individuals attend (Arcidiacono 2004; Belfield et al. 2018; Walker and Zhu 2008, 2018) and the educational outcomes they achieve (Feng and Graetz 2017; Jaeger and Page 1996; Jones and Jackson 1990) have an impact on future labour market outcomes.

There is substantial literature on the effect of peers on individuals' outcomes.[10] However, the literature on foreign peer effects is still small and has focused on lower levels of education. Ballatore et al. (2018) find negative foreign peer effects on language and math performance of second graders.[11] In contrast with this, Geay et al. (2013) show that, in England, the negative effect of foreign peers on natives is because of selection and that foreign peers produce a zero effect. Moreover, in their study of foreign peers on Israeli natives, Gould et al. (2009) find an overall no effect that, however, hides a negative effect for disadvantaged students. Finally, Ohinata and Van Ours (2013) study the effect of foreign peers on Dutch students finding only mild effects.

At higher education, Anelli, Shih, et al. (2017) study the effect of foreign peers on the probability of graduating in a science, technology, engineering and mathematics (STEM) major at a university in California. Their evidence shows that foreign peers have a negative effect on the probability of graduating in a STEM major. However, this has no negative effect on natives' earnings as immigrant peers displace native students into high earning social science majors. Also in the US, but using aggregate data, Orrenius and Zavodny (2015) provide evidence showing that increasing the share of immigrants in a cohort has a negative effect on the probability of STEM graduation for women. Finally, Chevalier et al. (2019) study the impact of linguistic diversity at a British university finding no effects on English-speaking students.

We add to the literature by estimating foreign peer effects in higher education on native educational and labour market outcomes for the whole of England. The main identification problem that we face is that certain institutions may attract particular native students and also attract foreign students (Ohinata and Van Ours 2013). We deal with self-selection into university and majors by using variation across cohorts within university and major. However, we show that this variation alone is not good enough to estimate the effect of foreign peers. This is because universities have ex-ante information about the quality of prospective students. Then, under capacity constraints, an increase in the number of prospective students, as given by a larger number of foreigners, increases the average ability of accepted students. We show that this is the case for EU students who share with natives the same cap on the number of subsidized students allowed to be enrolled by each university.[12] This produces that rising the share of EU students increases average ability of enrolled natives. However, non-EU students have no significant effects on average native ability. The same asymmetry holds for group size. Increasing the share of EU students does not produce any change in the size of the group, but increasing the share of non-EU increases group size.

[10] See Sacerdote (2011) review of the literature on education peer effects.

[11] Brunello and Rocco (2013) also provide evidence of negative effects but at the aggregate level.

[12] This was imposed by the Higher Education Funding Council for England (HEFCE) (see Machin and Murphy 2017).

Thanks to the quality of our data we can overcome this initial selection problem. This is because we observe a measure of natives' ability, their Universities and Colleges Admissions Service (UCAS) score. We show that, after conditioning on ability and group size, variation across cohorts within university and major is uncorrelated with a large set of pre-determined individual characteristics that are likely predictors of educational achievement.

The evidence we provide shows that there are only mild foreign peer effects and that these are heterogeneous across the ability distribution. The strongest foreign peer effects that we find are on the distribution of grades, in particular, on the probability of graduating with an upper- or lower-second for top ability students. We show that these effects on grades are not consistent with a grading-on-a-curve mechanism where foreign peers mechanically displace native students. Foreign peers, therefore, modify the human capital accumulation of native students. However, these effects are very mild. For example, the largest effect we find shows that a one percentage point increase on the foreign peer share reduces the probability of graduating with an upper second for top ability natives by .30 percentage points.

In terms of early labour market outcomes, we also find mild and typically statistically non-significant foreign peer effects. If anything, foreign peers increase the probability of natives working after graduation by reducing the probability of them going into further education; increase the probability of working in professional occupations and have a positive impact on wages of top ability natives.

Our research, therefore, is consistent with existing evidence providing mild foreign peer effects (Anelli, Shih, et al. 2017; Chevalier et al. 2019). However, it also suggests that foreign students may have an effect on the universities and majors that natives end up attending. This can, potentially, have a greater impact than foreign peer effects.

Chapter 2

Selection by Destination: Revisiting Stochastic Dominance in the Borjas Model

Abstract

Since Borjas (1987) seminal paper, the Roy model has been widely used to study immigrant self-selection. Particularly, the baseline two-location model has been a popular frame for self-selection research, even in situations where immigrants may face a larger choice set. I present an analysis of Borjas' model of selection extended to multiple locations. This extension shows that earnings dispersion alone determines selection only for the most and least disperse locations. Therefore, it highlights that, what determines selection, is the ranking of locations rather than the relative dispersion of earnings between home and destination. Using US internal migration, I provide evidence supporting the implications of the model.

2.1 Introduction

A long-standing empirical observation is that migrants are not randomly drawn from their population, they are selected in terms of their skills.

Seminal work in Borjas (1987) uses Roy (1951) occupational choice framework to explore this phenomenon. In the baseline model, individuals choose between two locations and decide where to work based on the returns to their skills and moving costs creating a selection mechanism that is entirely characterized by relative earnings dispersion or skill premium. Higher (lower) earnings dispersion at destination leads to positive (negative) selection. Thus, the selection mechanism behind this pattern can be thought of as low skilled individuals taking insurance from the more compressed distribution at the destination and high skilled taking advantage of improved opportunities in a more disperse location (Parey et al. 2017).

Recently, migrant selection has received renewed interest and several papers, most notably on Mexico-US migration, have tested the implications of the Borjas (1987) framework. Chiquiar and Hanson (2005) find medium to positive selection of Mexican migrants to the US, what is evidence against Borjas selection model as Mexico is a more unequal country than the US. However, later revision (Ambrosini and Peri 2012; Fernandez-Huertas Moraga 2011, 2013; Ibarraran and Lubotsky 2007; Kaestner and Malamud 2014) provide evidence supporting the implications of Borjas' model. Outside Mexico-US, Parey et al. (2017) use data on German graduates and provide evidence supporting the implications of the model. Borjas, Kauppinen, et al. (2018) provide further evidence in support of the Borjas using Danish data.

Several papers in the literature (e.g. Belot and Hatton 2012; Feliciano 2005; Fernandez-Huertas Moraga 2013; Parey et al. 2017) use the two-location model when studying selectivity of immigrants across multiple locations. A common interpretation of the main implication of the Roy model is that *"if the home country's income distribution is more unequal than in the ...[destination], immigrants will be negatively selected..."* (Feliciano 2005, p. 133). This interpretation guides, in some way, most of the analysis framed within the Borjas model. For example, Belot and Hatton (2012) regress measures of educational selectivity on differences in relative returns to education between home and destination location. My extension of the Borjas model including more than two locations highlights that selection is not driven by relative earnings dispersion between home and destination but by the ranking, in terms of earnings dispersion, of locations within the choice set. With more than two locations and unrestricted transferability of skills, my extension of the Borjas model predicts positive (negative) immigrant selection into the most (least) disperse location.

In the two-location model, the ranking and home-destination relative dispersion are equivalent and leads to the interpretation given in the literature. However, analysis of selection within the Borjas model looking at home-destination relative dispersion implicitly assumes that immigrants only consider these two locations. Although that could be the case in some migration phenomena, for example, Mexico-US, it can be restrictive for other where individuals are likely to decide among a wider range of alternatives. For example, US interstate migration (Kennan and Walker 2011) or emigration across European countries (Rojas-Romagosa and Bollen 2018).

Using US inter-state migration, I test whether migrant selection follows the pattern implied by my extension of the Borjas model. This migration phenomenon gives a perfect scenario to test the implications of the Borjas model as there are no border constraints that could contaminate the pattern of selection and individuals can be tracked even if they switch locations. This reduces measurement error, for example, due to illegal migration (Ibarraran and Lubotsky 2007), and allows observation of pre-migration characteristics. Furthermore, when engaging on internal migration, US workers face a large choice set, as highlighted by Kennan and Walker (2011), that matches well the set-up of my extension of the Borjas model. Finally, US local labour markets exhibit heterogeneity in their earnings distribution and such differences are long-lasting (Dahl 2002). Therefore, the selection drive presented in Borjas (1987) is at place.

For testing, I use data from the Survey of Income and Program Participation and Current Population Survey covering the period 1984 to 2013. Following Borjas, Kauppinen, et al. (2018), I cast the implications of the model in terms of stochastic dominance, which provides a stronger test than selection on means. In line with the model's predictions, I show that immigrants moving to the most (least) disperse location are positively (negatively) selected in terms of pre-migration earnings and immigrants to other locations show, somehow, intermediate selection. I show, also in line with the model, that the degree of selection increases with dispersion at the home location. Thus, I contribute to the literature by showing that a multiple location Roy model without imposing perfect transferability of skills (see Borjas, Bronars, et al. 1992) holds meaningful implications, at least for the least and most disperse locations, and that the implied pattern of selection is observed in the data.

2.2 Theoretical Framework

Seminal work in Borjas (1987) departs from the framework introduced in Roy (1951) and develops a two location model where individuals chose freely to migrate depending on earnings at origin and destination. The central implication of the baseline Borjas model is that, given appropriate assumptions about the relationship between migration costs and skill, selection is completely characterized by the relative dispersion of earnings.[1] With sufficient skill transferability, higher earnings dispersion at destination leads to positive selection and the opposite to negative.

[1]The baseline Borjas model assumes that migration cost and earnings are orthogonal so costs don't determine selection. Kaestner and Malamud (2014) provide some empirical evidence on migration costs and selection. Borjas (1991) and Chiquiar and Hanson (2005) introduce extensions with non-constant costs. In equation (2.1), one can let the amenity value be a random variable, if it is independent of skill then it won't change the implications of the model. Allowing for dependence between skill and amenities, one can produce arbitrary patterns of selection.

Recent work in Borjas, Kauppinen, et al. (2018) shows that the baseline Borjas model has not only implications for selection on mean skill but also for stochastic dominance. However, they keep the choice set constrained to two locations, origin and destination. Although this can be an adequate reduction for some migration events, e.g. Mexico-US, it seems too constrictive for settings with multiple competing destinations. Here, I explore the model when one allows for an arbitrary number of locations and provide conditions under which there is stochastic dominance.

To present the model, assume that individuals' utility from working in location l depends on local amenities μ_l and wages w_l

$$U_{il} = \mu_l + w_{il} \tag{2.1}$$

where for simplicity I assume that local amenities are deterministic and, critically, I impose a multivariate normal distribution on the vector of wages $\mathbf{w} \equiv [w_l]$.

Individuals are utility maximizers and choose to locate in the labour market holding the highest utility, leading to the following policy rule for location l

$$M_{il} \equiv \mathbb{1}[U_{il} - U_{ik} > 0 \ \forall \ k \neq l] \tag{2.2}$$

To characterize selection, let z be a measure of worker skill in the sense that the covariance vector $Cov(\mathbf{w}, z)$ has all strictly positive elements.[2] As in Borjas, Kauppinen, et al. (2018), I am interested in the implications of the model for stochastic dominance of migrants' skill. Thus, I want to derive under which conditions

$$\begin{gathered} F(z|M_l = 1) \leq F(z) \ \forall \ z \\ \text{with} \\ F(z|M_l = 1) < F(z) \text{ for some } z \end{gathered} \tag{2.3}$$

Thistle (1993) shows that a necessary and sufficient condition for stochastic dominance is that the negative moments are strictly ordered. Therefore, the stochastic dominance relation stated in (2.3) will hold if

$$m_{z|M_l=1}(-t) < m_z(-t) \ \forall \ t > 0 \tag{2.4}$$

where $m(.)$ is the moment generating function. Let me define two new vectors, one holding standardized wage differences

$$\mathbf{w}^{-l} = \left[\frac{w_l - w_1}{Var(w_l - w_1)}, \ldots, \frac{w_l - w_{l-1}}{Var(w_l - w_{l-1})}, \frac{w_l - w_{l+1}}{Var(w_l - w_{l+1})}, \ldots, \frac{w_l - w_L}{Var(w_l - w_L)}\right]$$

and the other holding amenity differentials standardized with the variance of the wage differences

[2] In the literature, there are multiple examples of skill measures. For example, Fernandez-Huertas Moraga (2011) uses pre-migration earnings, Parey et al. (2017) use predicted earnings and Belot and Hatton (2012) use education.

$$\boldsymbol{\Delta\mu}^{-l} = \left[\frac{\mu_l - \mu_1}{Var(w_l - w_1)}, \ldots, \frac{\mu_l - \mu_{l-1}}{Var(w_l - w_{l-1})}, \frac{\mu_l - \mu_{l+1}}{Var(w_l - w_{l+1})}, \ldots, \frac{\mu_l - \mu_L}{Var(w_l - w_L)}\right]$$

Then, an individual is observed in location l if the $L-1$ inequalities, $\mathbf{w}^{-l} > \Delta\boldsymbol{\mu}^{-l}$, are satisfied. This implies that one is interested in the distribution of $(z, \mathbf{w}^{-l})$ when $\mathbf{w}^{-l}$ is truncated from below. Given normality of the random variable $(z, \mathbf{w}^{-l})$, the moment generating function of the marginal truncated distribution can be readily obtained from results in Tallis (1961)[3]

$$m_{z|M_l=1}(t) = e^{\alpha t + \sigma^2 t^2/2} \frac{\Phi_{L-1}(\Delta\boldsymbol{\mu}^{-l} - \boldsymbol{\rho}^{-l} t)}{\Phi_{L-1}(\Delta\boldsymbol{\mu}^{-l})} \tag{2.5}$$

Where Φ_{L-1} is the mass under the right hand tail of a $L-1$ dimensional normal density, α and σ^2 are the mean and variance of z and $\boldsymbol{\rho}^{-l}$ is the vector of correlations between z and $\mathbf{w}^{-l}$. When these correlations are all zero the moment generating function in (2.5) reduces to that of a standard normal and with two locations it is the moment generating function provided in Arnold et al. (1993) and used in Borjas, Kauppinen, et al. (2018). It follows from (2.4) and (2.5) that the truncated distribution dominates the population if

$$\Phi_{L-1}(\Delta\boldsymbol{\mu}^{-l} + \boldsymbol{\rho}^{-l} t) < \Phi_{L-1}(\Delta\boldsymbol{\mu}^{-l}) \tag{2.6}$$

Thus, when all elements of ρ^{-l} are positive (negative) the truncated distribution dominates (is dominated) by the population independently of differences in local amenities. In turn, the sign of these correlations is determined by the covariance vector of the skill measure and wages that has k^{th} element

$$\mathrm{Cov}(z, w_l - w_k) = \sigma\rho_{zk}\sigma_k \left(\frac{\rho_{zl}\sigma_l}{\rho_{zk}\sigma_k} - 1\right) \tag{2.7}$$

Where ρ_{zl} is the correlation between the skill measure and wages in location l and σ_l is the standard deviation of wages in location l.[4] When skill transferability is homogeneous across locations ($\rho_{zl} = \rho \ \forall l$) the sign of the correlations in $\boldsymbol{\rho}^{-l}$ is determined by pair-wise comparison of skill prices across locations. Under these conditions, equations (2.6) and (2.7) imply that the skill distribution of those choosing the most (least) disperse location dominates (is dominated by) the population skill distribution. Thus, the implications for selection are analogous to Borjas (1987) canonical with a two location set-up, where the Borjas model implies positive selection into the market with the largest wage dispersion. However, working with more than two locations highlights one insight: it is not the ratio between the returns at destination with respect to home what determine selection, but the ranking of the destination

[3]Tallis (1961) provides the moment generating function of an L-dimensional normal, one needs to allow one of the truncation points go to minus infinite to get (2.5).

[4]Differences on wages standard deviations can also be interpreted as differences on skill prices for some unit variance normally distributed skill.

within the choice set. To see this, let the skill measure be the wage obtained at the home location, $z = w_h$. Then

$$\text{Cov}(w_h, w_l - w_k) = \sigma_h \rho_{hk} \sigma_k \left(\frac{\rho_{hl}\sigma_l}{\rho_{hk}\sigma_k} - 1 \right) \tag{2.8}$$

It follows from (2.8), that the home location plays a role on the direction of selection if there is heterogeneous transferability of skill across locations but not through differences in dispersion or returns to skill, i.e. through σ_h. Suppose that all locations are equally disperse, then the k^{th} element of the covariance vector is

$$\text{Cov}(w_h, w_l - w_k) = \sigma^2 \rho_{hk} \left(\frac{\rho_{hl}}{\rho_{hk}} - 1 \right) \tag{2.9}$$

and all $L - 1$ covariances in $\boldsymbol{\rho}^{-l}$ will be positive (negative) if the chosen location l is the one with the highest (lowest) transferability of skills. This is intuitive. If a worker has a good initial wage draw, she has an incentive to move to the location that preserves it. On the other hand, if a worker has a first wage that is low, she has an incentive to move into a location where the previous wage does not strongly determine the current one, so she can get a fresh start.[5]

If skill transferability is assumed to be homogeneous across locations (e.g. Borjas, Bronars, et al. 1992), then wage dispersion at the home location plays no special role in determining the direction of selection. Under homogeneous skill transferability the k^{th} element of the covariance vector is

$$\text{Cov}(w_h, w_l - w_k) = \sigma_h \rho \sigma_k \left(\frac{\sigma_l}{\sigma_k} - 1 \right) \tag{2.10}$$

Now the $L-1$ covariances are all positive (negative) if the chosen location, l, is the one with the highest (lowest) wage dispersion. It is clear from equation (2.10) that wage dispersion at the home location, σ_h, plays no special role in determining the direction of selection. Nonetheless, a higher wage dispersion at the home location will increase the degree of selection as it enlarges the elements of $\boldsymbol{\rho}^{-l}$. However, this effect of home earnings dispersion is given by me measuring skills with earnings at the home location. If instead I use an homoscedastic measure of skill, as in (2.7), the effect of home earnings dispersion disappears. In equation (2.11) what increases the degree of selection is the dispersion of the skill measure. This is intuitive. If all workers are the same, i.e. $\sigma = 0$, there can be no-selection. Moreover, the most heterogeneous workers are, i.e. the larger σ is, the most intensive selection can be. Thus the effect of home earnings dispersion in (2.10) is just an artefact of how I measure selection and not a result of the selection mechanism.

$$\text{Cov}(z, w_l - w_k) = \sigma \rho \sigma_k \left(\frac{\sigma_l}{\sigma_k} - 1 \right) \tag{2.11}$$

[5] In this discussion I rule out strange cases where $\rho_{hl} < 0$.

2.3 Data

To test the implications of the theory, I use data from the Survey of Income and Program participation (SIPP)and IPUMS-CPS (Flood et al. 2015) covering the period 1984-2013.[6] The SIPP is a sample of partially overlapping panels, all of them designed to be representative of the whole US. The duration of these panels goes from two to four years, although some panels were discontinued before reaching the two year mark. In my analysis, the main advantage of using the SIPP is that it keeps internal movers within scope, allowing me to observe people moving and record characteristics of movers before they move. This is critical for me as I use pre-migration measures to test migrant selection. In addition, the SIPP covers multiple cohorts and a long window of time, therefore, I also add to previous results in the literature (Borjas, Bronars, et al. 1992) using a particular cohort by showing whether the implications of the theory hold across multiple cohorts and periods of time.

Although there is a panel for each year between 1984-2013, I do not use data from panels 1984 and 1989. This is because I select observations into my estimation data conditioning on being US-born and not being enrolled in education. In panel 1989, there is no available information about migration history including state or country of birth and panel 1984 does not provide information about current school enrolment.[7]

From the original sample, I select observations for individuals who gave a valid interview, proxy included, are 18 to 64 years old, not self-employed, not enrolled in full-time education, have never been in the army and were born in the US. When I observe an individual enrolled in further education, I keep later observations if she leaves further education. I also eliminate from my estimation sample individuals that are living, at some point, in a non-individually identifiable state. These are Alaska, Idaho, Iowa, Maine, Montana, North Dakota, South Dakota, Vermont, and Wyoming.[8] Although I take these states into account when ranking locations. This is possible because I use CPS to produce location-specific earnings dispersion estimates.

When I merge individual characteristics with their migration histories, some observations are not matched.[9] These unmatched observations are due to individuals that were not present at the designated sample wave when migration histories were recorded. This can be because they left the survey, did not enter it yet or could not be contacted at that particular wave.[10] I disregard all observations without matched migration histories as these provide information about birth place.[11]

The resulting sample contains 264,813 individuals with an average of 29 observations per individual and a total of 13,339 interstate mobility events from which I

[6] I have used SIPP core and topical files, extraction programs and data dictionaries from NBER http://www.nber.org/data/survey-of-income-and-program-participation-sipp-data.html

[7] Panel 1989 was discontinued after three waves and topical modules were never released.

[8] I display state groups in table 2.6.

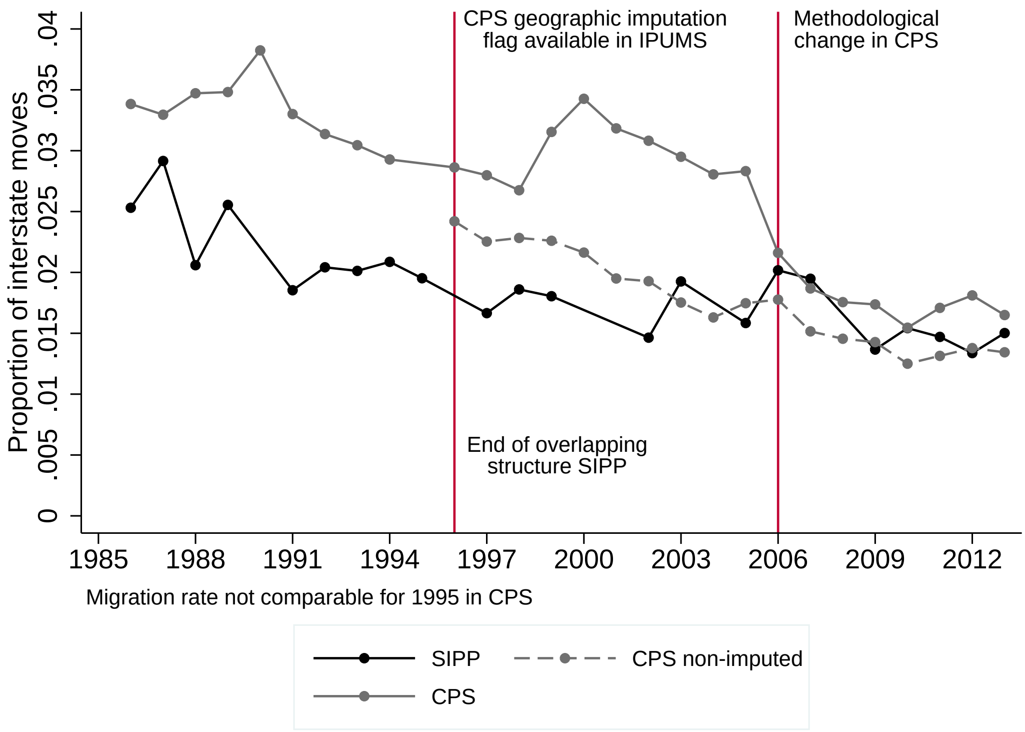
[9] For all panels but 1985 migration history questions were carried at wave 2. For panel 1985 this information was gathered at wave 4.

[10] For further information see Westat and Mathematica Policy Research (2001) chapter 13.

[11] How birthplace information is recorded varies across waves, see appendix 2.6.2. And the same is true for other variables, see appendix 2.6.2 for the homogenization I use.

eliminate 1,732 of them because I do not observe any pre-migration earnings and 2,390 that correspond to individuals moving to the State where they were born. To check the representativeness of the SIPP data, in figure 2.1, I display one-year migration rates drawn from SIPP and CPS.[12] Interstate migration rates have been decreasing for the last 30 years and this is apparent in both SIPP and CPS. However, the CPS data displays a sharp drop in year 2006 when imputed observations are not disregarded. Kaplan and Schulhofer-Wohl (2012) investigate the causes of such a drop and conclude that it can be mostly attributed to a methodological change on the missing value imputation procedure. When I drop imputed values, the data displays a much smooth decreasing behaviour, comparable with the one observed in SIPP.[13]

Figure 2.1: One-year interstate migration rate



2.3.1 Migrant Definition

I define as a mover those individuals for whom current state of residence differs from the previous one. The use of States to define alternatives in internal US migration is not new to the literature (e.g. Borjas, Bronars, et al. 1992; Dahl 2002; Kennan and Walker 2011) and it gives me an exhaustive and mutually exclusive division of mainland US that is stable across time. Furthermore, states are typically large enough that local labour markets do not cross state borders and the data shows that is across

[12]Construction of moving rates from SIPP require some work. See appendix 2.6.1 for an explanation of the procedure.

[13]IPUMS provides with an imputation flag from 1995 onwards. Before 1996 imputation rates were marginal (Kaplan and Schulhofer-Wohl 2017).

states and not at lower levels where labour motives dominate migration events (table 2.10).[14] Importantly, the use of States to delimit mobility events makes my results comparable with previous previous work on internal US migration selection (Borjas, Bronars, et al. 1992) and allows for easier replication as different datasets provide publicly available state identifiers but no smaller geographical areas (see Molloy et al. 2011).[15]

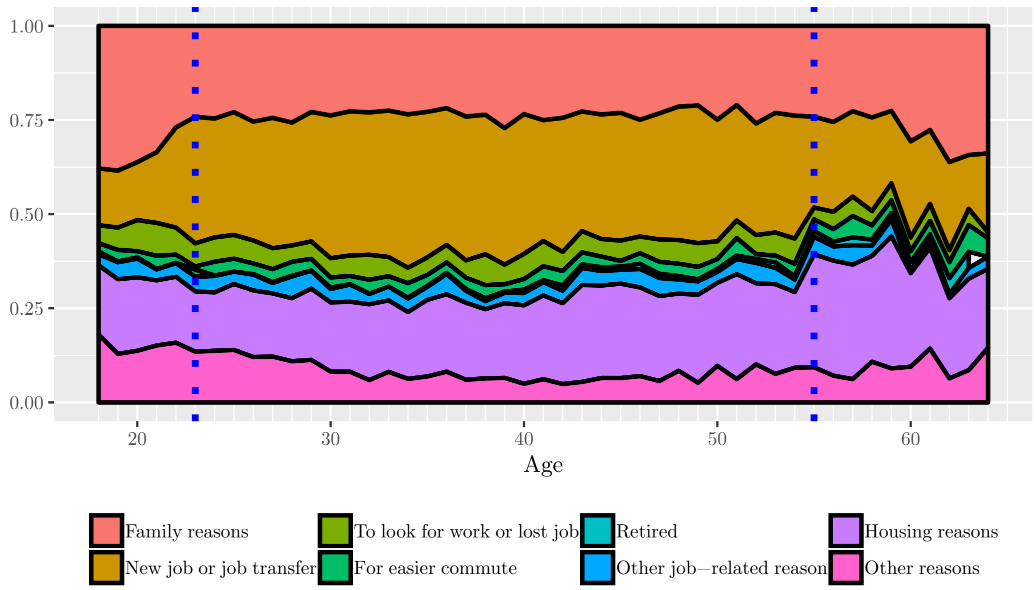
Within those that move across states, there is a marked age profile in terms of mobility motives (figure 2.2). Those movers that are 25-55 years old tend to report moving across states due to new job or job transfer. For older movers the dominant reason is housing reasons, or family reasons if above 60, and for younger movers the dominant motive is family reasons with a clear decreasing profile between age 16 and 25. In the analysis, I take this into account and produce estimates with all 16-65 working individuals and 25-55 only.

That few individuals declare having moved to look for work or due to job lost suggest that a large proportion of those who move due to job related reasons have engaged on non-local job search and, therefore, it is feasible to obtain ex-ante wage information at other locations. This is supporting evidence for a minimal condition for selection in the Borjas model: movers must know the distribution of wages at other locations.

[14]Nonetheless, states may cover multiple labour markets and some interstate migration events may not involve labour market changes (Kaplan and Schulhofer-Wohl 2017; Molloy et al. 2011).

[15]Other definitions of locations in the US migration literature include Standard Metropolitan Statistical Area (e.g. Bishop 2012; Gabriel and Schmitz 1995) and distance based measures (Ham et al. 2011).

Figure 2.2: Moving Reason Age Profile



Note: Pooled 1984-2013 CPS data, excluding 1995. In CPS an individual is defined as a migrant if one year ago she was living in a different state, with the exception of year 1995 where the time interval refers to five years prior to the interview (see Faber 2000). Those that move across states only.

Finally, what is more convoluted to define is the population against which one will compare the characteristics of movers. The usual procedure in the literature (e.g. Borjas, Bronars, et al. 1992; Fernandez-Huertas Moraga 2011; Parey et al. 2017; Spitzer and Zimran 2018) is to take a window of time, often times defined by the sample window, on which if an individual is observed moving she is classified as a migrant and otherwise as a stayer. Although, this is an undeniably interesting exercise, it does not quite match the layout of the model. Those that are not observed moving do not represent the population form which migrants are drawn and can be potentially much different if selection is strong. Actually, recovering the distribution of skills in the population can be a daunting challenge when there are many possible destinations, even with a highly parametrized framework as the Borjas model.[16] Instead of trying to recover the population distribution or using the population at risk of migration (as referred by Spitzer and Zimran 2018), I compare skill distributions of movers across destinations. This has the advantage of distributions been directly recoverable from the data and matches perfectly the layout of the model. In section 2.2, I show that the distribution of skills for movers to the least disperse location is dominated by the population while the distribution of those moving to the most disperse location dominates the population. Then, to test the implications of the

[16] Dahl (2002) provides a semi-parametric estimator to recover the population first moments that he uses to estimate differences in returns to education across US states in a multiple market Roy model.

model, I can just compare the distribution of skills for those moving to the least disperse location with the distribution of those moving to more disperse locations and, if the model gives a sufficiently good description of the selection mechanism, I should observe that selection turns positive as I increase dispersion at destination.

2.3.2 Measuring Skill and Dispersion

I follow Fernandez-Huertas Moraga (2011) and use pre-migration earnings measured at the last observation before migration as my measure of migrant skill. From pre-migration earnings I derive three measures of skill: predicted earnings, pre-migration standardized residuals and the aggregation of the two. To compute predicted earnings and residuals, similarly to Borjas, Bronars, et al. (1992), I use the following wage equation

$$w_{itj} = \boldsymbol{\beta}\mathbf{x_{it}} + \gamma_t + \omega_j + \epsilon_{itj} \tag{2.12}$$

Where $\mathbf{x}$ contains a constant, a quadratic on age and dummies for sex, race, highest education level attained, industry, and occupation. δ and γ are year and state dummies. I control for occupation and industry because individuals specialize within these, as one can observe in tables 2.1 and 2.11 where movers tend to find a job within the same industry or occupation as their last pre-migration job.

Table 2.1: Occupation Transitions of Movers
(*Row Proportions*)

	Current Occupation			
Previous Occupation	Manager	High/Medium Skilled	Low Skilled	Farmer
Manager	0.914	0.069	0.016	0.000
High/Medium Skilled	0.044	0.921	0.033	0.002
Low Skilled	0.028	0.089	0.876	0.007
Farmer	0.050	0.067	0.117	0.767

A possible concern is that the OLS estimates of returns to characteristics from an equation like (2.12) may be biased due to immigrant selection. However, estimators in the literature correcting for migrant selection (Dahl 2002; Parey et al. 2017) return estimates that are close to OLS. Thus I use OLS estimates from specification (2.12) to obtain a measure of predicted, $\boldsymbol{\beta}\mathbf{x}$, and residual, ϵ, pre-migration earnings. In the case of residual earnings, I follow Borjas, Bronars, et al. (1992) and divide them by the earnings variance of the state where the earnings were generated to account for differences in variance that are given by differences in unobservable skill pay-off across locations. Then, using the standardized measure of residual earnings (ϵ^*), I construct standardized pre-migration earnings as $w^* = \boldsymbol{\beta}\mathbf{x} + \epsilon^*$, this is my third measure of skill.

Another concern when using earnings to measure selection is whether they may be modified by or in anticipation of migration (Fernandez-Huertas Moraga 2011; Spitzer and Zimran 2018). Fernandez-Huertas Moraga (2011) and McKenzie et al. (2010) test whether there are earnings drops right before migration, similar to those in the program participation literature (Ashenfelter 1978), without finding any significant changes. In unreported results, I tested whether the difference between the earliest observation of earnings and the last observation before migration are significant. I find non statistically significant differences after controlling for experience and calendar year effects.

To rank locations I follow the model tightly and use earnings dispersion. I produce dispersion estimates using data from IPUMS-CPS because in the CPS all states are individually identified, while in the SIPP some states are grouped together. This implies that only with CPS data I can effectively rank all states according to their earnings dispersion. From the CPS I compute within year and state earnings dispersion and residual-earnings dispersion from a wage equation controlling for sex, ethnicity and education. Whether I choose to use earnings or residual-earnings to measure dispersion does not make much difference in terms of empirical results. This is expected given the strong correlation between the two measures. Nonetheless, what makes a difference is the temporal aggregation. Taking the average of the yearly earnings dispersion estimates for each state and regressing raw pre-migration wages on the dispersion ranking produces a strong positive estimate. Moving up one position in the ranking of the destination location correlates with a .4% increase on pre-migration earnings. However, the effect reduces to .1% if I rank locations according to yearly dispersion estimates instead of averages of yearly estimates (table 2.2). Measurement error may explain why I see smaller estimates when I rank location by year dispersion instead by decade or whole sample averages. This is because averaging observations across time may reduce the extend of measurement error and, therefore, the impact of attenuation bias.

The selection effect on wages is robust to controls for sex, ethnic group and age. Although it reduces to .2% when I introduce controls for whether the individual has no education, has completed up to 4^{th} grade, has a high school diploma or is a degree graduate. Additionally, controlling for industry and occupation reduces the selection effect further but only when ranking locations according to yearly dispersion or decade averages.

Table 2.2: Rank of Destination and Pre-Migration Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
Rank (fixed)	.004***	.004***	.004***	.004***	.002***	.002***
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Rank (decade)	.003***	.003***	.003***	.003***	.002***	.001**
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Rank (year)	.001*	.001**	.001**	.001**	.0001	−.0001
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Observations	9,217	9,217	9,217	9,217	9,217	9,217
			Controls			
Sex	N	Y	Y	Y	Y	Y
Ethnicity	N	N	Y	Y	Y	Y
Age	N	N	N	Y	Y	Y
Education	N	N	N	N	Y	Y
Industry	N	N	N	N	N	Y
Occupation	N	N	N	N	N	Y

Note: Each estimate in the table comes from an independent regression. Clustered standard errors by previous state in paranthesis. All columns include fixed effects for state of origin and year. Characteristics and earnings measured at last observation before migration. $p < .1$ $^{\cdot}$, $p < .05$ *, $p < .01$ **, $p < .001$ ***

Changes in the magnitude of selection effects estimates from column (4) to (5) in table 2.2 suggest that selection on observables might be stronger in terms of education, with mild or no selection in terms of sex, age and ethnicity. That is exactly what I observe when looking at those characteristics across location groups delimited by earnings dispersion (table 2.12). Individuals with a degree are over-represented among those moving to the most disperse group of locations (G5) as compared with those moving to the least disperse location group (G1, the baseline in table 2.12). The same is true for those in managerial occupations moving into the most disperse group of locations. In terms of sex and ethnicity I do not find statistically significant selection.

Throughout my analysis of stochastic dominance in section 2.4 I aggregate locations into groups according to earnings dispersion. I do this because otherwise sample sizes become too narrow as to be able to draw any meaningful inference. In particular, I aggregate locations into groups using earnings dispersion according to table 2.3. Reducing the number of possible destinations from 51 to 5 gives me sufficient sample size to test the implications of the model in terms of stochastic dominance.

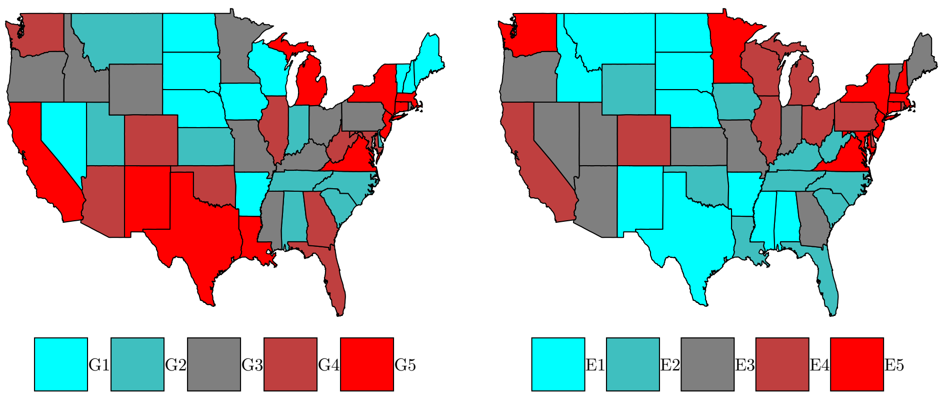
Table 2.3: Dispersion groups

Group	Definition
G1	if dispersion is below 20^{th} percentile
G2	if dispersion between 20^{th}-40^{th} percentile
G3	if dispersion between 40^{th}-60^{th} percentile
G4	if dispersion between 60^{th}-80^{th} percentile
G5	if dispersion above 80^{th} percentile

Within dispersion groups as defined in table 2.3, locations are not clustered geographical or in terms of average earnings (figure 2.3). In the most disperse group (G5) there are states from the west coast (California), south (New Mexico, Texas and Louisiana), the east coast (Virginia, New Jersey, New York, Connecticut and Massachusetts) and northern states (Michigan). Some of these states, those in the east coast, are also at the top in terms of average earnings, while others, the south, are at the bottom. This within group variation in terms of geographical location and average earnings helps on making sure that the selection patterns observed in the data are not driven by state characteristics other than earnings dispersion.

Figure 2.3: Ranking of Locations

(a) Dispersion Groups **(b)** Average Earnings



Before going into testing stochastic dominance relations, I test whether the selection patter that emerged in terms of average earnings in table 2.2 is still present when I aggregate locations into groups. In table 2.13, I regress pre-migration earnings on a set of dummies for each of the dispersion groups. As when using rankings, I observe that as the dispersion of the destination location increases average earnings monotonically increase. Also, as in table 2.2, introducing controls for education reduces the extend of selection and this gets further reduced when I control for occupation and industry. Thus the selection profile implied by the model is still there in terms of average earnings when testing it across earnings dispersion groups.

Finally, an implication from the theory that gets overseen is that earnings dispersion of the location where skills are measured, in this case the previous location, do not play a role in determining the direction of selection, but it does affect the intensity of selection. Higher dispersion at the previous location creates a stronger degree of selection, see (2.8). In table 2.4, I regress pre-migration earnings on the rank of the destination conditioning on dispersion of the home location. What I observe is precisely the pattern implied by the theory. Positive selection increases as the earnings dispersion increases and the magnitude of the selection effect increases with dispersion of the previous location. For example, for those that were in a location in the least disperse group (G1), a one position increase in the ranking of the destination location increases wages by .1%. The same increase in the ranking of the destination location produces a selection effect of .5% for those moving from a location in the most disperse group (G5).

Table 2.4: Selection Interacted with Location Group of Origin

	Home Location Dispersion Group				
	G1	G2	G3	G4	G5
Rank	.001	.002	$.003^{*}$	$.004^{***}$	$.005^{***}$
	(.002)	(.002)	(.002)	(.001)	(.001)
Observations	639	1,394	1,606	2,257	3,321

Note: Each estimate in the table comes from an independent regression. Clustered standard errors by previous state in paranthesis. All columns include fixed effects for year. Characteristics and earnings measured at last observation before migration. p < .1 ·, p < .05 *, p < .01 **, p < .001 ***

2.4 Stochastic Dominance

To provide a formal test of the model implications I use Davidson and Duclos (2013) restricted stochastic dominance test. Davidson and Duclos (2013) test departs from most of the previous literature on testing for stochastic dominance as it posts a null of non-dominance. Aside from its simplicity, testing a null of non-dominance effectively ranks the distributions under study. However, for continuous distributions the null of non-dominance can never be rejected at the tails of the distribution. This leads to a test of restricted stochastic dominance over an interval $[z^-, z^+]$. In my application, I set z^- to be the 5% quantile and z^+ the 99% and test whether I can reject the null of non-dominance when comparing the distribution of pre-migration earnings of migrants moving to the least disperse location group against each other location group. This is

$$F(z|j \in G1) \leq F(z|j \in Gk) \tag{2.13}$$

against the alternative hypothesis

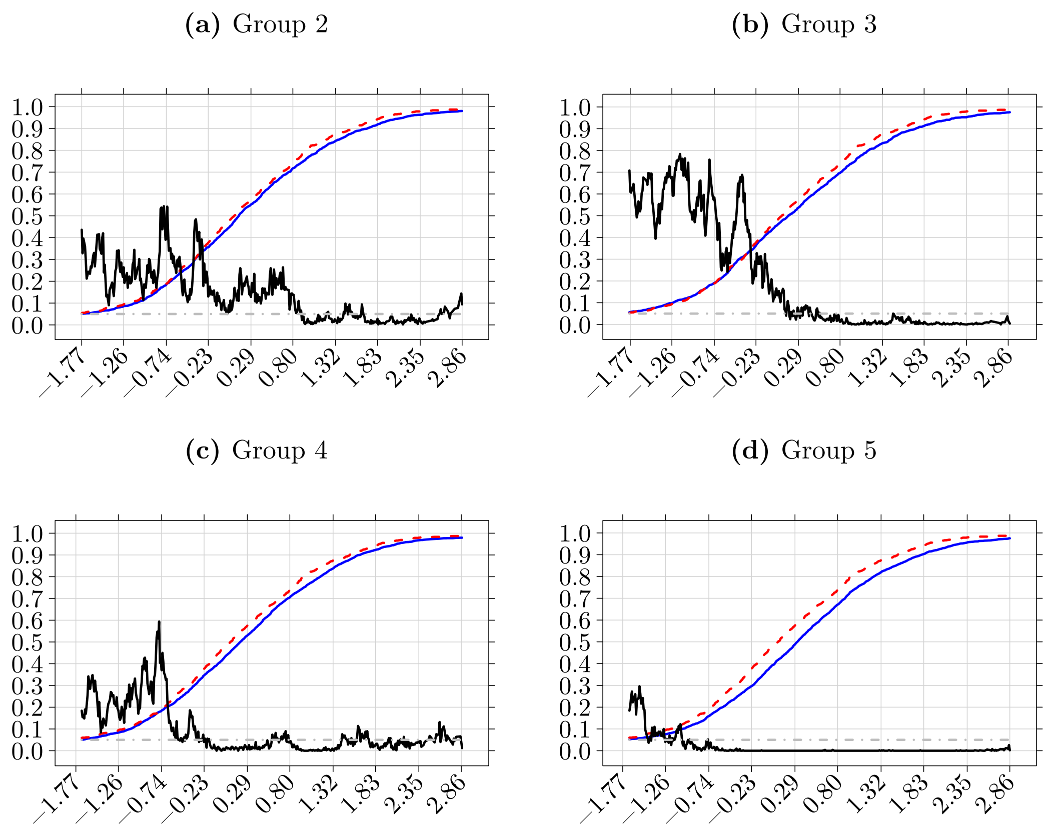
$$F(z|j \in G1) > F(z|j \in Gk)$$

at various z in the interval $[z^-, z^+]$ for each location group (k) other than the least disperse one. In what follows, I present figures comparing the distribution of skill measures of those moving to the least disperse group and all other groups, jointly with bootstrap p-values for Davidson and Duclos (2013) restricted dominance test.

In figure 2.4, I present evidence of selection in terms of pre-migration standardized earnings. In all sub-figures, I present the distribution of pre-migration standardized earnings for those moving into the least disperse location in a dashed red line and the distribution of those moving into the location group that I am comparing in solid blue. Jointly with these two distributions, I provide bootstrap p-values for the point-by-point null in (2.13) constructed with 500 replication samples for each point. The bootstrap samples are drawn using empirical likelihood probabilities constrained under the null as described in Davidson and Duclos (2013).

When I compare the distribution of pre-migration earnings of those moving to the least and second least disperse location (figure 2.4a) the null of non-dominance can only be rejected at small and discontinuous points of the distribution. This means that there is no evidence of dominance between these two distributions. However, as I increase the dispersion of the destination location the dominance relations implied by the model become apparent. The earnings distribution of those moving into the third least disperse location group dominates the distribution of those moving to the least disperse location from the 70% quantile and up. This threshold decreases to around the 20% quantile when looking at those that move into the second most disperse group and close to the 10% quantile when comparing movers to the least disperse group with movers to the most disperse group. The data, therefore, displays the selection pattern predicted by the model: the earnings of those moving to the most disperse location dominates the earnings of those moving to the least disperse and for those locations in the middle there is somehow intermediate selection. As noted by Borjas, Kauppinen, et al. (2018), selection in terms of stochastic dominance provides stronger evidence for the implications of the model than selection in terms of average wages. This is because stochastic dominance implies ordering of the first moments but not the opposite.

Figure 2.4: Pre-Migration Earnings CDF



Notes: Red dashed line is the CDF of pre-migration earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

The same selection pattern on pre-migration earnings is displayed by the data if I select those that are 25-55 years old, figure 2.7. Individuals in this age range, as I show in table 2.2, are the ones most likely to be moving due to labour market motives. Furthermore, selection is driven by both male, figure 2.8, and female, 2.9, movers. For females positive selection at the top of the pre-migration earnings distribution is only found for those moving to the most disperse location group. Females moving to the third least disperse exhibit positive selection but only, roughly, between the 50-90% quantiles and for females moving to the second most disperse group selection happens, roughly, between the 70-90% quantile. For male movers, I observe the same pattern as with the whole population, although p-values are more volatile.

Another possible way on which one could use Davidson and Duclos (2013) restricted dominance test, is to search for the longest interval, $[z^-, z^+]$, on which one can reject the null of non-dominance at a given significance level. This is interesting because it implies that the closer the bonds of the rejection interval are to the bottom and top of the joint support the more power the test has (Davidson and Duclos 2013). When carrying this test the null is over the maximum statistic in the interval $[z^-, z^+]$

$$\max_{z\in[z^-,z^+]}(F(z|G5)-F(z|G1))\geq 0$$

In table 2.5, I present p-values for this null for various intervals. Comparing the distribution of earnings of movers to the least (G1) and second least disperse (G2) location groups shows that there is no clear ordering of the two distributions. There is no interval on which the null of non-dominance can be rejected. A similar conclusion is true for comparison of the least disperse group with the third least disperse group (G3). Although, in that case, the null of non-dominance can be rejected to the right of the median. For the most (G5) and second most disperse (G4) groups, the null of non-dominance can be rejected from the median to the 99% quantile. To the left of the median the null of non-dominance can be rejected at the 90% confidence level over the interval [.3, .5] for the second most disperse group and over the interval [.2,.5] for the most disperse group. Thus the nested test confirms what was already found with the point by point test: the distribution of earnings for movers follows the selection pattern implied by the theory.

Table 2.5: Rejection Regions
Pre-Migration Earnings

$F(z\|G1) \leq F(z\|G2)$				$F(z\|G1) \leq F(z\|G3)$			
Right of the Median				Right of the Median			
$[0.5, 0.6]$	$[0.5, 0.7]$	$[0.5, 0.8]$	$[0.5, 0.9]$	$[0.5, 0.6]$	$[0.5, 0.7]$	$[0.5, 0.8]$	$[0.5, 0.9]$
{0.184}	{0.134}	{0.124}	{0.096}	{0.058}	{0.066}	{0.056}	{0.044}
Left of the Median				Left of the Median			
$[0.4, 0.5]$	$[0.3, 0.5]$	$[0.2, 0.5]$	$[0.1, 0.5]$	$[0.4, 0.5]$	$[0.3, 0.5]$	$[0.2, 0.5]$	$[0.1, 0.5]$
{0.104}	{0.410}	{0.416}	{0.328}	{0.250}	{0.682}	{0.658}	{0.530}
Around the Median				Around the Median			
$[0.4, 0.6]$	$[0.3, 0.7]$	$[0.2, 0.8]$	$[0.1, 0.9]$	$[0.4, 0.6]$	$[0.3, 0.7]$	$[0.2, 0.8]$	$[0.1, 0.9]$
{0.164}	{0.336}	{0.312}	{0.300}	{0.232}	{0.654}	{0.662}	{0.536}
$F(z\|G1) \leq F(z\|G4)$				$F(z\|G1) \leq F(z\|G5)$			
Right of the Median				Right of the Median			
$[0.5, 0.6]$	$[0.5, 0.7]$	$[0.5, 0.8]$	$[0.5, 0.9]$	$[0.5, 0.6]$	$[0.5, 0.7]$	$[0.5, 0.8]$	$[0.5, 0.9]$
{0.010}	{0.068}	{0.042}	{0.030}	{0.000}	{0.000}	{0.000}	{0.000}
Left of the Median				Left of the Median			
$[0.4, 0.5]$	$[0.3, 0.5]$	$[0.2, 0.5]$	$[0.1, 0.5]$	$[0.4, 0.5]$	$[0.3, 0.5]$	$[0.2, 0.5]$	$[0.1, 0.5]$
{0.008}	{0.076}	{0.284}	{0.506}	{0.000}	{0.000}	{0.008}	{0.112}
Around the Median				Around the Median			
$[0.4, 0.6]$	$[0.3, 0.7]$	$[0.2, 0.8]$	$[0.1, 0.9]$	$[0.4, 0.6]$	$[0.3, 0.7]$	$[0.2, 0.8]$	$[0.1, 0.9]$
{0.006}	{0.044}	{0.254}	{0.500}	{0.000}	{0.000}	{0.008}	{0.082}

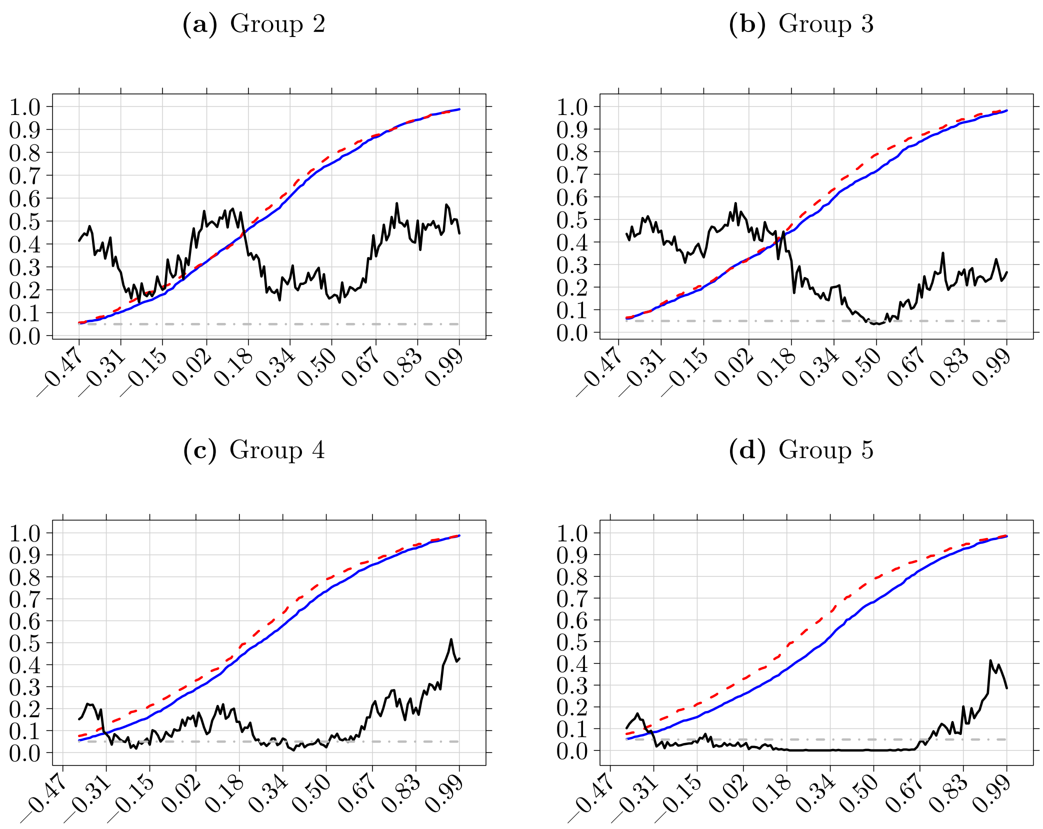
Note: Bootstrap p-values from 500 replications in braces.

A relevant question is whether unobservable or observable characteristics drive selection, as this is informative about transferability of observed and unobserved skills. My estimates of selection on average wages in table 2.2 and 2.13 suggest that both play a role, as controlling for observable characteristics reduces the extend of selection but does not eliminate it completely.

To study selection on observables, in figure 2.5 I display analogous plots to the ones I produced for pre-migration earnings but using pre-migration predicted earnings from specification (2.12). Thus what I compare is the distribution of an index composed of individual characteristics (sex, race, education, occupation and industry) weighted by their returns at the national level. As in the case of pre-migration earnings, selection increases with dispersion at destination. When I compare predicted earnings of those moving to the least disperse location with those moving to the second least disperse the null of non-dominance cannot be rejected at any point. For those moving to the third least disperse group there is a small region where the null can be rejected at the 90% confidence level. This region gets enlarged when I compare those moving to the second most disperse location and even further for those moving to the most disperse location. For those that move into the most disperse location there is positive selection from the 10% quantile up to, roughly, the 85%.

That selection in terms of observables follows the pattern implied by the theory was already suggested by results in table 2.12, where I showed that those moving to the most disperse location tend to be university graduates working in managerial occupations. Results in figure 2.5 confirm the pattern.

Figure 2.5: Pre-Migration Predicted Earnings



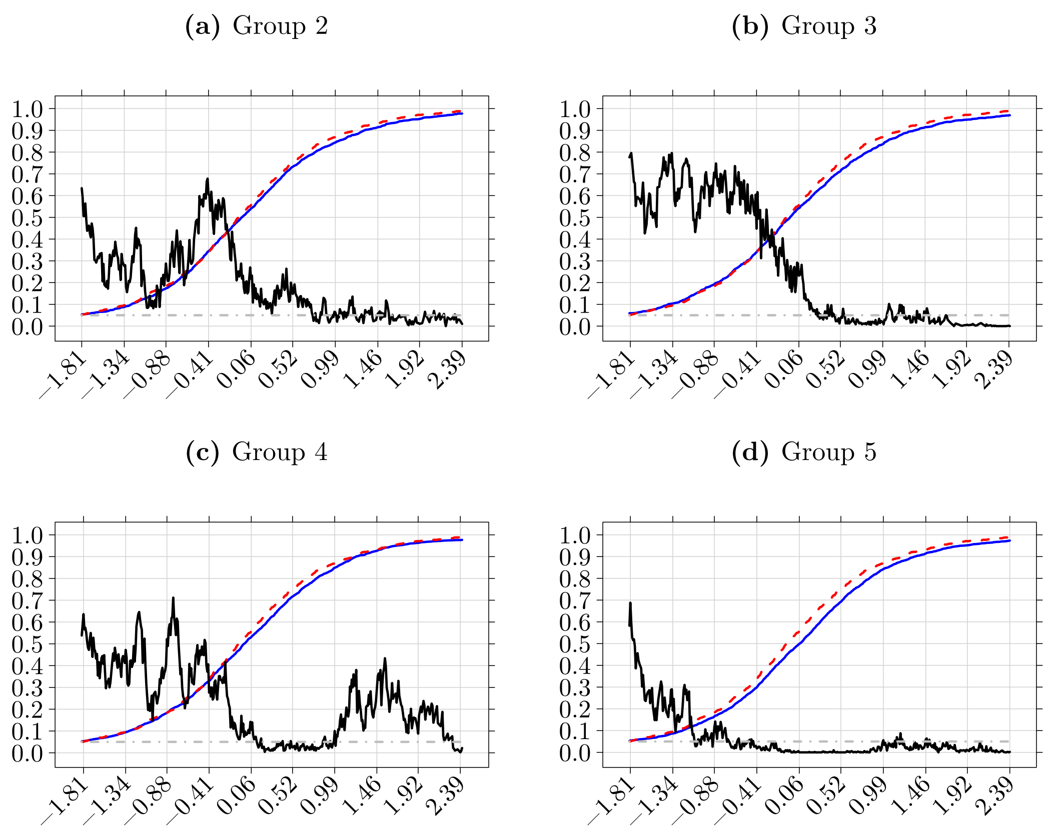
(a) Group 2 **(b)** Group 3 **(c)** Group 4 **(d)** Group 5

Notes: Red dashed line is the CDF of pre-migration Pred of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Finally, the pattern of selection implied by the model is also observed in terms of residual earnings. The distribution of residual earnings for those moving to the least and second least disperse location are fairly similar. Rejection intervals are small an discontinuous, although the p-value for the null of non-dominance drops from about the median and up. This changes when I compare the distribution of those moving to the third least disperse location group. For this sub-set of movers, their residual earnings distribution dominates the distribution of movers to the least disperse location from the 70% quantile and up. However, selection of those moving to the second most disperse location is at the middle of the residual distribution: between the median and, roughly, the 85% quantile. For those moving to the most disperse location group I can reject the null of non-dominance at the 95% level from the 40 to 85% quantile, and at the 90% level from the 10% to the 85%.

As in the case of pre-migration and predicted earnings, the pattern of selection implied by the Borjas model can be observed in residual earnings and it holds for the whole sample (figure 2.6), those that are most likely to move due to labour motives (figure 2.13), males (figure 2.14) and females (figure 2.15).

Figure 2.6: Pre-Migration Residuals



Notes: Red dashed line is the CDF of pre-migration Resid of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

2.5 Conclusion

Since Borjas (1987) seminal paper on migrant selection, a sizeable part of the literature within migration economics has been dedicated to the study of immigrant selection. Most of the existing literature testing the implications of the Borjas (1987) model use a two-location model even in situations where immigrants may face a larger choice set. I extend on the baseline Borjas model and include an arbitrary number of competing locations. I show that, without imposing more structure than in the baseline model, the theory still holds meaningful implications for the least and most disperse locations. For these, the extension implies that the earnings distribution of migrants moving to the least (most) disperse location is dominated by (dominates) the population distribution. For locations not ranked at the extremes, selection depends on a combination of location specific average earnings and migration costs. Selection is, therefore, partially determined by the ranking of locations within the choice set, instead of differences in dispersion between home and destination as typically interpreted in the literature.

To provide empirical evidence, I use data from US interstate migration drawn from the Survey of Income and Program Participation. US internal migration is of interest in its own right given that between one to four percent of the population move states every year during my period of study. But, most importantly, the use of within country migration implies that there are no border constraints and that individuals remain under scope when they move. The absence of border constraints is critical to test the model, as often times countries put in place skill-based border schemes that will contaminate the underlying selection mechanism. Furthermore, that movers stay within scope improves representativeness of the data for migrants and allows me to observe measures of worker skill before they actually move.

I show that positive selection on average wages increases with earnings dispersion of the destination state, and is stronger for those leaving from a state with large earnings dispersion. These two patterns are predicted by the model. Moreover, given that the implications of my extension are expressed in terms of stochastic dominance relations, providing a stronger test for selection than averages, I test whether these dominance relations are observed in the data. For this, I rank all states according to their earnings dispersion and use Davidson and Duclos (2013) restricted stochastic dominance test. I find that the distribution of pre-migration earnings of those moving to the least disperse locations never dominates the distribution of movers to other locations and is dominated by those moving to the most disperse location. This is the pattern implied by the model and is driven by both observables, measured using predicted earnings as a function of education, occupation, industry, sex and ethnicity; and unobservables, measured with residual earnings. Furthermore, the pattern holds for males, females and those between 25-55 years old that, I show, are the ones most likely to move due to labour market motives.

2.6 Appendix

2.6.1 SIPP and CPS moving rates

To make SIPP migration rates comparable with CPS I apply a procedure similar to the one applied in Kaplan and Schulhofer-Wohl (2017). I compare the state FIPS code in the first interview month with the FIPS code in the same month a year after, if these differ then the individual is classified as a migrant. Due to the rotational design of SIPP there are months in which some rotation groups are not interviewed and I drop them. Additionally, due to the overlapping design in place until panel 1996, there might be years with observations from two panels. If that happens I take the average of the two. I follow the same steps for subsequent years. Note that individuals who move and return to the initial location within a given observational year are classified as non-migrants. Finally, I perform the sample selection on the date the migrant identifier is assigned, i.e. end of the observational year, making the sample selection comparable with the one I perform on the CPS data.

2.6.2 Changes on Recorded Information

Across SIPP panels there are different changes on which and how information is recorded. For those variables that I have found differences I proceed as follows.

Birth place

The universe for asking about birthplace in panels 1986-1993 are those 15 and older who have moved at some point. For 1985 and 1996 onwards the universe are those 15 and older independently of whether they have moved or not. For individuals in panels prior to 1996, other than 1985, that appear as not in universe I assign as birth location the state identifier at the time of the interview asking migration history questions, this is typically the reference month of the second wave. For panels 1985 and 1996 onwards there is no change to be made as I have selected into my sample only those who are 18-64. Even with this imputation some observations still have missing birth place.

State of Residence Identifiers

To protect confidentiality, some states with small sample sizes are grouped together in some panels

Table 2.6: State Groups Across Panels

State Grouping	Panels
Maine, Vermont	1985-2001
North Dakota, South Dakota, Wyoming	1996-2001
Alaska,Idaho,Montana,Wyoming	1985-1993
Iowa,North Dakota,South Dakota	1985-1993

Highest Education Attained

For panels 84 to 93 there is no variable directly recording highest education attained. I use information given by *HIGRADE: What is the highest grade or year of regular school this person attended?* and *GRDCMPL: Did he/she complete that grade (HIGRADE)?*.[17] If the grade is not completed then I assign the immediate lower education level. The educational classification provided in panels 1984 to 1993 is

Table 2.7: Highest grade or year of regular school 1984-1993

Label	Value	Re-Coded
Under 15, did not attend or attended only kindergarten	0	0
Elementary	1-8	1
High School	9-12	2
College	21-26	3

Later panels, 1996-2008, provide a more detailed disaggregation. To make both comparable, I aggregate the educational groups provided in panels 1996 to 2008. For this I follow the school level organization given in U.S. Department of Education (2008) and apply the following criteria:[18]

[17] Variable's names suffer modifications across panels. These refer to the 1993 panel.
[18] Labels for the re-coded values are given in table 2.7

Table 2.8: Highest degree received or grade completed 1996-2008

Label	Value	Re-Coded
Under 15	-1	0
Less Than 1st Grade	31	0
1st, 2nd, 3rd or 4th grade	32	0
5th Or 6th Grade	33	1
7th Or 8th Grade	34	1
9th Grade	35	1
10th Grade	36	1
11th Grade	37	1
12th grade, no diploma	38	1
High School Graduate (diploma or GED or equivalent)	39	2
Some college, but no degree	40	2
Diploma or certificate from a vocational, technical, trade or business school beyond high school	41	2
Associate degree in college* (occupational/vocational program)	42	3
Associate (2-yr) college degree (academic degree)	43	3
Bachelor's degree (e.g. BA, AB, BS)	44	3
Master's degree (e.g. MA, MS, MEng, MEd, MSW, MBA)	45	3
Professional School degree (e.g. MD(doctor), DDS(dentist), JD(lawyer))	46	3
Doctorate degree	47	3

*In panels 2004 and 2008, 42 is merged into 43

Race

How race is recorded also varies across panels. I homogenize this characteristic by creating three broad groups White, Black, and Other.

Table 2.9: Race

2004-2008		1984-2001		Re-Coded	
Label	Value	Label	Value	Label	Value
White alone	1	White	1	White	1
Black alone	2	Black	2	Black	2
Asian alone	3	Asian or Pacific Islander	4	Other	3
Residual	4	American Indian, Eskimo or Aleut	3		

2.6.3 Table Appendix

Table 2.10: Moving Reason by Type of Move
(*Current Population Survey*)

	Between States	Within State	Difference	Total
Family reasons	0.279	0.250	0.029***	
	{8045}	{6719}	{1326}	16090
New job or job transfer	0.162	**0.324**	**-0.163*****	
	{4663}	**{8722}**	**{-4059}**	9326
To look for work or lost job	0.029	0.057	-0.028***	
	{830}	{1526}	{-696}	1660
For easier commute	0.085	0.031	0.054***	
	{2442}	{824}	{1618}	4884
Retired	0.002	0.005	-0.002***	
	{66}	{128}	{-62}	132
Other job-related reason	0.023	0.038	-0.015***	
	{656}	{1028}	{-372}	1312
Housing reasons	**0.354**	0.199	0.155***	
	{10216}	{5346}	{4870}	20432
Other reasons	0.068	0.097	-0.029***	
	{1949}	{2604}	{-655}	3898
Total	28867	26897	1970	57734

Note: 1984-2013 CPS data, excluding 1995. Those that move between states or within state across counties only. Significance symbols for two-sided differences in proportions: $^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 2.11: Industry Transitions of Movers
(*Row Proportions*)

	Current Industry								
Previous Industry	Agriculture	Construction	Finance	Manufacturing	Mining	Public Administration	Services	Trade	Transportation
Agriculture	0.760	0.031	0.000	0.042	0.000	0.010	0.125	0.031	0.000
Construction	0.005	0.853	0.005	0.029	0.005	0.010	0.034	0.045	0.013
Finance	0.002	0.004	0.886	0.005	0.000	0.000	0.063	0.029	0.011
Manufacturing	0.005	0.013	0.008	0.864	0.000	0.001	0.057	0.041	0.012
Mining	0.027	0.054	0.000	0.054	0.838	0.000	0.027	0.000	0.000
Public Administration	0.000	0.003	0.016	0.016	0.000	0.877	0.044	0.022	0.022
Services	0.002	0.007	0.010	0.015	0.000	0.006	0.911	0.040	0.008
Trade	0.004	0.015	0.010	0.030	0.001	0.002	0.069	0.852	0.017
Transportation	0.000	0.007	0.005	0.020	0.000	0.004	0.054	0.039	0.871

Table 2.12: Differences in Characteristics
Across Destination Dispersion

	G2	G3	G4	G5
		Education		
No-Education	-0.344	-0.190	-0.394·	-0.427·
	(0.288)	(0.324)	(0.212)	(0.258)
Below High-School	1.080	-0.470	-1.956·	-2.026*
	(0.955)	(1.114)	(1.091)	(1.018)
High School	0.437	-0.309	0.956	-9.361***
	(2.298)	(1.999)	(1.669)	(2.120)
Degree	-1.174	0.968	1.394	11.814***
	(2.244)	(2.339)	(1.927)	(2.441)
		Sex		
Male	1.946	2.275	1.588	0.947
	(1.445)	(1.450)	(1.269)	(1.288)
		Ethnicity		
White	-1.205	-1.806	-0.887	-2.331·
	(1.091)	(1.141)	(1.360)	(1.207)
Black	1.129	1.842*	1.062	1.952·
	(0.954)	(0.913)	(1.050)	(1.089)
Aisan	0.436	0.315	-0.421	-0.066
	(0.501)	(0.523)	(0.515)	(0.558)
Other	-0.361	-0.351	0.246	0.445
	(0.296)	(0.338)	(0.368)	(0.319)
		Occupation		
Managers	-1.549	1.463	1.000	8.128***
	(1.814)	(1.982)	(1.612)	(1.723)
High/Med. Skilled	0.942	-1.406	0.746	-1.719
	(1.953)	(2.025)	(1.981)	(1.819)
Low Skilled	0.631	0.351	-1.258	-5.455***
	(1.704)	(1.538)	(1.368)	(1.381)
Farm Workers	-0.024	-0.408	-0.488	-0.954**
	(0.377)	(0.309)	(0.311)	(0.303)

Note: Differences in proportion with respect to least disperse location group. Clustered standard errors by previous state in paranthesis. Characteristics measured at last observation before migration. $p < .1$ ·, $p < .05$ *, $p < .01$ **, $p < .001$ ***

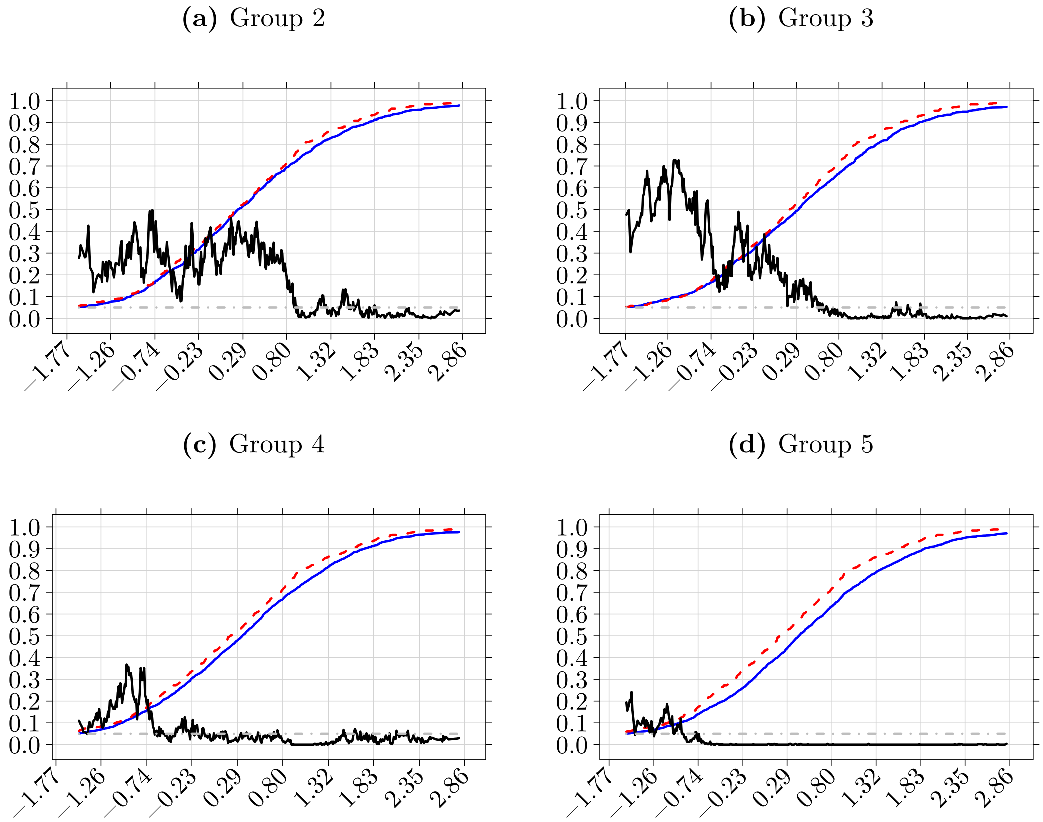
Table 2.13: Wages and Dispersion Groups

	(1)	(2)	(3)	(4)	(5)
G2	0.010	-0.001	-0.003	0.006	0.012
	(0.026)	(0.028)	(0.027)	(0.029)	(0.026)
G3	0.029	0.018	0.020	0.019	0.020
	(0.031)	(0.031)	(0.031)	(0.033)	(0.030)
G4	0.038	0.029	0.029	0.025	0.026
	(0.027)	(0.027)	(0.027)	(0.027)	(0.023)
G5	0.116***	0.109**	0.110***	0.063*	0.047·
	(0.033)	(0.033)	(0.033)	(0.031)	(0.025)
			Fixed Effects		
Sex	N	Y	Y	Y	Y
Ethnicity	N	N	Y	Y	Y
Education	N	N	N	Y	Y
Industry	N	N	N	N	Y
Occupation	N	N	N	N	Y
Obs.	11632	11632	11632	11632	11632

Note: Differences in wages with respect to least disperse location group. All columns include fixed effects for previous state, year and their interaction. Clustered standard errors by previous state in paranthesis. Characteristics measured at last observation before migration. $p < .1$ ·, $p < .05$ *, $p < .01$ **, $p < .001$ ***

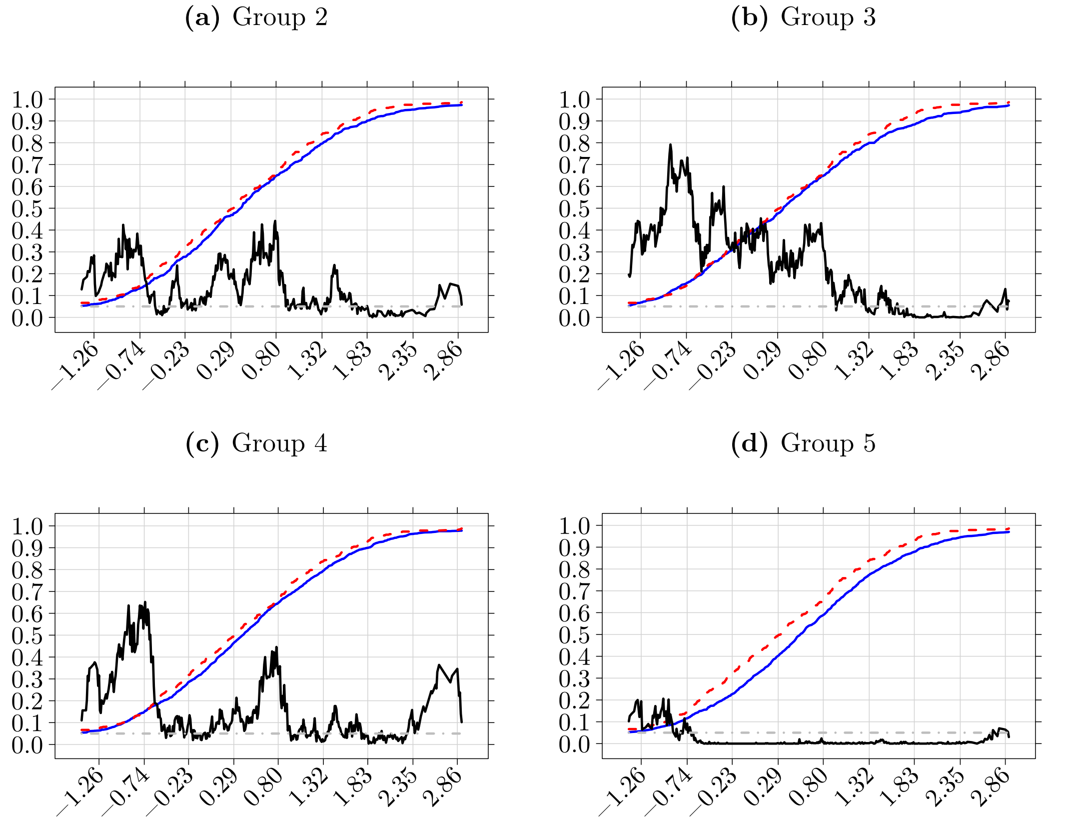
2.6.4 Figure Appendix

Figure 2.7: Pre-Migration Earnings CDF
25-55 Year Old



Notes: Red dashed line is the CDF of pre-migration earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.8: Pre-Migration Earnings CDF
Males

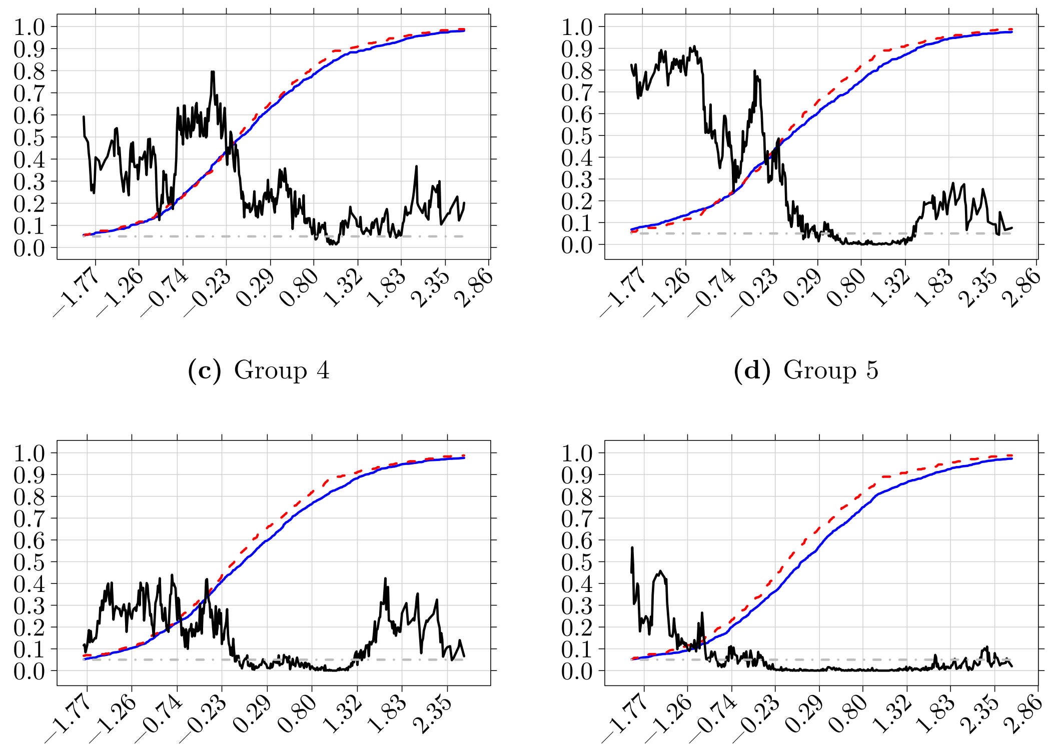


Notes: Red dashed line is the CDF of pre-migration earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.9: Pre-Migration Earnings CDF
Females

(a) Group 2 **(b)** Group 3

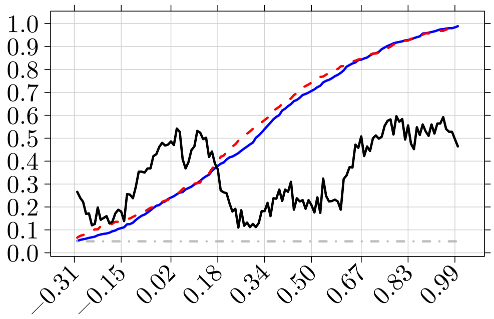
(c) Group 4 **(d)** Group 5



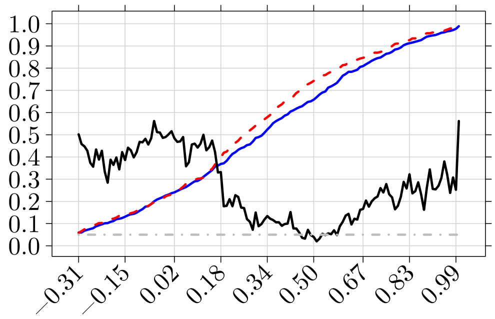
Notes: Red dashed line is the CDF of pre-migration earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.10: Pre-Migration Predicted Earnings
25-55 Year Old

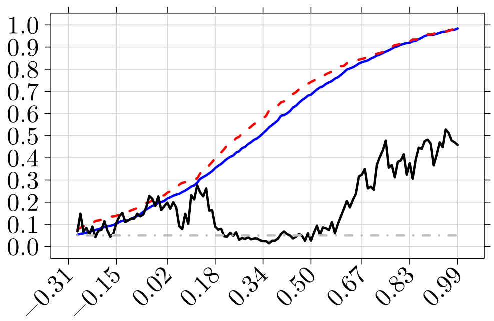
(a) Group 2



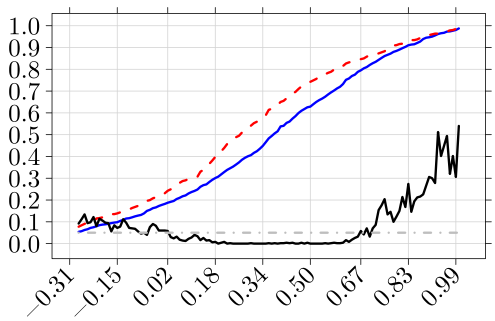
(b) Group 3



(c) Group 4

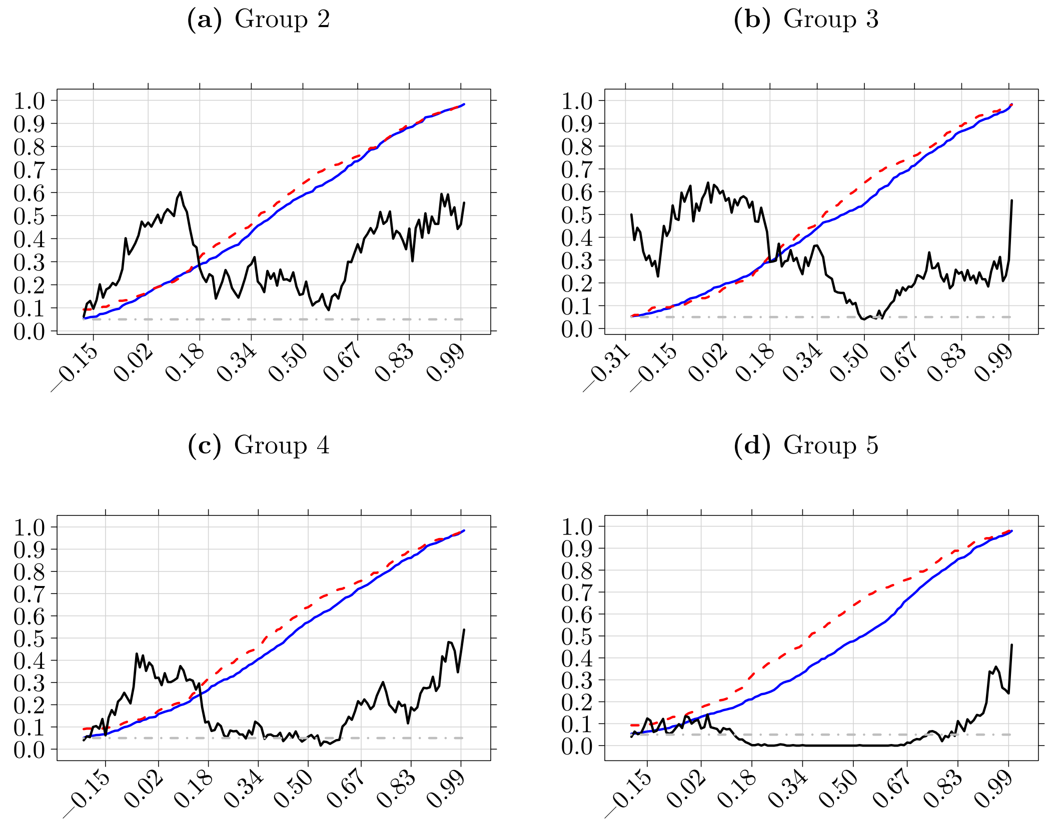


(d) Group 5



Notes: Red dashed line is the CDF of pre-migration predicted earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

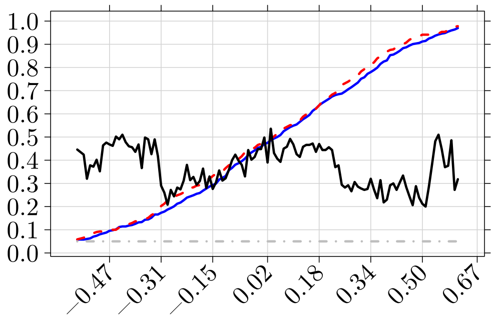
Figure 2.11: Pre-Migration Predicted Earnings
Males



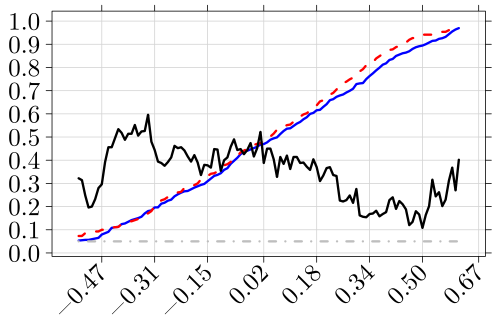
Notes: Red dashed line is the CDF of pre-migration predicted earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.12: Pre-Migration Predicted Earnings
Females

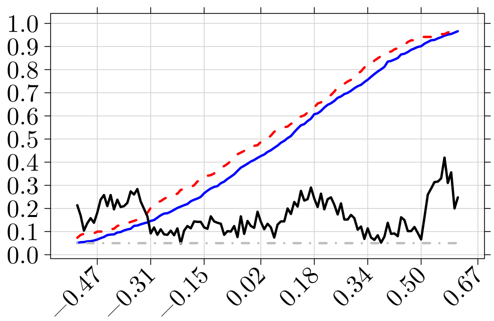
(a) Group 2



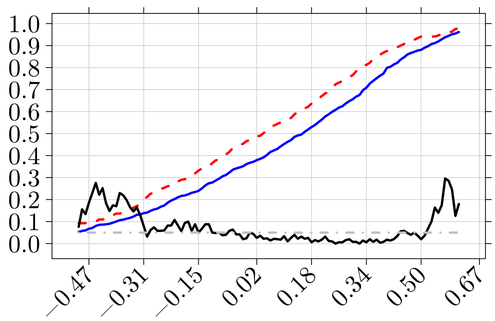
(b) Group 3



(c) Group 4



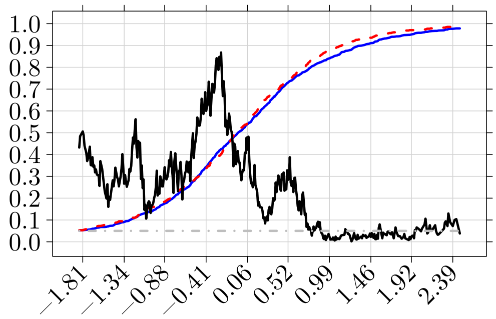
(d) Group 5



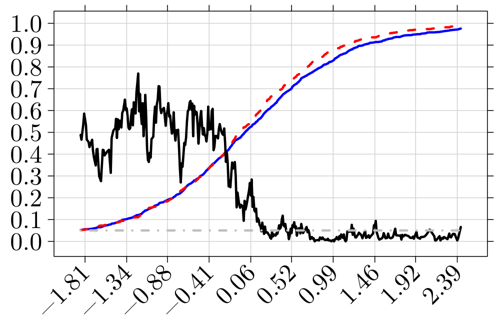
Notes: Red dashed line is the CDF of pre-migration predicted earnings of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.13: Pre-Migration Residuals
25-55 Year Old

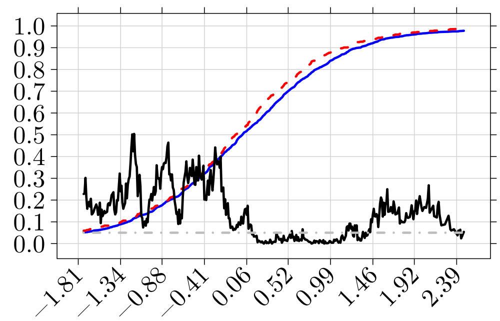
(a) Group 2



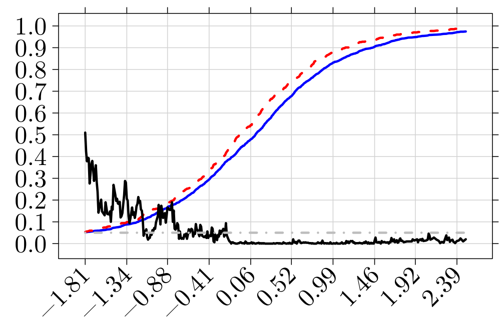
(b) Group 3



(c) Group 4



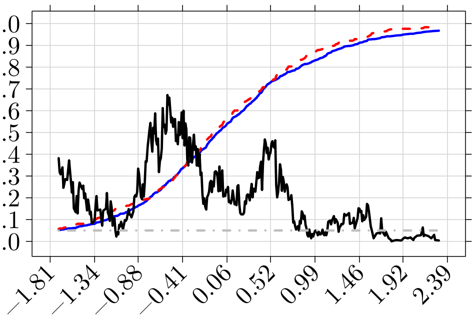
(d) Group 5



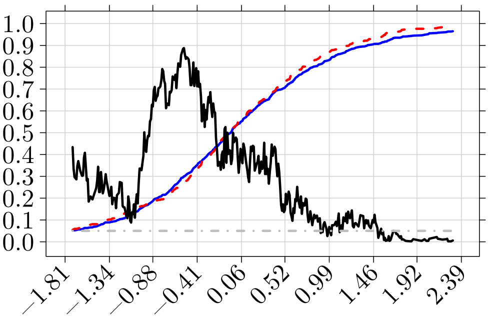
Notes: Red dashed line is the CDF of pre-migration standardized residuals of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.14: Pre-Migration Residuals
Males

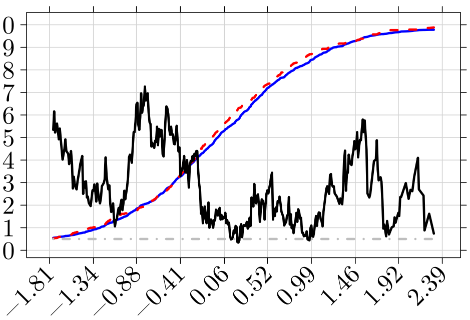
(a) Group 2



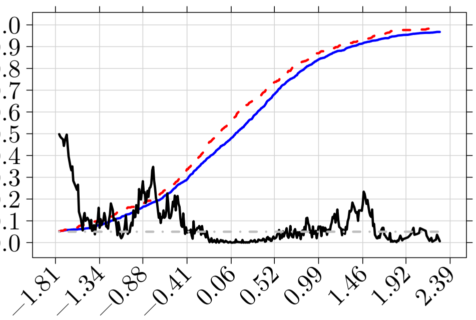
(b) Group 3



(c) Group 4



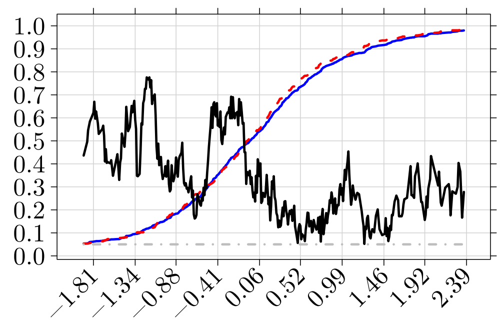
(d) Group 5



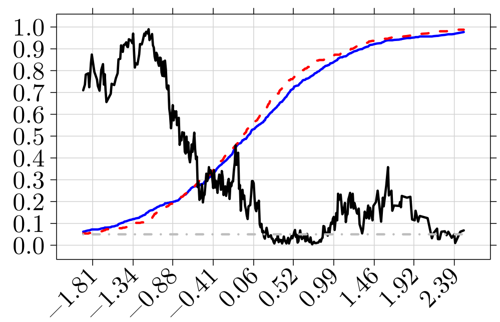
Notes: Red dashed line is the CDF of pre-migration standardized residuals of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Figure 2.15: Pre-Migration Residuals
Females

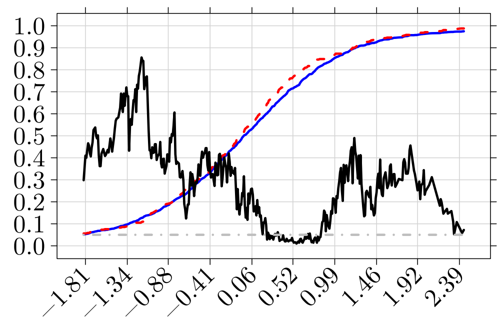
(a) Group 2



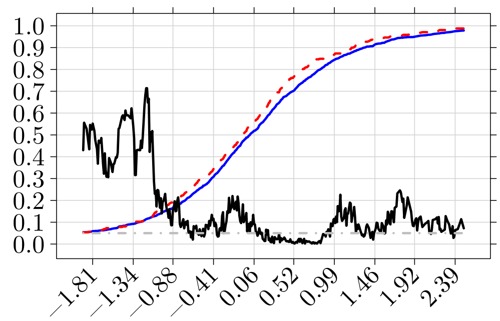
(b) Group 3



(c) Group 4



(d) Group 5



Notes: Red dashed line is the CDF of pre-migration standardized residuals of those moving into the least disperse locations. The black line is the point-by-point bootstrap, 500 replications per point, p-value of Davidson and Duclos (2013) stochastic dominance test. Dashed grey line at .05

Chapter 3

Immigration and Labour Productivity: Evidence from the UK

Abstract

Understanding the effects of immigrants in their receiving economies is a main issue in the labour economics literature and the agenda of policymakers. This is specially true for the UK where immigrant labour shares have more than doubled over the last 15 years. Using data from the UK Labour Force Survey and regional accounts, I investigate the effect of immigrants on a relatively unexplored margin, labour productivity. I show that immigration has a positive effect on output per worker. A ten percentage point increase on immigrant relative supply increases output per worker by a 6.7%. This reduced form effect is composed of differences in marginal returns to immigrant labour and the effect of immigrants on capital, native augmenting technologies, native labour and native skill distribution. I show that immigrants have a positive effect on capital stocks, native labour and skills of natives. Nonetheless, the positive effect of immigrants on labour productivity goes beyond effects on these inputs. In this direction, comparison of two alternative production function estimates, one of them allowing for endogenous native augmenting technologies, suggest that immigrants trigger the development of technologies that complement migrant labour.

3.1 Introduction

With immigrant labour shares more than doubling over the last 15 years, immigration is a central issue in the UK. This has motivated the rise of a rich literature that studies the effect of immigration in various aspects of the British economy. For example, the literature has studied the effect of immigrants on the labour market (Dustmann, Fabbri, et al. 2005; Dustmann, Frattini, et al. 2013; Hatton and Tani 2005; Manacorda et al. 2012; Wadsworth 2010), public services (Preston 2014; Wadsworth 2013) and attitudes towards migrants (Dustmann and Preston 2007; Viskanic 2017). However, there is little evidence on the effect of immigrants on native labour productivity.

The scarcity of evidence on the effect of immigrants on labour productivity contrasts with the centrality that both have for British policymakers. For example, the Secretary of State for Business, Energy and Industrial Strategy has noted the importance of promoting productivity on *"raising living standards, providing funds to support ...public services and improving the quality of life for all ...citizens"* (BEIS 2017, p. 6). Going further, the British Home Office links immigration and productivity as it has made clear its intention to put forward an immigration system that promotes productivity(Home Office 2018).[1]

I contribute to the literature by providing evidence on the effect of immigrant labour on aggregate production and labour productivity in the UK. To produce evidence, I combine data from the UK Labour Force Survey with regional accounts provided by the Office for National Statistics. Combining data from these sources, I create a panel of output, employment and capital stocks across regions and industries from 1998 to 2014. With these data, I estimate reduced-form effects of immigrants on labour productivity.

A common concern when studying immigrant effects is that changes in immigrant stocks and flows are likely endogenous (e.g. Altonji and Card 1991; Card 1990, 2001; Ottaviano, Peri, and Wright 2018; Peri 2012). In my set-up, it is likely that economic shocks jointly determine production levels and demand for immigrant and native labour. If these shocks don't produce parallel effects on immigrant and native labour, for example, because of different responses of native and foreign workers (Dustmann, Glitz, et al. 2010), they will generate an endogenous relative supply of immigrant labour. Or, equivalently, a spurious correlation between immigrant-native labour ratios and labour productivity. I follow a well-established tradition in migration economics (e.g Bell et al. 2013; Dustmann, Frattini, et al. 2013; Ottaviano, Peri, and Wright 2018; Peri 2012) and use previous settlement of immigrants to instrument current stocks.

Altonji and Card (1991) and Card (1990) pioneered using past immigrant settlement as an instrument. The relevance of this instrument is supported by evidence

[1] Another signal of the policy relevance of the effect I explore, is that this paper originates from a Migration Advisory Committee commission to investigate the effect of immigrants on labour productivity. This commission fitted more broadly as part of Government consultancy to the Migration Advisory Committee in the context of the UK leaving the EU.

showing that immigrants locate in areas where workers from their country of origin are located (see Bauer et al. 2007; Jaeger 2007; Munshi 2003). However, the instrument isn't free of validity concerns. Borjas (2014, p. 86) notes that past settlement of immigrants may respond to economic incentives at the time, implying that serial correlation in economic conditions may render the instrument invalid. In this line, Cadena and Kovak (2016) and Jaeger (2007) show that immigrants respond to economic incentives, making serial correlation a threat to validity. I address this concern in two ways. First, I use previous stocks of immigrant labour at the industry-region level measured ten years before the starting of my estimation sample. If the correlation process is stationary, dependence on past conditions will vanish as time goes on. Second, in combination with the instrument, I introduce a comprehensive set of region, industry and year fixed effects. With these fixed effects, in my preferred specification, identification comes from variation within industry and region across time. Causal interpretation, therefore, relies on shocks within industry-region not being strongly serially correlated.[2] This is possible because I project past immigrant stocks at the industry and region level using national level growth, what gives me variation in the instrument within industry and region. Using this empirical strategy, I find that a ten percentage point increase in the relative supply of immigrant labour produces a 6.7% increase in output per worker. When evaluated at the average labour productivity, this implies a £5,600 increase in annual output per worker.

To better understand what the reduced form estimate actually estimates, I introduce a Constant Elasticity of Substitution (CES) production function. This technology produces output combining immigrant and native labour and these two inputs may have different productivity levels and constant, but otherwise unconstrained, elasticity of substitution. In the short run, the reduced form effect of immigrants on output depends on the relative supply of immigrant labour, its marginal product and the effect on native labour. This, therefore, calls for caution when extrapolating reduced-form estimates from other countries or even other periods and levels of aggregation, as they depend on the distribution of immigrant and native labour.

Once one moves from the short run, immigrants can also influence output through changes induced in other inputs, as, for example, the skill distribution of natives (Hunt 2017; Llull 2017), development of new technologies (Peri, Shih, et al. 2015) or the accumulation of capital. I use the CES framework to decompose the reduced form effect estimate of immigration on output. Using data on capital stocks and skill distribution of natives, I show that immigrants have a positive effect on both.[3]

[2]Here not being strongly correlated means that recovery from industry-region specific shocks must take, at most, ten years.

[3]The positive effect of immigration on the skill distribution of native workers is consistent with evidence for the UK in Dustmann, Frattini, et al. (2013) showing that immigrants depress the wages of natives at the bottom of the wage distribution and have a positive effect at the top. Given that low-skilled workers tend to be at the bottom of the wage distribution, the evidence in Dustmann, Frattini, et al. (2013) suggests that immigration may alter incentives to accumulate skill. This is also consistent with evidence from the US in Hunt (2017) who shows that immigrants have a positive effect on the educational achievement of natives, with a stronger effect for children of low-educated parents.

The effect of immigrants on output through technologies that complement native labour is more complex to study. This is because these technologies are typically not observed. Instead of directly studying the effect of immigrants on native augmenting technologies, I provide suggestive indirect evidence. Using my decomposition of the reduced form effect I show that, for a wide range of production function parameters, the evidence I produce is consistent with immigrants triggering the development of technologies that complement them. This is consistent with the endogenous technological change framework developed by Acemoglu (1998, 2002) showing that, when inputs are gross substitutes, increasing the relative supply of one input creates incentives to develop technologies that complement it. The literature (Manacorda et al. 2012; Ottaviano and Peri 2012) provides evidence of gross substitution between immigrant and native labour. Although available estimates are not free of concern (see Dustmann and Preston 2012; Dustmann, Schönberg, et al. 2016), the general understanding is that available substitution estimates are underestimated.

Other than through the reduced-form decomposition, I produce further suggestive evidence for the effect of immigrants on native augmenting technologies by comparing production function and reduced form estimates. To start with, I estimate the CES production function restricting native augmenting technologies to be constant. This estimate produces a marginal effect of immigrants on labour productivity that is only positive when allowing effects on capital stocks. This is not consistent with my reduced form estimates controlling for capital stocks, as they show that, controlling for capital and native skill supply, there is, still, a positive and significant effect of immigrants on labour supply. Specifically, rising immigrant relative supply by ten percentage points increase labour productivity by 5%. Therefore, it suggests that the production function is misspecified and that native labour augmenting technologies depend on relative immigrant supply. This is further supported by my comparison of the structural estimates with alternative production function estimates using a polynomial approximation. Where the latter is compatible with Acemoglu (1998, 2002) endogenous input augmenting technologies. The polynomial approximation produces marginal effects that are positive and always larger than their structural counterparts. Therefore, it further suggests that immigrants trigger the development of technologies that complement them. However, this piece of evidence should be taken with caution, as the null of equality of marginal effects can only be rejected for a small set of points.

In terms of existing literature, my work closely relates to research in Boubtane et al. (2016), Ortega and Peri (2014), and Ottaviano, Peri, and Wright (2018). Boubtane et al. (2016) and Ortega and Peri (2014) provide cross-country evidence, including the UK, showing that immigration has a positive effect on labour productivity. I add to their work in that I exploit variation across industries and regions within the UK. This implies that my estimates are representative for the UK, what is important given that reduced-form estimates depend on the precise input mix and this varies across countries, regions and industries.

Ottaviano, Peri, and Wright (2018) investigate the effect of immigration on British service firms, showing that immigration has a positive effect on labour productivity. The effect of immigration on the service sector can be quite different from the average effect of immigration on aggregate labour productivity. For example, Ottaviano, Peri, and Wright (2018) highlight the possible role of immigrants on providing knowledge or connections with their home country and the impact this may have when selling services that are country specific, such as economic consultancy. Nonetheless, these knowledge and connections might be of lesser importance for industries producing goods that do not strongly dependent on the market of destination or focus on the interior market. My research, therefore, adds to Ottaviano, Peri, and Wright (2018) in that my estimates reflect the average effect of immigrants on output and labour productivity across the whole British economy.

More broadly my research relates to work for the US in Ottaviano and Peri (2006) and Peri (2012) showing that immigration has a positive impact on labour productivity. Peri (2012) exploits a shift-share instrument and distance from the Mexico-US border to estimate the effect of immigrants on various components of an output decomposition using a CES technology combining high a low skilled labour. He reports a positive effect in total factor productivity and the adoption technologies that augment the productivity of low-skilled labour. Ottaviano and Peri (2006), on the other hand, explore the effect of diversity across American cities on wages and rents. They provide evidence showing that increasing diversity has a positive effect on both wages and rents, and argue that these are consistent with a positive effect on the productivity of natives. I add to Ottaviano and Peri (2006) and Peri (2012) in that my baseline estimates do not depend on any assumed functional for the production function nor on assumptions about how wages, and rents, are set.

Finally, through estimation of the parameters of the production function, I also contribute to the literature estimating elasticities of substitution between immigrant and native labour (Card 2009; Manacorda et al. 2012; Ottaviano and Peri 2012). My main contribution to this literature is that my estimates are robust to gaps between contribution to production and labour remuneration. This is because I do not maintain any assumption about how wages are set. This generalization is relevant as a large body of evidence in the UK (Dustmann, Frattini, et al. 2013; Dustmann, Schönberg, et al. 2016) and elsewhere (Aydemir and Skuterud 2008; Eckstein and Weiss 2004; Izquierdo et al. 2009) shows that immigrants earn lower wages than comparable natives.[4] My elasticity point estimates are between 2.6-6.1. This covers estimates for the UK in Manacorda et al. (2012) and is below US estimates in Card (2009) and Ottaviano and Peri (2012). However, it should be noted that I cannot reject a null of perfect substitution.

[4]Dustmann and Preston (2012) and Dustmann, Schönberg, et al. (2016) show how immigrant downgrading can bias elasticity estimates obtained from wage differentials across education cells.

3.2 Data and Empirical Strategy

I obtain measures of immigrant and native labour from the UK Labour Force Survey (LFS). The LFS is the largest household survey in the UK and it is the source for official measures of employment produced by the Office for National Statistics (ONS). Initially implemented in 1973, the LFS started as a biennial survey, becoming annual in 1984 and quarterly from 1992 onwards.

Since its introduction, LFS respondents are asked about a number of socio-economic characteristics, including country of birth, employment status, region of work, occupation and industry. This allows obtaining employment measures at different levels of detail and has made the LFS a popular survey to study the effect of immigration on the British labour market (Dustmann, Fabbri, et al. 2005; Dustmann, Frattini, et al. 2013; Manacorda et al. 2012; Ottaviano, Peri, and Wright 2018).

In my analysis, I combine data from all quarters between 1998 to 2014 to compute average annual employment by immigrant status, region of work, occupation and industry. In most of my analysis I define employment as worker headcount. However, in section 3.3, I show that may main results are robust if I measure employment as total hours of work and use predicted wages to weight the headcount measure.

Following existing literature on the effect of immigration on the British economy,[5] I classify workers into native and immigrant status according to workers' country of birth. I start by classifying workers into nine groups: UK born, Irish, EU pre-2004 (excluding Irish), Eastern European, other European Economic Area, Asian, North American or Caribbean, African and Other. From these nine groups I define natives (UK born and Irish) and immigrants (all others).[6] Once I classify workers according to their country of birth, I use regions and industries to define markets as a region-industry pair and aggregate total immigrant and native labour for each market. The resulting sample of employment figures comprises thirty-one industries (table 3.6) in each of the twelve regions used by the ONS, including Northern Ireland, Scotland and Wales (figure 3.1), from 1998 to 2014.[7] To this sample of employment, I add my outcome of interest, gross domestic product, from ONS production side regional figures covering the same period.[8]

Finally, to compute capital stocks I use ONS regional gross fixed capital formation data covering the period 2000 to 2014. Because the available regional gross fixed capital formation data is aggregated into eleven industry groups, I impute the regional gross fixed capital formation from these groups into the thirty-one industries in the employment data by using national level gross fixed capital formation. Specifically, I impute gross fixed capital formation and compute capital stocks as follows. From national gross fixed capital formation I obtain the proportional contribution of each industry in each group and use these to allocate regional gross fixed capital formation

[5] For example, Bell et al. (2013), Dustmann, Fabbri, et al. (2005), Dustmann, Frattini, et al. (2013), Manacorda et al. (2012), and Ottaviano, Peri, and Wright (2018)

[6] I classify Irish as natives due to their geographical proximity and historical ties.

[7] I disregard all employment from activities of households and extraterritorial organizations.

[8] All in 2013 pounds.

figures across the thirty-one industries. After this, I compute initial capital stocks in 1999 from national data assuming that industries across different regions operate at the same capital-employment ratio. I combine the initial capital stock, the gross fixed capital formation figures and national level depreciation rates to compute stocks from 2000 onwards as previous capital stock plus current capital formation minus consumption of capital. With this I produce capital stock figures for every market, i.e. region-industry, for the period 2000-2014 that I add to the employment and output dataset. The resulting dataset is a panel of markets for which I annually observe the number of immigrant and native workers, output and capital stocks.

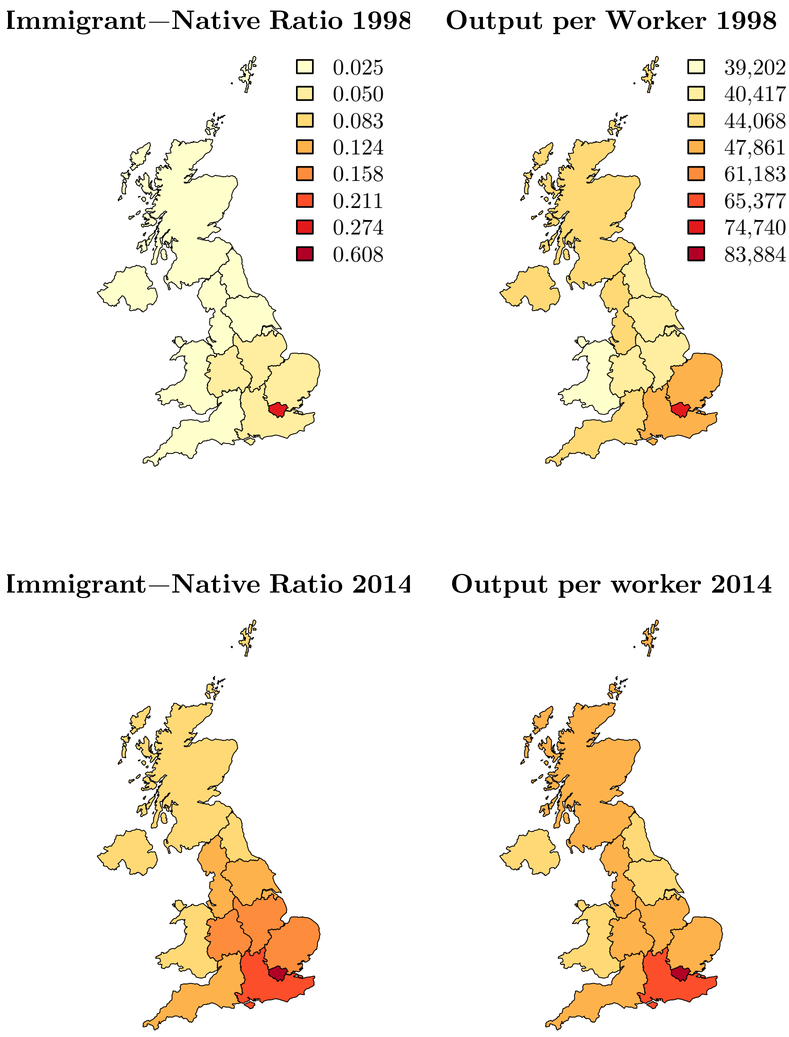
3.2.1 Descriptives

Most of the variation on immigrant employment happens across regions. London has, by a large margin, the largest immigrant concentration. In 1998, in the LFS sample, for every four native workers in London there was a foreign born worker. This is well above the national average at fifteen natives per immigrant worker. Over the period 1998-2014, immigrant-native ratios have been increasing in all UK regions. This increase has been fairly even across regions over time, with persistently large immigrant-native ratios in London and neighbouring regions, particularly the South-East (see figure 3.1). Having said this, with an average annual growth of 8.5% and 11%, Northern Ireland and Scotland have been the UK regions with the largest growth rate of immigrant relative labour supply. On the other side of the spectrum, London with an average annual growth of 5% and Wales with 6% are the UK regions that have experienced the smallest average annual growth. However, London has experienced, by far, the largest year-to-year variation in terms of levels of workers. On average around 60,000 additional immigrant workers entered the London labour force every year between 1998-2014. That is 120% larger than in the South East of England, the region with the second largest year-to-year variation, and sixteen times larger than in Northern Ireland, the region with the smallest year-to-year level variation. Despite this, immigrant geographical concentration has diminished over time. In 1998, 42% of the total immigrant work force was concentrated in London. By 2014 this proportion decreased to 34.6% of the immigrant workforce and most other regions experienced an increase in their shares of the total immigrant workforce. Measuring immigrant concentration across regions using the Gini index returns a measure of .55 in 1998, declining to .47 by 2010, and stable thereafter.

The regional distribution of output per worker has some similarities with the distribution of immigrant labour (figure 3.1). With annual average output per worker at £83,283, London has a worker productivity 33% larger than the South East of England and 90% larger than Wales, the regions with the second largest and the smallest worker productivity. These regional differences have been stable across time. For example, measuring regional inequality across time using the Gini index returns an index that is always around .1 and has no clear time trend. Thus the UK presents larger regional inequality in terms of immigrant labour distribution than in terms of output per worker and both have followed different time trends.

It is well known (e.g. Blundell et al. 2014; Pessoa and Van Reenen 2014) that output productivity in the UK has stagnated since the beginning of the Great Recession. However, the picture across regions is heterogeneous. Output per worker was growing in London between 1998 and 2007, rapidly decreasing between 2007 and 2009 and stagnant from then onwards. A similar pattern holds for the East of England, Northern Ireland and Yorkshire and The Humber. But output per worker stagnation seems to precede the Great Recession in most other regions (figure 3.4).

Figure 3.1: Regional Differences and Evolution



In comparison with differences across regions, variation on immigrant labour relative supply across industries is less pronounced. Immigrant relative supply across broad industrial groups range from, roughly, ten native workers per immigrant in electricity and water supply to five natives per immigrant in manufacturing. This is shown in the first panel of table 3.1, where I use I to refer to immigrant labour and N for native. With the exception of extractive industries, immigrant relative supply is also homogeneous across industries and occupations (third and fourth panel in table 3.1.[9] Although immigrants are, at least marginally, over-represented in high skilled occupations in all industry groups except for manufacturing and services.

In terms of output per worker, differences across industries are larger than differences across regions. Average output per worker in real estate is at £858,646, this is £725,349 per worker higher than in Electricity and Water Supply, the industry group with the second highest worker productivity. However, as with regions, differences in labour productivity across industries have decreased over time. For example, in 1998 output per worker in real estate was 4 times higher than in electricity, gas and and air-conditioning supply; and 40 times higher than labour productivity in manufacturing of textiles.[10] The same figures reduce to 3 and 13 times in 2014. This reduction in differences on labour productivity across industries was fully general, as indicated by the Gini index declining from 0.55 in 1998 to 0.38 from 2012 onwards.

[9] I classify workers in one of the occupational groups by mapping four digit occupations into National Qualification Framework levels. High skilled occupations correspond, roughly, to a bachelors degree or higher. The occupational mapping was provided by the Migration Advisory Committee.

[10] Electricity, gas and and air-conditioning supply is included inside Electricity and Water Supply in table 3.1. Textile manufacturing is included inside Manufactures.

Table 3.1: Native-Immigrant Labour and Output
(*Means and Standard Deviations, all in thousands*)

	Total	Extractive	Manufactures	Eletricity and Water Supply	Services	Real Estate
I	7.810	1.395	2.669	1.343	15.332	2.318
	{16.516}	{1.826}	{2.928}	{1.440}	{23.258}	{3.750}
N	59.558	14.271	19.508	12.896	116.940	15.793
	{75.123}	{11.687}	{15.409}	{6.096}	{86.457}	{9.990}
I/N	0.147	0.129	0.180	0.108	0.125	0.112
	{0.220}	{0.215}	{0.260}	{0.129}	{0.186}	{0.113}
				High Skilled Occupations		
I	2.322	0.392	0.606	0.381	4.773	0.424
	{6.920}	{0.854}	{0.827}	{0.565}	{10.135}	{0.977}
N	15.600	1.583	4.547	2.908	31.701	2.752
	{26.650}	{2.365}	{3.795}	{1.824}	{35.013}	{2.137}
I/N	0.155	0.307	0.168	0.145	0.121	0.132
	{0.494}	{1.637}	{0.352}	{0.335}	{0.163}	{0.259}
				Medium/Low Skilled Occupations		
I	5.488	1.003	2.063	0.962	10.559	1.895
	{11.936}	{1.416}	{2.510}	{1.143}	{16.944}	{2.905}
N	43.958	12.688	14.961	9.989	85.239	13.042
	{57.013}	{11.023}	{12.449}	{4.898}	{67.917}	{8.192}
I/N	0.152	0.127	0.195	0.102	0.123	0.113
	{0.296}	{0.301}	{0.385}	{0.128}	{0.201}	{0.118}
				Output per Worker		
Native	94.747	57.460	65.969	147.649	55.210	949.631
	{183.727}	{33.667}	{77.185}	{87.201}	{27.854}	{412.327}
All	83.831	52.752	56.807	133.297	48.302	858.646
	{166.970}	{32.516}	{74.169}	{77.526}	{19.056}	{374.361}

Note: Pooled Sample. Extractive includes Agriculture, Mining and Quarrying. Standard deviations in braces.

3.2.2 Empirical Strategy

The existence of persistent differences on immigrant mix and output across markets points towards fixed market-specific unobserved characteristics. Such unobserved characteristics might attract inputs either from other markets or outside the UK leading to endogenous input mix. As a naive example, if one compares output and immigrant-native ratios in London and Wales, would arrive to the conclusion that immigration has a positive impact on output and labour productivity. Nonetheless, it is not hard to imagine that agglomeration economies (Ottaviano and Puga 1998) might boost production at the same time that reduce the cost of immigration, for example through lower transportation costs or denser immigration networks.

Accounting for fixed unobserved heterogeneity, therefore, is an important part of the analysis in section 3.3. In my analysis, my first step is estimation of the reduced form effect of immigrant-native mix on output, section 3.3. To see where my identifying variation comes from let y_{mt} be (log-)output in market m and period t. My interest is on estimation of the output semi-elasticity (β) in the following equation

$$y_{mt} = \beta \frac{I_{mt}}{N_{mt}} + \tau_m + \nu_{ti(m)} + \omega_{tr(m)} + \epsilon_{mt} \tag{3.1}$$

Where $i(m)$ and $r(m)$ are the industry and region of the m^{th} market and τ, ν and ω are market, year-industry and year-region fixed effect. Estimation of (3.1) with a fixed effect estimator implies that β is identified from within market variation across time. Market fixed characteristics and time varying shocks at the industry and region level are differenced out, implying that causal interpretation of the fixed effect estimate of β from an equation like (3.1) relies on immigrant-native mix being uncorrelated with industry-region-year specific shocks. In table 3.8, I display market transitions from Understanding Society data.[11] From year to year, 7.6% of natives and 9.5% of immigrants change market, i.e. they change either industry, region or both. Among those that move markets, the vast majority, above 80%, only change industry. For both natives and immigrants around 4% of those that change industry or region change both at the same time. This suggest that labour has a higher mobility response to industry rather than region shocks. Nonetheless, the fact that some workers change region and industry simultaneously means that there can be endogenous responses to industry-region specific shocks. To deal with this source of endogeneity I use the shift-share instrument pioneered in the migration literature by Altonji and Card (1991) and Card (2001) and widely applied since then (for UK examples see Bell et al. 2013; Ottaviano, Peri, and Wright 2018). The rationality behind this instrument comes through the effect of networks, where workers that are already settled in the receiving market provide connections to newcomers shifting their moving costs.

To create the instrument, I use the annual LFS from 1985-1988 to compute the average stock of immigrants from each of nine country-regions of origin in each market. By measuring the stock of immigrants ten years before the start of my period of analysis I aim to obtain variation that is independent from current economic conditions. In this sense, in my set-up, validity of the instrument relies on industry-region-time specific shocks not being strongly serially correlated, as all other region, industry, market (i.e. region-industry), region-year and industry-year shocks are captured by the fixed effects.

Once I have past employment stocks, I project them into the current year using employment growth from every country of birth at the national level. When computing national growth employment, I produce a different growth figure for each market. These differences across markets come from me leaving out employment from the market on which I'm going to use the growth figure to project employment. I use this leave-out projection to address validity threats derived from computing projections that include own employment. This is because it could be the case that

[11] The LFS has a longitudinal dimension but is more limited in terms of temporal span and limits the available information due to anonymity concerns.

immigration from some country is driven by demand forces in a particular market,[12] therefore, challenging the exogeneity requirement of the instrument. This would be the case even if the base employment is independent from current economic conditions (Goldsmith-Pinkham et al. 2017).

The combination of the leave-out projection and market specific settlement gives me variation on the instrument across years within market. This variation is crucial to identify the semi-elasticity (β) in (3.1) with the full set of fixed effects. Otherwise, variation would be only at the market level and this is absorbed by market fixed effects.

Finally, the aggregation of survey data into narrow markets may lead to measurement error on the market's employment measures and, thus, on the immigrant-native mix.[13] In the context of immigration wage effects, Aydemir and Borjas (2011) have shown that attenuation does play a role when using aggregate figures and that its impact grows fast as underlying sample sizes decrease. In a linear-on-parameters set-up, as in (3.1), the IV estimator deals with both the endogeneity and attenuation bias problem.

3.3 Immigrant Labour and Output

Here I present reduced form estimates of the effect of immigration on output. Table 3.2 displays estimates from specification (3.1) where I progress from a baseline specification that includes region, industry and time fixed effects, into a more comprehensive specification with market (region-industry), industry-year and region-year fixed effects. Identifying variation for the parameter of interest is at the market level, so all tables provide market clustered standard errors.

The left half of table 3.2 displays OLS estimates both without (columns (1) and (2)) and with controls for total employment ((3) and (4)). All of the OLS estimates show a negative, typically non-significant, effect of immigration on output that reduces in size when I include the interacted fixed effects. On the other hand, the IV estimates on the right half of table 3.2, show a positive and significant effect. With the full set of fixed effects, the IV estimates show that a ten percentage points increase on immigrant relative supply increases output by a 9.8%. If evaluated at the averages in table 3.1, it implies that an increase of 596 immigrant workers, i.e. a 7.6% increase on the average immigrant stock, leads to a .98% increase on output. This 7.6% increase on immigrant employment is, roughly, equivalent to the annual growth rate of the average immigrant stock observed between 1998 and 2014, see figure 3.6.

That increasing immigration increases output is unsurprising. This is because increasing the relative supply of immigrant labour increases the stock of productive

[12]For example, during the 60s-70s Pakistani workers were driven into particular areas of UK by the textile sector,(Finney and Simpson 2009, chapter 3).

[13]See figure 3.5 where I compare my LFS employment measures with employment from Workforce Jobs. Workforce Jobs combines data from employers' surveys, the LFS and administrative sources. However, it cannot be used directly for my analysis as no information on nationality or country of birth is provided.

inputs as long as immigrants do not substitute native labour one-to-one and do not crowd-out capital accumulation. Therefore, a more interesting question to explore in the data is whether rising the relative supply of immigrants while keeping total labour constant increases production. Thus, in columns (7) and (8) of table 3.2, I control for (log-)total employment. The inclusion of (log-)total employment ($log(I + N)$) as an additional control introduces an additional endogenous variable that I instrument with the log of projected past employment. Both instruments, for immigrant relative supply and total employment, exploit similar variation on past employment. Thus, it is not surprising that the first stage becomes weaker. Nonetheless, in the preferred specification with the full set of fixed effects, column (8), the F-statistic is above usual rules of thumb. In regards to this, the rather large increase in precision between columns (5) and (6) is due to the much better fit produced by the specification with market specific fixed effects. For example, the market fixed effects account for large structural differences on input levels between real estate in London and Wales. Without these fixed effects these input differences create large residuals producing larger standard errors. This can be seen by comparing standard errors in columns (7) and (8). Once I control for total employment, the introduction of market fixed effects does not produce that much of a difference in terms of precision.

The estimate in column (8) shows that increasing immigrant relative supply by ten percentage points while keeping total employment constant increases output by 6.7%. This effect is actually the effect of immigrants on labour productivity, as the estimate from regressing output on immigrant-native ratios controlling for total employment is identical to the estimate when the left hand side is output per worker. Differences are only in terms of the estimated effect of increasing employment on output.[14] At the averages in table 3.1, increasing the relative immigrant labour supply by ten percentage points increases output per worker by £5610.

[14]The effect of increasing employment on (log-)output per worker is just the effect of increasing employment on (log-)output minus one.

Table 3.2: Immigrant-Native Mix and Output

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I/N	-0.236	-0.080	-0.225	-0.102*	2.588*	0.978*	0.733*	0.670*
	(0.205)	(0.051)	(0.156)	(0.047)	(1.190)	(0.398)	(0.300)	(0.320)
$log(N+I)$			0.911***	0.135***			1.320***	0.324
			(0.076)	(0.032)			(0.102)	(0.221)
	First-Stage (I/N)							
$\widetilde{I/N}$					1.015***	0.781***	1.115***	1.105***
					(0.297)	(0.164)	(0.290)	(0.221)
$\widetilde{log(N+I)}$							-0.083***	-0.429
							(0.018)	(0.223)
	First-Stage ($log(N+I)$)							
$\widetilde{I/N}$					1.015***	0.781***	0.465*	-0.049
					(0.297)	(0.164)	(0.206)	(0.284)
$\widetilde{log(I+N)}$							0.799***	1.049**
							(0.071)	(0.328)
F-Stat					11.708	22.565	7.749	17.286
Observations	6,310	6,310	6,310	6,310	6,310	6,310	6,310	6,310
	Fixed Effects							
Market	N	Y	N	Y	N	Y	N	Y
Region-Year	N	Y	N	Y	N	Y	N	Y
Industry-Year	N	Y	N	Y	N	Y	N	Y

Note: Market clustered standard errors in parenthesis. All regressions include year, region and industry fixed effects. F-Stat is the Kleibergen and Paap (2006) reduced rank statistic. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The effect of immigrants on output that I estimate in table 3.2 is robust to changes in employment measures. In table 3.3 I replicate the estimates in table 3.2 measuring employment using hours of work and weighting head-counts with individual specific predicted wages. The use of predicted wages to account for underlying differences in productive characteristics is standard in the literature (see Acemoglu and Autor 2011; Autor et al. 2008). To create the weights, I use British nationals in the first quarter of 2012-2014 LFS to estimate a wage equation on a quadratic of age and years of education. Then, for every individual in the LFS, I produce fitted wages and use these to weight every individual observation when aggregating employment head-counts.

With predicted wages the estimated effect is at a 5% increase on output per worker per ten percentage points increase on immigrant labour relative supply. With total hours of work the effect increases to 9.5% per ten percentage points increase on immigrant relative supply. These changes can be due to changes in scaling. Particularly, one will expect these movements on the size of the estimated effect if immigrants have more desirable skills, as measured by fitted wages, and tend to work less hours. This is because, in such case, any given increase on the head-count measure translates into a larger (smaller) increase on the predicted wage weighted (hours) measure. Plotting the three measures (figure 3.7) this is exactly what one observes. Immigrants tend to have more productive characteristics and work less hours.[15]

Table 3.3: Immigrant-Native Mix and Output
(*Alternative Employment Measures*)

	OLS				IV			
	Wage Weights		Hours		Wage Weights		Hours	
I/N	-0.043	-0.066	0.034	0.023	0.798**	0.544*	1.185**	0.954*
	(0.050)	(0.044)	(0.071)	(0.065)	(0.297)	(0.264)	(0.457)	(0.389)
$log(N+I)$		0.127***		0.132***		0.301		0.171
		(0.032)		(0.029)		(0.215)		(0.123)
F-Stat					24.769	16.623	21.971	14.580
Observations	6,310	6,310	6,310	6,310	6,310	6,310	6,310	6,310
	Fixed Effects							
Market	Y	Y	Y	Y	Y	Y	Y	Y
Region-Year	Y	Y	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y	Y	Y

Note: Market clustered standard errors in parenthesis. All regressions include year, region and industry fixed effects. F-Stat is the Kleibergen and Paap (2006) reduced rank statistic. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, Jaeger, Ruist, et al. (2018) show that shift-share instruments, as the one I use, most times estimate an effect that is a weighted combination of short- and long-run effects. This is because the shift-share is not only a good predictor for current immigrant stocks but also for past ones. Jaeger, Ruist, et al. (2018) propose including lags of the immigrant stock instrumented with lags of the instrument to separate these two effects. In table 3.9, I display IV estimates including one lag of the immigrant stock. Without market, year-region and year-industry fixed effects, my multiple instrument estimate shows that increasing the relative supply of current immigrant labour increases output by 1.96%. The same increase on the lag produces a .894% increase on output that, however, is not statistically significant. Once I introduce market, region-year and industry-year fixed effects, the effect of current immigrants reduces to .852% and is not statistically significant. Also, with the full set of fixed effects the first stage is substantially weaker. The Kleibergen and Paap (2006) reduced rank statistic is at 2.89, well below typical rules of thumb. Therefore, my estimates in table 3.9 suggest that the effect I capture is not only driven by long-run effects. However, they are not completely conclusive.

[15] Differences on hours could be due to miss-report of hours. This is because I measure total hours instead of ours per-worker, if a larger proportion of immigrants tend to not report hours my hour measure will underestimate the total work time they contribute.

3.4 Decomposition of the Reduced-Form Effect

3.4.1 Short-run decomposition

Given the positive effect of immigrants on output that I estimate in table 3.2, the question is what is the mechanism behind it. In the short run, the effect of immigrants on output is given by changes in the immigrant-native labour mix and the extend of the effect depends on the technology used to combine both types of labour and the effect of immigrants on native labour.

To explore how the reduced form effect in (3.2) depends on the production technology, let output be produced by combining immigrant and native labour using the following Constant Elasticity of Substitution Environment (CES)[16]

$$y_{mt} = \frac{1}{\rho} \log \left(\Delta + (1-\Delta) \left(\frac{I_{mt}}{N_{mt}} \right)^{\rho} \right) + \log(N_{mt}) + \xi_{mt} \tag{3.2}$$

This CES production function allows for relatively more productive native or immigrant labour ($\Delta \neq .5$), forces a constant elasticity of substitution taking the form $1/(1-\rho)$ and is subject to time-market specific shocks ξ. Assuming that this CES production function provides a good description of the production technology, one can explore what is included inside the reduced form estimate in table 3.2. Abstracting from endogeneity problems and taking a second order Taylor expansion,[17] the estimate obtained from regressing output generated according to (3.2) on immigrant relative supply has the following decomposition

$$\hat{\beta} \approx \underbrace{\frac{(1-\Delta)\,\mathbb{E}\left[I/N\right]^{\rho-1}}{(1-\Delta)\,\mathbb{E}\left[I/N\right]^{\rho} + \Delta}}_{I/N \text{ Output Semi-Elasticity}} + \underbrace{\frac{Cov\left[\log(N), I/N\right]}{Var\left[I/N\right]}}_{\text{Effect of Immigrants on Native Labour}} \tag{3.3}$$

Equation (3.3) highlights two aspects of the reduced form estimate. First, it is inherently local as it depends on the distribution of labour inputs and this changes across time and location (see figure 3.1). Second, the extend of the effect of immigrants on output depends on the elasticity of substitution between immigrant and native labour (determined by ρ) and the relative productivity of immigrant and native labour (Δ). Given that immigrant labour is relatively scarce, in the sense of $\mathbb{E}\left[I/N\right] < 1$, the effect of immigration on output becomes more positive as the elasticity of substitution decreases. Moreover, independently of immigrant-native mix, the effect of immigration on output increases if the economy becomes relatively more efficient on using immigrant labour, i.e. if Δ decreases.

[16] The CES framework has been used extensively in the migration literature (e.g. Manacorda et al. 2012; Ottaviano and Peri 2012; Peri 2012).

[17] The approximation is similar for the IV estimator, $\beta_{IV} \approx \sum_s y_s \frac{Cov(s,z)}{Cov(I/N,z)}$ where y_s is the output semi-elasticity of input s. This approximation can be extended to account for additional controls by using the Frisch-Waugh theorem.

In the short run, when only labour varies, immigration has a negative effect on output if immigrants displace a sufficiently large amount of native labour.[18] The second summand in equation (3.3) is actually the OLS estimate from regressing (log-)native employment on the relative supply of immigrant labour. If one runs that regression, obtains the estimates displayed in table 3.4. OLS estimates of (log-)native employment on immigrant-native ratios show a negative effect of immigration on native employment that would signal towards immigrants crowding-out native employment. However, the IV estimator returns a positive effect indicating that immigrants promote native employment. Thus, comparison of OLS and IV estimates suggest that immigrants sort into markets where native employment is decreasing, but they do not crowd-out native employment.[19]

3.4.2 Long-run decomposition

Once one moves from the short-run, immigrants may shape other inputs. For example, a larger immigrant labour force may modify incentives to accumulate human capital by altering returns to skill (Llull 2017) and, therefore, modify the underlying distribution of native skill.[20] Immigrants can also promote development of new technologies that complement immigrant labour (Peri 2012; Peri, Shih, et al. 2015) or allow faster capital accumulation through reductions in production costs (Ottaviano, Peri, and Wright 2018). In such case, the positive effect of immigrants on output per worker studied in (3.3) will also be a function of the effect of immigrants on native skill mix, technologies and capital. To see this, let me augment the production function in (3.2) and write[21]

$$\begin{aligned} y_{mt} =& \frac{(1-\alpha)}{\rho} \log\left(\Delta_{mt}\left(\frac{N^L_{mt}}{N_{mt}} + \phi\frac{N^H_{mt}}{N_{mt}}\right) + (1-\Delta_{mt})\left(\frac{I_{mt}}{N_{mt}}\right)^{\rho}\right) \\ &+ (1-\alpha)\log(N_{mt}) + \alpha\log(K_{mt}) + \xi_{mt} \end{aligned} \tag{3.4}$$

Where N^H and N^L are high and low skilled native employment, K are capital stocks, ϕ is the relative marginal product of skilled labour and native augmenting technologies (Δ) may change over time and across markets. Inside this framework the Taylor approximation of the reduced form effect takes the form

[18] The same holds for output per worker, although then the negative effect would work through a sufficient negative effect on the share of native labour. When the left hand side is output per worker, the second term in (3.3) is $Cov(log(N/(I+N)), I/N)/Var(I/N)$.

[19] Sorting could be produced by immigrants entering markets with stagnated or decreasing native labour, but also by immigrants staying in markets that natives are leaving.

[20] See also the literature on the effect of immigrants on native wages (e.g. Dustmann, Fabbri, et al. 2005; Dustmann, Frattini, et al. 2013; Manacorda et al. 2012).

[21] For the sake of my exposition I do not include the skill differential ϕ for immigrants. But later on, when I estimate the production function, I redefine $I = I^L + \phi I^H$.

$$
\begin{aligned}
\hat{\beta} \approx & \frac{\partial y}{\partial I/N} + \frac{\partial y}{\partial N^H/N} \underbrace{\frac{Cov[N^H/N, I/N]}{Var[I/N]}}_{\text{Effect on Native Skill Mix}} \\
& + (1-\alpha) \underbrace{\frac{Cov[\log(N), I/N]}{Var[I/N]}}_{\text{Effect on Native Employment}} + \alpha \underbrace{\frac{Cov[\log(K), I/N]}{Var[I/N]}}_{\text{Effect on Capital Stocks}} \\
& + \frac{\partial y}{\partial \Delta} \underbrace{\frac{Cov[\Delta, I/N]}{Var[I/N]}}_{\text{Effect on Native Augmenting Tech.}}
\end{aligned}
\tag{3.5}
$$

The second to third elements in (3.5) are the effects of immigrants on output through changes in the skill mix of native labour, total native labour supply and capital stocks. The size of the skill mix effect depends on the parameters of the production function, the distribution of labour inputs and technologies and the effect of immigration on native mix ($Cov(N^H/N)/Var(I/N)$). Obtaining the sign and magnitude of the effect of immigrants on the skill mix of natives, $Cov(N^H/N)/Var(I/N)$, is straight forward. In table 3.4, I provide both OLS and IV estimates for this effect. The IV estimate shows that a ten percentage points increase on the relative supply of immigrant labour produces a 1% increase on the skill supply of natives. This positive effect of immigrants on native skill supply will have a positive effect on output if high skilled natives are relatively more productive than low skilled natives, i.e. $\phi > 1$. Assuming that natives are paid their marginal contribution to production one can recover ϕ from relative wages of natives in high and low skilled occupations.[22] In the sample, the average relative wage is at $\phi = 2.84$.[23] Thus immigrants rise output and the productivity of native labour by increasing the skill supply of native labour.

The positive effect of immigrants on the skill mix of natives is consistent with evidence in Dustmann, Frattini, et al. (2013) showing that immigrants in the UK reduce the wages of natives at the bottom of the wage distribution and have a positive effect for natives at the top. Given that unskilled workers tend to be at the lower tail of the distribution, the evidence in Dustmann, Frattini, et al. (2013) suggest that immigrants may reshape the relative returns to high skilled jobs, producing incentives for natives to accumulate skills that allow them to move into these high skilled jobs.[24]

[22]The assumption of perfect competition in the labour market is, somehow, more credible for native than for immigrant labour. This is because there is abundant literature (e.g Aydemir and Skuterud 2008; Eckstein and Weiss 2004; Izquierdo et al. 2009) showing that immigrants earn lower wages that comparable natives (see also Dustmann and Preston 2012; Dustmann, Schönberg, et al. 2016).

[23]This is the average across markets and years of the average wage offered in skilled occupations over the average wage in unskilled occupations. I measure wages from gross weekly earnings information contained in the Labour Force Survey.

[24]For the US, Llull (2017) shows that natives respond to immigration by adjusting their educational and occupational choices. Hunt (2017) provides evidence showing that immigrants entering the US have a positive effect on educational outcomes of native children, where the effect is stronger for children with low-educated parents.

The effect of immigrants on capital stocks, $Cov\left(\log(K), I/N\right)/Var\left(I/N\right)$, can be easily estimated, too. In the last four columns of table 3.4, the IV estimate shows that increasing the relative supply of immigrant labour by ten percentage points produces a 5% increase in capital stocks and a 4% increase in per worker capital stocks. Thus showing that immigrants also contribute to production through the accumulation of capital.

Table 3.4: Immigrant-Native Mix and Input Supply

	$log(N)$		H/N		$log(K)$			
	OLS	IV	OLS	IV	OLS	OLS	IV	IV
I/N	-0.403***	0.349^{+}	0.049^{+}	0.116^{+}	-0.002	-0.009	0.522**	0.395**
	(0.044)	(0.201)	(0.026)	(0.061)	(0.019)	(0.019)	(0.174)	(0.137)
$log(N+I)$						0.045**		0.129
						(0.014)		(0.098)
F-Stat		22.565		22.565			31.362	22.720
Observations	6,310	6,310	6,310	6,310	5,567	5,567	5,567	5,567
				Fixed Effects				
Market	Y	Y	Y	Y	Y	Y	Y	Y
Region-Year	Y	Y	Y	Y	Y	Y	Y	Y
Industry-Year	Y	Y	Y	Y	Y	Y	Y	Y

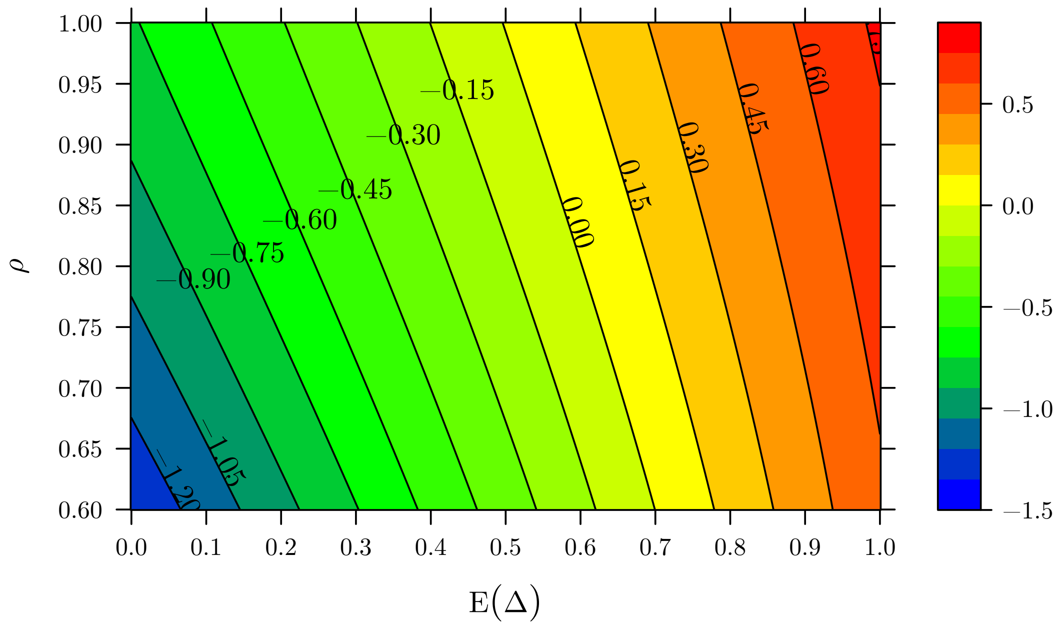
Note: Market clustered standard errors in parenthesis. F-Stat is the Kleibergen and Paap (2006) reduced rank statistic. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Because Δ is not directly observable, testing the effect of immigrants on the accumulation of native augmenting technologies is more convoluted. However, using the decomposition in (3.5) and estimates in table 3.2 it is possible to obtain an approximation of the effect of immigrants on native augmenting technologies. In figure 3.2, I display an approximation for $Cov(\Delta, I/N)/Var(I/N)$ obtained from the expansion in (3.5) at a set of values for ρ and $\mathbb{E}[\Delta]$ setting ϕ at the average high skilled relative wage of natives.

It follows from figure 3.2 that, if native and immigrant labour are gross substitutes, say at $\rho = .6$, rising the share of immigrant labour has a negative impact on native augmenting technologies if, with equal labour shares, natives are, at most, 133% more productive than immigrants, i.e. $\mathbb{E}[\Delta] \leq .7$. On the other hand, if immigrant and native labour are perfect substitutes, i.e. $\rho = 1$, increasing the share of immigrant labour has a negative effect on native augmenting technologies if, with equal labour shares, natives are, at most, as productive as immigrants, i.e. $\mathbb{E}[\Delta] \leq .5$.

A negative effect of immigrants on technologies that complement native labour is consistent with the endogenous technological change framework developed by Acemoglu (1998, 2002). Acemoglu (2002) shows that, when native and immigrant labour are gross substitutes, i.e. $\rho > 0$, increasing the share of immigrant labour produces a negative effect on the development of technologies that complement immigrant labour.[25] Where the effect is driven by the market effect derived from the larger number of workers that can benefit from using technologies that complement immigrant labour.

Figure 3.2: Immigrant-Technology Regression Coefficient (at $\phi = 2.84$)



The decompositions of the reduced form estimator highlights that one can gain further understanding about the effect of immigrants on output and labour productivity by estimating the parameters governing the production function. The existing literature estimating parameters of production functions with native and immigrant labour (Card 2009; Manacorda et al. 2012; Ottaviano and Peri 2012) use relative wages to identify production function parameters. This has the advantage of delivering linear on parameters estimating equations. However, a vast literature reporting wage gaps between immigrants and comparable natives (e.g. Aydemir and Skuterud 2008; Chiswick 1978; Eckstein and Weiss 2004; Izquierdo et al. 2009) suggest that immigrants my not be paid their marginal contribution to production. In turn, this implies that relative wages do not directly reflect the parameters of the production function. Thus, instead of pursuing estimation of production function parameters through immigrants' relative wages, my aim is to learn something about the production function directly from production data. For this let me rewrite the production function in (3.4) as

[25] Acemoglu (1998, 2002) frames it as any two inputs.

$$
\begin{aligned}
y_{mt} - \log(\tilde{N}_{mt}) =& \frac{(1-\alpha)}{\rho} \log\left(\Delta(x_{mt}) + (1-\Delta(x_{mt}))x_{mt}^{\rho}\right) + \alpha \log(K_{tm}/\tilde{N}_{mt}) \\
& + \xi_{mt} \qquad (3.6) \\
=& g(x_{mt}) + \alpha \log(K_{tm}/\tilde{N}_{mt}) + \xi_{mt}
\end{aligned}
$$

where $\tilde{N} = N_L + \phi N_H$, $x = (I_L + \phi I_H)/\tilde{N}$ and I assume that ϕ is known or can be recovered on a first step.[26] Furthermore, that native augmenting technologies are determined by the relative supply of immigrant labour, i.e. $\Delta_{mt} \equiv \Delta(x_{mt})$, follows from the literature on directed technological change (see Acemoglu 1998, 2002).

The reformulation of the production function in (3.6) has the convenient property that it produces a quasilinear function with an unknown function in a scalar variable. I exploit this to produce two estimates. On one hand, I estimate g imposing the CES structure and assuming exogenous native augmenting technologies, i.e. $\Delta(x)$ is a constant. On the other, I estimate g using a linear equation with a cubic polynomial on the relative supply of immigrant labour. This can be interpreted as a third order Taylor series approximation for g, assuming that $\Delta(x)$ is a smooth function.[27]

As with estimation of the reduced form effect, endogeneity is still a concern when estimating g. Here I exploit the same variation on past settlement of immigrants across markets that I use for the reduced form. For the cubic polynomial approximation I use an IV estimator using powers of the shift-share instrument. In estimation of g imposing the CES structure and $\partial\Delta(x)/\partial x = 0$, I use a control function (see Terza et al. 2008; Wooldridge 2015) non-linear least squares estimator.[28]

In table 3.5, I display estimates of production function parameters and marginal relative contribution of immigrant labour, i.e.

$$
M\hat{R}TS = \frac{\sum_{m,t} \frac{(1-\Delta)}{\Delta} \left(\frac{I_{mt}}{N_{mt}}\right)^{\rho-1}}{\sum_{m,t} 1} \qquad (3.7)
$$

The non-linear least squares elasticity estimate is between .62-.84, depending on the value of the relative marginal product of skilled and unskilled labour. At .84 this estimate is close to Manacorda et al. (2012) estimate at $\rho \approx .872$ and smaller than US estimates in Card (2009) and Ottaviano and Peri (2012).[29] However, my confidence intervals are broad and I cannot reject a null of perfect substitutes, i.e. $\rho = 1$. This

[26] For example, from relative wages of high and low skilled natives.

[27] The smoothness assumption of g could be relaxed by using a non-parametric estimator (Horowitz 2011). However, given that the validity of g relies on Acemoglu (1998, 2002) framework and this produces a smooth function I will not gain much robustness.

[28] In a linear on parameters set-up, the control function approach returns estimates that are analogous to usual two-stage least squares IV (Terza et al. 2008). However, in non-linear set ups the two-stage least squares is generally not available.

[29] Sample sizes in table 3.5 are smaller than in previous tables because I must exclude markets on which no immigrant worker is observed and capital stocks are only available from 2000 onwards.

increase on the dispersion of the estimate, as compared with previous estimates in Manacorda et al. (2012) where the 95% confidence interval is $[.781, .962]$, can be given by different sources of identifying variation. Manacorda et al. (2012) exploit variation within education-age cells assuming that this is exogenous. I exploit variation within markets using an instrumental variables strategy and this produces noisier estimates.

The relative productivity of native labour, Δ, is estimated above .5 for all specifications but the one where I allow for free ψ. In this specification I set ψ to be the relative wage of high skilled natives in the market in the given year. That $\Delta > .5$ means that, when immigrant relative supply is equal to one, natives are more productive than immigrants. However, in none of my specifications I can reject the null of equality of productivity. Moreover, the combination of gross substitution between immigrant and native labour, differences in productivity and relative scarcity of immigrant labour leads to large differences in marginal products. However, again, these differences are not statistically significant. I can not reject that the average immigrant relative marginal contribution is equal to one.[30]

Table 3.5: Production Function Estimates
Control Function

	$\phi = 1$	$\phi = 2.84$	Free ϕ
ρ	0.687	0.617	0.836
	[0.306 , 1.099]	[0.416 , 2.724]	[0.312 , 1.796]
Δ	0.552	0.610	0.389
	[0.230 , 0.910]	[0.087 , 1.000]	[0.074 , 0.915]
α	0.304	0.345	0.337
	[0.232 , 0.429]	[0.249 , 0.465]	[0.279 , 0.479]
	Immigrant-Native Relative Marginal Contribution		
	1.828	1.846	2.322
	[0.609 , 6.301]	[0.000 , 5.193]	[0.269 , 8.121]
Observations	5100	5100	5026

Note: 95% Bootstrap confidence intervals, 500 replications. Δ estimated through a sigmoid transformation. All estimates include year, region and industry fixed effects. The last column sets ϕ to be the relative native skill wage in every market. Diffrences in sample sizes are due to missing relative wages.

Using my non-linear least squares estimates I can compute the marginal effect of the relative supply of immigrant labour. In figure 3.3, I display this marginal effect holding total labour and capital constant, solid line, and allowing capital to move with immigrant relative supply using estimates from table 3.4, discontinuous

[30] The parameter that is most tightly estimated is the Cobb-Douglas output elasticity of capital, α, that is estimated between .3-.34. This is larger than the IV estimate from the cubic polynomial, 0.15-0.31, and close to the value of .33 typically used in the growth accounting literature and in Peri (2012) decomposition of the effect of immigrants on labour productivity in the US.

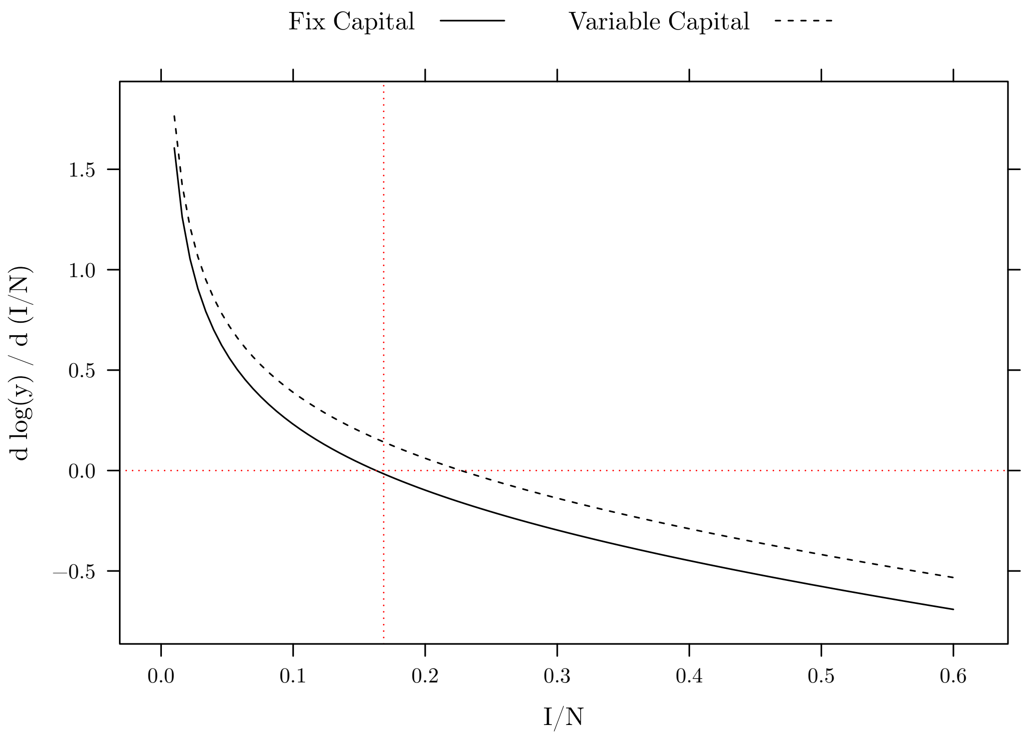
line. At the average immigrant supply, vertical red line, the non-linear least squares marginal effect of immigrants is, roughly, zero. This is holding total employment constant, thus, it will imply a zero effect on output per worker. If I include the effect on capital stocks, the marginal effect from the non-linear least squares evaluated at the average immigrant relative supply is positive. This suggests that, if native augmenting technologies are exogenous, the positive effect of immigrants on output per worker works through the effect in other inputs.

To produce further evidence on whether this is the case, in table 3.10, I estimate the reduced form effect controlling for capital stocks and native skill supply. The baseline effect, first column of table 3.10, is the one estimated earlier in table 3.2. A ten percentage points increase on the relative supply of immigrant labour increases output per worker by 6.7%. Evaluated at the average this is a £5,610 increase on output per worker. Controlling for capital stocks reduces the effect to 5% per ten percentage points increase on the relative supply of immigrant labour. Evaluated at the average this is a £4,270 increase per worker. Including additional controls for the share of native employment in high skilled occupations marginally increases the effect of immigrants on output per worker. This is due, surprisingly, to a negative correlation between output per worker and the share of skilled native workers. In any case, after controlling for other inputs there is, still, a positive and significant effect of immigrants on output per worker. Specifically, around 76% of the baseline effect remains. This further suggest that increasing the relative supply of immigrant labour has a positive impact on the development of technologies that complement immigrant workers.

Finally, in figure 3.8, I compare the marginal effect of immigrants, holding total labour constant, from the non-linear least squares with the cubic polynomial approximation. The cubic polynomial approximation produces marginal effects that are always larger than the non-linear least squares. Given that the polynomial approximation does not restrict native augmenting technologies to be constant,[31] this difference between the polynomial approximation and the non-linear least squares further suggests that immigrants have a positive impact on the development of technologies that complement them. However, this should be taken with caution. Bootstrap confidence intervals are wide and I can only reject the null of equality over a small set of points.

[31]And assuming that the CES production function is otherwise a good model for the actual production technology.

Figure 3.3: Relative Supply Immigrant Labour Marginal Product (at $\phi = 1$)



3.5 Conclusions

With immigrant labour rapidly growing over the last 15 years immigration has become central for British policy makers. I contribute to the current discussion of the effect of immigrants in the British economy by estimating the effect of immigrants on production and labour productivity.

Using data from the UK Labour Force Survey and regional accounts I show that rising the relative supply of immigrant labour has a positive effect on output and output per worker. To produce this estimate I exploit variation on past settlement of immigrants across industries and regions in an instrumental variables estimator with a comprehensive set of year, industry and region fixed effects. My baseline estimates show that increasing the relative supply of immigrant labour by ten percentage points increases output per worker by 6.7%. When interpreted at average levels this implies an additional £5,610 per worker.

To explore the composition of the reduced form effect I introduce a Constant Elasticity of Substitution (CES) production function. I show that the reduced form not only depends on the technology been used but also on the joint distribution of immigrant labour and other inputs, including the development of new technologies. This highlights that reduced form estimates of the effect of immigrants on output are inherently heterogeneous and warns against extrapolating results across periods or countries. In this direction, I show that part of the positive effect of immigrants on output works through positive effects on the stock of capital, native employment and the skill distribution of natives. Increasing immigrant relative labour supply by ten percentage points increase native employment a 3.5%, the relative supply of high skilled native labour by 1% and capital stocks per worker by 4%.

Using the decomposition of the reduced form estimate I show that, for a wide set of parameter values, the evidence I present is consistent with immigrants triggering accumulation of technologies that complement immigrant labour. This is in line with the directed technological change literature (see Acemoglu 1998, 2002). Furthermore, comparing production function and reduced form estimates I produce further, suggestive, evidence on the effect of immigrants on the development of native augmenting technologies. Specifically, estimates of the CES production function assuming that native augmenting technologies are constant, produce a marginal effect of immigrants on output per worker that is only positive when I allow for immigrant effects on capital stocks. This is not consistent with reduced form estimates where I control for capital stocks, which display a positive and significant effect. Specifically, a ten percentage points increase on immigrant labour supply rises output per worker by 5%, i.e. an additional £4,270 per worker. I interpret this as a signal of misspecification of the production function. Particularly, against the assumption that native augmenting technologies are constant.

To produce further evidence, I estimate a cubic polynomial approximation of the production function that is consistent with the CES function with endogenous technological change á la Acemoglu (1998, 2002). Marginal effects computed with this specification are always larger than marginal effects from the CES specification holding technologies constant. This further suggest that immigrants trigger the development of technologies that complement them. However, this piece of evidence should be taken with caution as I only reject the null of equality of marginal effects over a small set of points.

3.6 Appendix

3.6.1 Classifications

Table 3.6: Industries

Code	Code Number	Name
A	1	Agriculture, forestry and fishing
B	2	Mining and quarrying
CA	3	Food products, beverages and tobacco
CB	4	Textiles, wearing apparel and leather products
CC	5	Wood and paper products and printing
CD	6	Coke and refined petroleum products
CE	7	Chemicals and chemical products
CF	8	Basic pharmaceutical products and preparations
CG	9	Rubber and plastic products
CH	10	Basic metals and metal products
CI	11	Computer, electronic and optical products
CJ	12	Electrical equipment
CK	13	Machinery and equipment not elsewhere classified
CL	14	Transport equipment
CM	15	Other manufacturing and repair
D	16	Electricity, gas, steam and air-conditioning supply
E	17	Water supply; sewerage and waste management
F	18	Construction
G	19	Wholesale and retail trade; repair of motor vehicles
H	20	Transportation and storage
I	21	Accommodation and food service activities
J	22	Information and communication
K	23	Financial and insurance activities
L	24	Real estate activities
M	25	Professional, scientific and technical activities
N	26	Administrative and support service activities
O	27	Public administration and defence; compulsory social security
P	28	Education
Q	29	Human health and social work activities
R	30	Arts, entertainment and recreation
S	31	Other service activities
T	32	Activities of households
U	33	Activities of extraterritorial organisations and bodies

Table 3.7: Understanding Society Industries

Code	Name
Agriculture and Forestry	1
Fisheries	2
Energy/Water	3
Mining	4
Chemicals	5
Synthetics	6
Earth/Clay/Ston	7
Iron/Steel	8
Mechanical Eng.	9
Electrical Eng	10
Wood/Paper/Prit	11
Clothing/Text.	12
Food Industry	13
Construction	14
Constr. Relate	15
Wholesale	16
Trading Agents	17
Retail	18
Train System	19
Communication/Entertainment	20
Other Trans.	21
Financial Inst	22
Insurance	23
Restaurants	24
Service Indust	25
Trash Removal	26
Educ./Sport	27
Health Service	28
Legal Services	29
Other Services	30
Volunt./Church	31
Priv. Househld	32
Public Admin.	33
Social Sec.	34

3.6.2 Tables

Table 3.8: Market Movers

	UK Born		**Foreign Born**		
	Market Stayer	Market Mover	Market Stayer	Market Mover	Total
	0.924	0.076	0.905	0.095	
	{11821}	{970}	{2042}	{215}	15048
	Conditional on Moving		Conditional on Moving		
	Same Industry	Different Industry	Same Industry	Different Industry	
Region Stayer	0.000	0.891	0.000	0.870	
	{0}	{864}	{0}	{187}	1051
Region Mover	0.068	0.041	0.093	0.037	
	{66}	{40}	{20}	{8}	134
Total	11887	1874	2062	410	16233

Note: Data from Understanding Society waves 3 and 4 (2011-2014). All individuals with consecutive interviews, known industry of work, region of residence and country of birth. Due to to data constraints the industrial classification is slightly different of the one I use in the rest of the paper, see table 3.7

Table 3.9: Output per Worker
IV estimates, lagged values

	IV	
	(1)	(2)
I/N	1.976*	0.852
	(0.904)	(0.566)
I/N(lag)	0.894	0.882
	(0.612)	(0.469)
	First-Stage (I/N)	
$\widetilde{I/N}$	1.227***	1.138***
	(0.199)	(0.137)
$\widetilde{I/N}$ (lag)	-0.282	-0.614***
	(0.190)	(0.168)
	First-Stage (I/N (lag))	
$\widetilde{I/N}$	-0.493	-0.660*
	(0.274)	(0.273)
$\widetilde{I/N}(lag)$	1.546***	1.274***
	(0.237)	(0.180)
F-Stat	33.249	2.891
Observations	5,932	5,932
	Fixed Effects	
Market	N	Y
Region-Year	N	Y
Industry-Year	N	Y

Note: Market clustered standard errors in parenthesis. All regressions include year, region and industry fixed effects. F-Stat is the Kleibergen and Paap (2006) reduced rank statistic. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.10: Output per Worker
Input Controls, IV estimates

	Baseline	With Capital Stocks	With Capital Stocks and Native Skill
I/N	0.670*	0.509*	0.533*
	(0.320)	(0.257)	(0.261)
$log(N+I)$	-0.676**	-0.730***	-0.739***
	(0.221)	(0.185)	(0.183)
F-Stat	17.286	20.556	21.355
Observations	6,310	5,567	5,567
		Fixed Effects	
Market	Y	Y	Y
Region-Year	Y	Y	Y
Industry-Year	Y	Y	Y

Note: Market clustered standard errors in parenthesis. F-Stat is the Kleibergen and Paap (2006) reduced rank statistic. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.6.3 Plots

Figure 3.4: Output per Worker Evolution
Regions

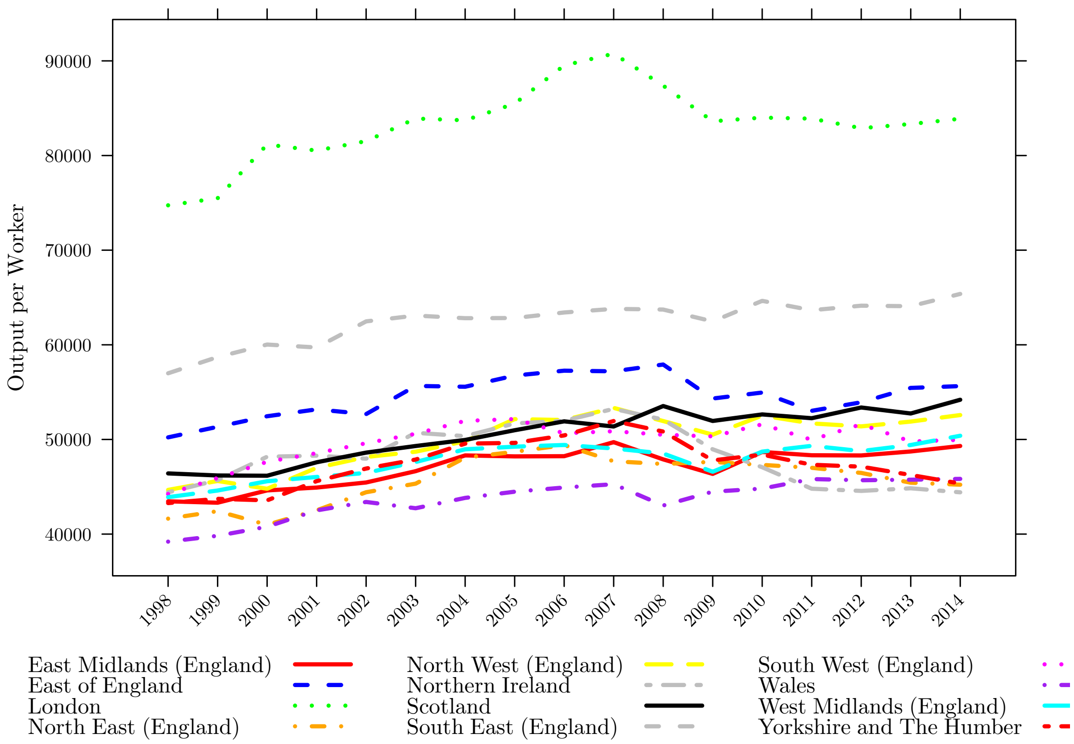


Figure 3.5: Workforce Jobs and LFS: Measures of Employment

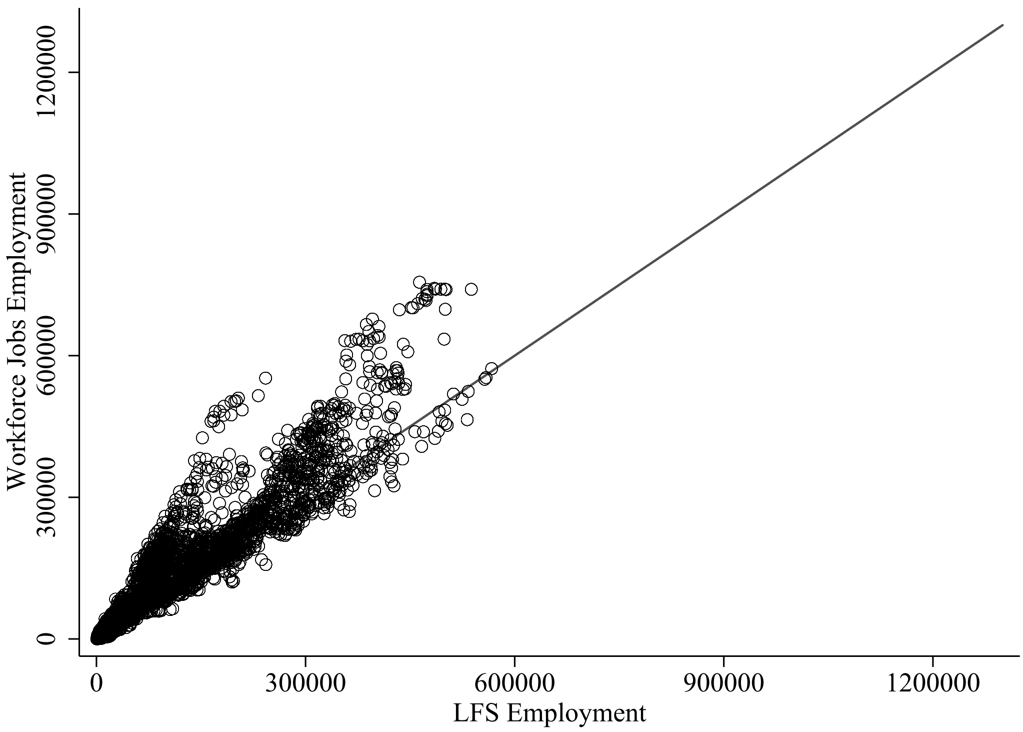


Figure 3.6: Immigrant Stock Evolution

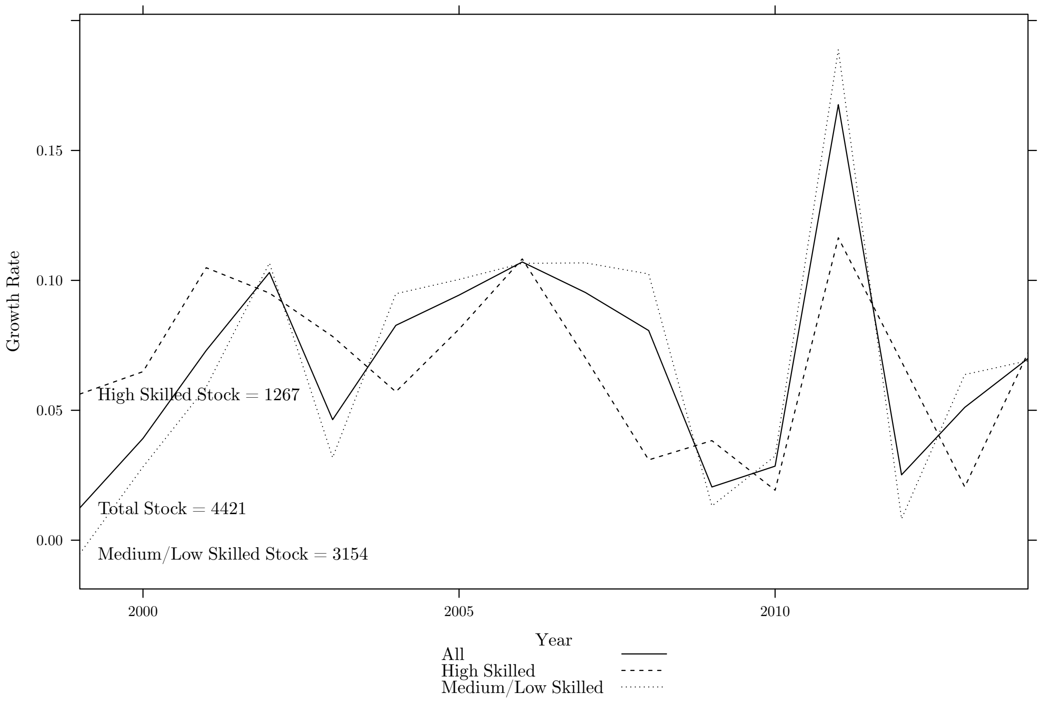


Figure 3.7: Immigrant-Native Ratios
(*Comparison of Measures*)

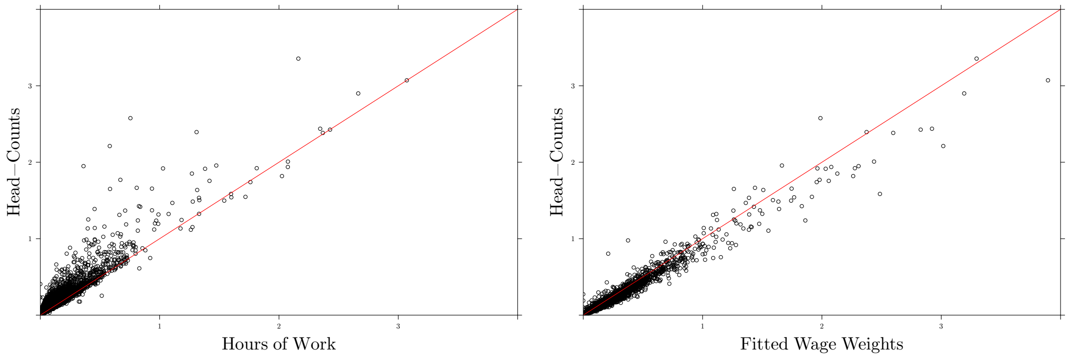
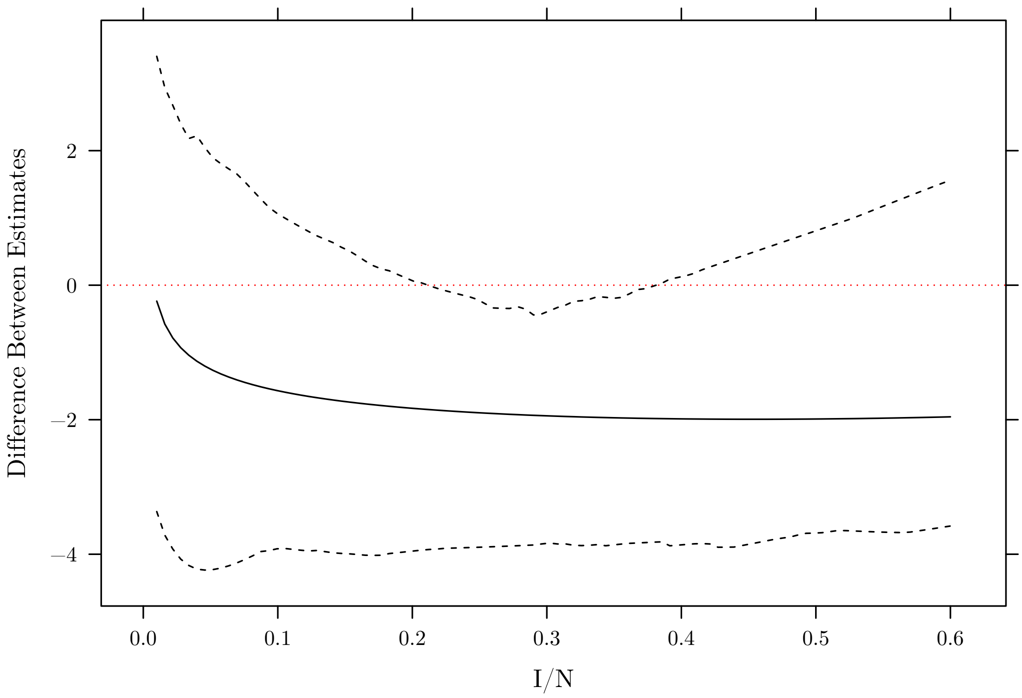


Figure 3.8: Difference in Immigrant Labour Marginal Product
(at $\phi = 1$)



Note: The solid line is the difference between the marginal effect from non-linear least squares and cubic polynomial. Dashed lines are 95% confidence levels from 500 bootstrap replications. For each replication I compute the marginal effect of both estimators and its difference.

3.6.4 Data Preparation: Changes in Industries and Occupations Coding

In 1991 the new SOC90 coding scheme was introduced in the LFS. In this year there are two variables reporting occupation, one is *KOS* that was used until 1991 and the other is *SOC/SOCMAIN* that uses the SOC90 coding scheme and is in place until 2001. Thanks to this, I can create a mapping between the old coding and the new by cross tabulating the two variables using LFS 1991. As I am using aggregate quantities I choose to employ a probability mapping between schemes. Furthermore, some translations were imputed manually.[32]

In 2001 there is a new change and occupations were coded under SOC00. During this year both the old *SOCMAIN* variable and the new *SOC2KM* were coded. However, only the latter is provided in the end of user quarterly datasets available at the UK Data Archive, with no occupational information provided at all in the first quarter. To overcome this issue, I use the last quarter of LFS 2000 and the second of 2001. Thanks to the panel dimension of LFS I can match observations that belong to the same individual in both quarters, and identify those who have not changed jobs. Then I use these observations to construct the mapping between occupational codes. While creating this map I have imputed manually some one-to-one correspondences.[33]

Finally, in the first quarter of 2011 there is a new change and occupations are now coded following SOC10. During this year in the end of user access datasets from the UK Data Archive there are available two occupational variables, one codes under SOC00, *SC102KM*, and the other under SOC10, *SOC10M*. Allowing me to create a mapping.

With the three maps at hand I translate everything into SOC10 coding and then use MAC's NQF classification[34] to put workers in one of the 4 different groups, ranking 2, 3, 4 and 6. SOC 2010 occupations: (1171) Officers in armed forces, (2444) Clergy, (3311) NCOs and other ranks, (3314) Prison service officers (below principal officer), (3441) Sports players, and (3442) Sports coaches, instructors and officials; are given no classification as these are not eligible for Tier 2 visas.

For industries there is a change between codes for the base years 1985-1990 and my first year of study, 1998; and between 2008 and 2009. I use proportional mappings from ONS and Dr. Jennifer Smith,[35] to code pre 2008 LFS figures into SIC07. Due to availability constraints the matching for years 1985-1990 is performed using two digits industries while the matching for 1998-2008 is done using four digits.

[32] See the Do-files.

[33] See the Do-files.

[34] See https://www.gov.uk/guidance/immigration-rules/immigration-rules-appendix-j-codes-of-practice-for-skilled-work

[35] Available at http://www2.warwick.ac.uk/fac/soc/economics/staff/jcsmith/sicmapping/resources/proportional/.

Table 3.11: NQF Classification and Major Occupation

	NQF				
	2	3	4	6	Total
ADMINISTRATIVE AND SECRETARIAL OCCUPATIONS	**0.720**	0.240	0.040	0.000	
	{18}	{6}	{1}	{0}	25
ASSOCIATE PROFESSIONAL AND TECHNICAL OCCUPATIONS	0.033	**0.443**	0.377	0.148	
	{2}	**{27}**	{23}	{9}	61
CARING, LEISURE AND OTHER SERVICE OCCUPATIONS	**0.615**	0.385	0.000	0.000	
	{16}	{10}	{0}	{0}	26
ELEMENTARY OCCUPATIONS	**0.931**	0.069	0.000	0.000	
	{27}	{2}	{0}	{0}	29
MANAGERS, DIRECTORS AND SENIOR OFFICIALS	0.000	0.286	0.200	**0.514**	
	{0}	{10}	{7}	**{18}**	35
PROCESS, PLANT AND MACHINE OPERATIVES	**0.905**	0.095	0.000	0.000	
	{38}	{4}	{0}	{0}	42
PROFESSIONAL OCCUPATIONS	0.000	0.000	0.014	**0.986**	
	{0}	{0}	{1}	**{69}**	70
SALES AND CUSTOMER SERVICE OCCUPATIONS	**0.778**	0.167	0.056	0.000	
	{14}	{3}	{1}	{0}	18
SKILLED TRADES OCCUPATIONS	0.175	**0.825**	0.000	0.000	
	{10}	**{47}**	{0}	{0}	57
Total	125	109	33	96	363

Note: Totals and frequencies refer to number of four digit SOC 2010 occupations

Table 3.12: Occupations without NQF Match

Officers in armed forces	1171
Clergy	2444
NCOs and other ranks	3311
Prison service officers (below principal officer)	3314
Sports players	3441
Sports coaches, instructors and officials	3442

Chapter 4

Foreign Peer Effects in Higher Education

Joint work with Greta Morando

Abstract

We estimate the effect of foreign peers on educational and early labour market outcomes of undergraduate students. Using data on the universe of higher education students in England, we estimate causal effects by exploiting cross-cohort variation within major and institution. First, we show that increasing the share of EU students alters the ability composition within peer group. This is due to EU students enlarging the pool of individuals from which universities draw while operating under funding constraints. After controlling for this selection at entry, we show that foreign students have only mild effects on native educational and labour market outcomes and that these are heterogeneous across the distribution of ability. Finally, we show that the strongest effects that we find, which are those on grades, are not consistent with a grading on a curve mechanism. This suggest that foreign students mildly shape the human capital accumulation of natives.

4.1 Introduction

Since Grossman (1982) seminal paper, economists have amply investigated the effect of migrants on labour market outcomes of natives. However, the economic literature has paid much less attention to the effects of migrants on the human capital of natives (e.g. Hunt 2017; Llull 2017). We contribute to narrowing this gap in the literature by studying the causal effect of foreign students on the human capital of native peers.

In particular, we study the effect that foreign peers have on native students' higher education and early labour market outcomes. This is of policy relevance as higher education is a critical point in the human capital accumulation of workers. At this stage, individuals make career decisions in terms of which specialization to follow and which institutions to attend that have long-lasting effects (e.g. Arcidiacono 2004; Belfield et al. 2018; Walker and Zhu 2008, 2018).[1] Besides the institutions they attend, the educational performance of students also impacts their labour market outcomes. For example, Jaeger and Page (1996) find that individuals that hold a diploma earn higher wages than workers with the same years of education but without a diploma. These diploma effects exist all the way from high-school to doctoral degrees and, for those that graduate, the grades they get have a positive relation with labour market performance (Jones and Jackson 1990).[2] In this vein, using data from a British university, Feng and Graetz (2017) provide evidence showing that graduating with a first class has a positive effect on industry and earnings of recent graduates.

Increasing the number of international students may increase competition for native students to enter certain institutions and sharing major with foreign students may change the learning environment. For example, foreign students may change the ability composition of the group with which natives compare themselves, leading to natives changing their perception about their own ability and their effort provision (Elsner and Isphording 2017; Murphy and Weinhardt 2018). Moreover, foreign students may create a more diverse pool of skills that may help natives when engaging in cooperative work.[3] In addition, foreign students may slow down the development of seminars and the skill accumulation of peers due to language barriers (Anelli, Shih, et al. 2017) or increase competition for top grades when grades are awarded based on relative position. All these mechanisms push in opposite directions and make the investigation of foreign peer effects an eminently empirical matter.

Despite the potential economic importance of foreign effects in higher education, this question has not been widely explored in the literature. We provide a complete characterization of the effect of foreign students in higher education: from the effect of foreigners on the composition of natives at entry in higher education to their effect on natives educational performances and longer-term consequence on labour market outcomes.

[1] See also Ehrenberg (2004) review of the literature.
[2] French et al. (2015) provide evidence on the correlation between high school grades and earnings.
[3] Similar to effects of skill mix on firm productivity (Iranzo et al. 2008).

Given the characteristics of the financial system in higher education in the UK we start by providing a theoretical model of how foreign peers affect students composition at entry in higher education. In our context, this is simultaneously determined by two elements which are whether non-natives students are from a country belonging to the EU and the type of university considered. In England, home (i.e. British and EU) and non-EU students face different funding constraints imposed by the funding system. This is because universities are funded by the state through the Higher Education Funding Council for England (HEFCE) that, during our period of study, set caps on the numbers of subsidized students that each university can enrol. These caps were binding as universities have to face monetary fines, offsetting direct gains, if they enrolled students above the cap (see Machin and Murphy 2017). This creates an asymmetry between home and non-EU students, as the later are only subject to visa restrictions and capacity constraints of the institution.[4] We show that this asymmetry in the treatment of international students creates a heterogeneous foreign effect on the enrolment of native students across selective and non-selective universities. Selective universities are able to select the best students and this induces a negative foreign effect on the number of natives enrolled from increasing the number of both EU and non-EU prospective students. Non-selective universities, on the other hand, use non-EU students to fill capacity and this creates an asymmetry. A larger number of EU students crowds-out native students but a similar increase on non-EU has no effect on natives.

We then empirically test whether our predictions are supported in the data. We show that cross-cohort variation on foreign shares correlates positively with average ability of incoming native students and negatively with the number of enrolled natives. In non-selective universities both effects are driven by EU students. However, in selective universities these effects are driven by both EU and non-EU students. This is exactly what our mechanism predicts.

Without controlling for this selection at entry, one can only estimate an effect that is a combination of selection plus foreign peer effects. This selection effect is different from the self-selection effect that is the main concern in most of the literature (e.g. Carrell, Fullerton, et al. 2009; Sacerdote 2001). Similar to Lavy and Schlosser (2011), we argue that students may know the average number of foreign students in a major, university or even in a major within a university. But are unlikely to know, ex-ante, the number of students within a cohort in a university and major. In this situation, exploiting variation across cohorts within university-major deals with student self-selection. However, universities selecting students on the basis of their ability makes cross cohort variation within university and major non quasi-experimental.[5] Hanushek et al. (2003) note the importance of controlling for school selectivity and use a set of fixed effects, including individual fixed effects to control for selectivity. Nonetheless, the inclusion of individual fixed effects may

[4]There have not been any changes on the student visa that we are aware of in the period considered in the paper.

[5]This and capacity constraints explain the effect of EU peers on native average ability.

not be without concerns. This is because individual fixed effects are estimated using within individual variation on outcomes measured after interaction with the peer group. Thus, there is no guarantee that one is not eliminating part of the peer effect. One of the strengths of our analysis is that, thanks to the quality of our data, we are able to control for selection using a measure of ability that pre-dates interaction with peers. This measure is the Universities and Colleges Admissions Service (UCAS) tariff score. The UCAS score is what universities observe in students' UCAS application and the main characteristic that they use when selecting individuals. Furthermore, the UCAS score is composed of students previous educational outcomes and, as such, is pre-determined by the time students enter in contact with the peer group.

We show that conditional on ability, cross-cohort variation on peer-shares is uncorrelated with a comprehensive set of pre-determined individual characteristics that are likely predictors of educational attainment.[6]This provides strong evidence of the share of foreign students to be as good as random, thus supporting the causality of our estimates (Cools et al. 2019; Lavy and Schlosser 2011). Therefore we follow the literature (e.g. Anelli, Shih, et al. 2017; Angrist and Lang 2004; Carrell, Hoekstra, et al. 2018; Carrell, Malmstrom, et al. 2008; Cools et al. 2019; Lavy and Schlosser 2011) and exploit cross-cohort variation within university and major to deal with student self-selection.[7] For this, we use data on the universe of English undergraduate students and construct foreign shares at the major (e.g. Economics) and higher education institution level (e.g. University of Oxford). In our estimating equation, university and major fixed effects eliminate spurious correlations given by stable characteristics of universities and majors. Furthermore, university and major fixed effects interacted with cohort fixed effects eliminate transitory shocks that are common to either universities or majors. These account for shocks, for example, on labour demand for specific majors that may drive the flow of international students at the same time that shift the composition of native students (Taylor and Rampino 2014).

In addition, the other main challenge that one faces on estimation of peer effects is reflection (Carrell, Hoekstra, et al. 2018). The reflection problem (Manski 1977) is given by the joint determination of individual and peer outcomes. This relates to the reverse causality problem (e.g. Miguel et al. 2004) and is given by the difficulty of distinguishing the effect of peers on the individual from the effect of the individual on her peers (Carrell, Fullerton, et al. 2009). In our set-up, however, reflection is not a problem. This is because foreign status is determined before the focal individual enters in contact with her peers.

In our analysis, we find mild foreign peer effects. Our estimates have magnitudes that are typically smaller than half of a percentage point per percentage point increase in foreign shares. Even more, for the average student, we find no statistically significant effects on any of our higher education outcomes. These are: graduating,

[6]These include sex, ethnicity and socio-economic background.

[7]The other main branch of the peer effect literature uses random assignment of individuals to peer groups (e.g. Anelli and Peri 2017; Carrell, Fullerton, et al. 2009; Lyle 2007; Sacerdote 2001). See also Sacerdote (2011) review of the literature.

dropping out, changing from a STEM to a non-STEM major,[8] graduating with a first, graduating with an upper second, graduating with a lower second, graduating with a third, graduating with a lower qualification or failing to graduate.

Across the ability distribution, we show that foreign peers decrease the probability of graduating for native students at the middle-top of the ability distribution. For these students, a one percentage point increase in the share of foreign peers decreases the probability of graduating by.014-.027 percentage points. This change in the probability of graduating works mostly through foreigners increasing the probability of dropping out for natives above the ability bottom quartile. For these, a one percentage point increase in foreign shares increasing the probability of dropping out by.004-.015 percentage points.

We also find heterogeneous foreign peer effects on grades across the ability distribution. Nonetheless, these effects can be mechanical if teachers grade on a curve.[9] This is because, under grading on a curve, differences in the quality of foreign and native students mechanically displace native grades. We explore formally this mechanism and provide evidence showing that grading on a curve, alone, cannot explain the effects we find. This suggests that foreign students mildly affect the human capital acquisition of native students. In the theoretical model outlined at the beginning of the paper we also illustrate how foreigners could affect natives effort provision in their study, and consequently, their educational performances. We show that this effect can be heterogeneous across students ability distribution and we conclude that determining its direction is purely an empirical question. Our estimates indeed show that foreign peer effects are heterogeneous across the ability distribution.

For a sample of students that graduate in higher education, we observe their labour market outcomes six months after graduation.[10] With this sample we estimate the effect of foreign peers on the probability of continuing studying, working, working in a professional occupation, having a graduate job and average earnings. We find mild foreign peer effects. Increasing the foreign share by a one percentage point increases the probability of working by.064 percentage points for native students at the top of the ability distribution. Mostly by driving them out of postgraduate education. Increasing the foreign share also increases the probability of working in a professional occupation for all native students and has no effect on the probability of holding a graduate job. Finally, increasing the foreign share by a one percentage point increases natives earnings at the third and top ability quartiles by around .17%.

Finally, this paper aims to contribute to the current policy debate in the UK. In 2017, around 82% of all higher education students in the UK were enrolled in an English university and the UK is the second country in the world, after the US, in terms of foreign students in higher education. About one-fifth of all students in the

[8]STEM stands for science, technology, engineering and mathematics.

[9]Under grading on a curve, teachers do not report actual grades. Instead, they transform the actual distribution of grades to have some distribution of their choice.

[10]We show that the sample with matched labour market outcomes is representative of the population of higher education graduates.

UK higher education system come from outside the UK: 6% from the rest of the EU and 13% from the rest of the world.[11] This implies that exposition to foreign peers is non-negligible for most native students. Furthermore, there is substantial variation on the concentration of international students across different stages of education and fields of study. For example, in science (non-science) subjects, international students nowadays constitute 13% (16%) of undergraduate students, 26% (39%) of postgraduate taught students, and 42% (42%) of postgraduate research students.[12] This variation across majors, higher education institutions and time is crucial for our identification strategy. In addition, the effect of foreign students in the UK is a matter of current policy interest (MAC 2018). Our research, therefore, informs British policy-makers about what could be the possible effects of changing foreign student rates in British universities on natives educational and labour market outcomes. This is at a time where foreseeable labour market restrictions and higher fees for EU students may make it harder for the UK to keep up with mainland European competition.[13]

Our work contributes to the existing literature in several ways. We add to the previous literature estimating the effect of immigration on natives education (e.g. Betts 1998; Betts and Fairlie 2003; Hunt 2017; Llull 2017) and the effect of immigrants on natives labour market outcomes (e.g. Butcher and Card 1991; Dustmann, Frattini, et al. 2013; Manacorda et al. 2012; Ottaviano and Peri 2012) in that we provide the effect of direct interaction with immigrants. In this literature, the effect of immigrants works through general equilibrium arguments. Changes in the supply of labour, induced by immigrants, shift the distribution of wages (Dustmann, Frattini, et al. 2013; Manacorda et al. 2012) and these shifts change incentives to acquire human capital (Llull 2017). Our work estimates a different margin of immigrant effects. Instead of general equilibrium effects, we estimate the effect of sharing education with foreign peers. This can be given by direct competition with foreign students or externalities in education. The existing literature studying these margins is scarce. Anelli, Shih, et al. (2017) study the effect of foreign peers on the probability of graduating in a STEM major using data from an introductory math course in a Californian university. They find that foreign students lower the probability of natives graduating in STEM majors. However, this negative impact on STEM major graduation does not produce negative effects on natives wages as they move into high earnings social sciences. Chevalier et al. (2019) use data from economic students in an English university and show that increasing in-classroom linguistic diversity has no effect on English-speaking students. Braakmann and McDonald (2018) consider all English universities, but they focus on the impact of exposure to socio-economic diversity on students educational outcomes. We add to Anelli, Shih, et al. (2017), Braakmann and McDonald (2018), and Chevalier et al. (2019) in that our estimates provide the average effect for the whole of England for a wider set of outcomes. Thus, by covering

[11] https://www.ukcisa.org.uk/

[12] (https://www.hesa.ac.uk/data-and-analysis/students/where-from).

[13] The UK is already losing market share against other competitor countries such as Australia, Germany or France (Universities UK 2018).

multiple universities and degrees, we avoid possible external validity concerns from using data from a single institution and field of study.

More generally, our work relates to the literature on foreign peer effects at lower levels of education. The most relevant paper for us, as it studies foreign peer effects in English primary schools, is Geay et al. (2013). They find a negative effect of foreign peers on English natives that is driven by selection. Once selection is taken into account they find a zero foreign peer effect. Our findings at higher eduction are consistent with these findings. Gould et al. (2009) find an overall no effect on Israeli student and a negative effect for disadvantaged Israeli students. Ohinata and Van Ours (2013) find mild foreign peer effects on Dutch students. However, Ballatore et al. (2018) find negative foreign peer effects on maths and science performance of second graders.

4.2 Theoretical Framework

To fix ideas here we present a simple framework of student admission, spill-overs in grade production and peer determined effort costs.[14] Imaging that there is a mass H and I of home and international students,[15] that we call cohort sizes. Every student is characterized by her educational achievement up to higher education (w), that we call ability, and whether she is home or international. Ability for home and international students has distribution F_H and F_I and common support $[0, 1]$.

4.2.1 Admission

Every institution chooses cut-off grades, $\underline{w}_H$ and $\underline{w}_I$, above which all students receive an offer that accept with probability $\alpha_H(w)$, $\alpha_I(w)$. The chosen cut-off grade is such that it maximizes the following constrained optimization problem[16]

$$\max_{\underline{w}_H, \underline{w}_I} H \int_{\underline{w}_H} \psi(w)\alpha_H(w)dF_H(w) + I \int_{\underline{w}_I} \psi(w)\alpha_I(w)dF_I(w) \tag{4.1}$$

$$s.t.$$

$$H \int_{\underline{w}_H} \alpha_H(w)dF_H(w) \leq c_1 \tag{4.2}$$

$$H \int_{\underline{w}_H} \alpha_H(w)dF_H(w) + I \int_{\underline{w}_I} \alpha_I(w)dF_H(w) \leq c_2 \tag{4.3}$$

$$w_H, w_I \geq 0 \tag{4.4}$$

[14] We leave the analysis of foreign peers effects through relative grades to section 4.6.3 where we also provide a test for it.

[15] At minimal risk of confusion we also use H and I to refer to home and international students.

[16] This framework is inspired on Bhattacharya et al. (2017). A caveat about (4.1) is that universities do not take into account the effect of peers, thus we assume that universities ignore the existence of peer effects. This assumption may not be as restrictive as it sounds. If peer effects were to be well know it would be hard to motivate the extensive literature on peer effects (see Sacerdote 2011). Nonetheless, ultimately, we make this assumption for simplicity.

Where $\psi(w)$ is the weight that the university puts on a student of ability w. For example, this can reflect expected outcomes of a student of ability w. However, our only assumption about ψ is that it is monotonically increasing on ability. The first constraint in (4.2) reflects the funding cap for home students, British and EU, so it doesn't apply to international students. The second constraint in equation (4.3) reflects the capacity constraint of the university and, as such, it applies to both home and international students.[17]

We work under the condition that, for all universities, there is always a home cut-off small enough but larger than zero such that the funding constraint binds.[18] Then depending on the level of the restrictions, the ability distributions and the supply of home and international students, we have three candidates of the optima that we use to classify universities.

Selective Universities are able to select the best home and international students and they set a common entry ability cut-off, $w_H = w_I = \underline{w}$ for both. This entry cut-off is such that the capacity constraint, (4.3), binds and it must be the case that the following inequality holds

$$c_2 - c_1 < I \int_{\underline{w}} \alpha_I(w) dF_I(w) \tag{4.5}$$

Equation 4.5 is quite illustrative. Selective universities must be able to attract international students. Otherwise, if $\alpha_H(w) = 0 \ \forall w$, the necessary condition in (4.5) will only hold if the capacity constraint is smaller than the funding constraint. What seems a rather odd situation. In the opposite direction, universities which offers are always accepted, $\alpha_H(w) = 1 \ \forall w$, are the best candidates to be of selective type. Under this later condition, the difference between the funding and the capacity constraint needs to be smaller than the number of international students with ability $\underline{w}$ or higher, i.e. $c_2 - c_1 < I(1 - F_I(\underline{w}))$.[19]

In **Non-Selective Universities** both the funding and the capacity constraint bind. Moreover, home and international students have different ability cut-offs. To understand why, it is illustrative to note that the shadow price of the funding constraint is $\psi(w_H) - \psi(w_I)$. Given that we assume $\psi'(w) > 0$, this implies the necessary condition $w_H > w_I$. This is intuitive. If the home cut-off were to be smaller than the international one, the university could be better off by increasing the home cut-off. This will cut the number of home students accepted and will allow the university to increase the number of international students, therefore, increasing the ability of the marginal student. Thus, non-selective universities use international students to fill capacity and, for this, they need to set a lower cut-off for international students, as compared with natives.

Supply Constrained Universities are not able to fill capacity using international students and cannot enrol more home students because they are already at the

[17]For example, given by availability of space and teachers.
[18]This reduces the number of candidates for optimal points and leaves our cases of interest.
[19]The cut-off $\underline{w}$ must also be such that the funding constraint for home students doesn't bind.

funding constraint. This leads to different cut-off grades for home and international students. The cut-off grade for home is such that the funding constraint holds with equality and the cut-off grade for international students is $w_I = 0$. Thus supply constrained universities send offers to all international students. Furthermore, in supply constrained universities the following inequality holds

$$I \int_{w_I} \alpha_I(w) dF_I(w) < c_2 - c_1 \tag{4.6}$$

This necessary condition is similar to the selective university counterpart. However, in this case those universities that are most-likely to be supply constrained are those that face small international student supplies, i.e. $\alpha_H(w) \approx 0$. Imaging the opposite, if all international students were to accept offers from supply constrained universities we will have the necessary condition in (4.7). This means that the university would have to have capacity to take in the whole immigrant cohort plus c_1 natives. This is highly unlikely.

$$\frac{c_2 - c_1}{I} > 1 \tag{4.7}$$

4.2.2 Cohort Size, Number of Students Enrolled and their Ability

The effect of international peers on the number of enrolled home students is heterogeneous across and within university types. From (4.8), where $h(w_H)$ is the number of home students enrolled, it follows that international students only have a negative effect on the number of enrolled home students in selective universities. For the other universities, the home ability cut-off does not move with the cohort size of international students and (4.8) is zero. In other words, the demand of home students only depends on the supply of international students for selective universities. In addition, within selective universities, the displacement of home students by enrolled international students can be heterogeneous. For example, consider two selective universities that set the same cut-off, but the first one faces a larger home demand than the second, i.e. $\alpha^1_H(w_H) > \alpha^2_H(w_H)$. Then the displacement of home students is larger for the first, more demanded, university.

$$\frac{dh(w_H)}{dI} = -H\alpha_H(w_H) f_H(w_H) \frac{dw_H}{dI} \tag{4.8}$$

On the other hand, the effect of increasing cohort size on the distribution of ability and average ability is similar for selective and non-selective universities. In these universities increasing the cohort size of either home or international students pushes the mass of the ability distribution towards the top. This can be seen in equation (4.9), where G is the distribution of ability for enrolled students and the equivalence follows from defining $\phi_K(1, w) = \int_w^1 \alpha_K(\omega) dF_K(\omega)$ and using the fact that selective and non-selective universities always fill their capacity. However, the mechanism driving this positive effect on ability of enrolled students is different for selective and non-selective universities. In selective universities, increasing the cohort size of international students rises the ability cut-off for both home and international students. This reduces the number of enrolled home students and increases the ability of the marginally enrolled student. In non-selective universities, however, increasing the cohort size of international students only increases the ability cut-off of international students. This implies that neither the number of enrolled international nor home students changes, but the marginally enrolled international students is of higher quality. Therefore, the share of international students enrolled stays constant but the ability of international students increases, resulting in a push of the ability mass towards the top of its distribution.[20]

$$\begin{aligned} \frac{dG(w|w_H, w_I)}{dK} &= -\frac{\int_w^1 \alpha_K(\omega) dF_K(\omega)}{H \int_{w_H} \alpha_H(\omega) dF_H(\omega) + I \int_{w_I} \alpha_I(\omega) dF_I(\omega)} \\ &\equiv -\frac{\phi_K(1, w)}{c_2} < 0, \;\; K = H, I \end{aligned} \tag{4.9}$$

$$\frac{d\,\mathbb{E}(w|w_H, w_I)}{dK} = \frac{\int_{w_K}^1 [w - w_K]\, \alpha_K(w) dF_K(w)}{c_2} > 0, \;\; K = I, H \tag{4.10}$$

In supply constrained universities the effect of increasing the cohort size of home students on the ability of enrolled students is analogous to the effect experience by selective and non-selective universities.[21] However, the effect of increasing the cohort size of international students is undetermined. Increasing the size of the international cohort will push the ability mass towards the bottom of the distribution and, therefore, decrease the average ability of enrolled students if the stochastic dominance relation in (4.11) is met. But nothing guarantees that (4.11) is going to be met. For example, imaging that international students have higher ability in the population, in a first order stochastic dominance sense, and have the same supply as native students, i.e. $\alpha_I(w) = \alpha_H(w)\ \forall w$. Then increasing the size of the international cohort can have a heterogeneous effect. We represent this situation in figure 4.5. What happens is that, in supply constrained universities, an increase on the international cohort size increases the share of international students enrolled but keeps constant their ability distribution. This increases the mass of enrolled students bellow the home cut-off, independently of whether international students are of higher or lower ability. However, if international students are of higher ability, as assumed in figure 4.5, increasing the cohort size of international students will also increase the number of students towards the top of the distribution. Thus condition (4.11) is not met all across the ability distribution. Which of this two effects is the strongest will determine whether the average ability of enrolled students increases or decreases with the

[20] A similar discussion holds for increasing the cohort size of home students.

[21] With the caveat that the number of enrolled students is smaller than the capacity constraint. In (4.9) and (4.10) one needs to substitute c_2 with the number of enrolled students, i.e. $H \int_{w_H}^1 \alpha_H(\omega) dF_H(\omega) + I \int_0^1 \alpha_I(\omega) dF_I(\omega)$.

cohort size of international students.

$$\frac{dG}{dI} > 0 \Leftrightarrow \frac{\phi_I(w,0)}{\phi_I(1,0)} > \frac{\phi_H(w,w_H)}{\phi_H(1,w_H)} \tag{4.11}$$

4.2.3 Effort Provision

Once enrolled, students choose how much effort, e, and leisure, l to consume by solving the following problem

$$\begin{aligned} &\max_e f(\bar{w})e^{\sigma} - eC(w,\bar{w}) \\ &s.t.\ e + l = T \end{aligned} \tag{4.12}$$

Where $f(\bar{w})e^{\sigma}$ is the expected grade given effort level e and G is the distribution of ability of enrolled students. The grade production function in (4.12) is similar to Moretti (2004) in that it allows spill-over effects given by the average quality of peers. The cost function, on the other hand, is inspired by the work of Elsner and Isphording (2017) and Murphy and Weinhardt (2018), showing that within school rank has a positive effect on students outcomes. Therefore, we assume that $\partial C/\partial w < 0$. Based on the findings of Elsner and Isphording (2017) and Murphy and Weinhardt (2018) we interpret this negative effect of own ability on the marginal cost of effort as reflecting improved self-esteem. Assuming an interior solution, the optimal allocation of effort is given in (4.13) and we are interested on the effects that increasing cohort sizes have on it, i.e. equation (4.14).

$$e^* = \left(\frac{c}{f\sigma}\right)^{\frac{1}{\sigma-1}} \tag{4.13}$$

$$\epsilon_K \equiv \frac{d\log(e^*)}{dK} = \frac{1}{1-\sigma}\left[\frac{1}{f}\frac{\partial f}{\partial \bar{w}}\frac{d\bar{w}}{dK} - \frac{1}{c}\frac{dC}{dK}\right] \tag{4.14}$$

To learn more about the possible direction of the effect in (4.14) we introduce peer ability as an effort augmenting technology, see equation (4.15). For the cost function we introduce two specifications. In the first one, equation (4.16), individuals compare themselves with the average peer. This induces that top students are the ones whose marginal effort cost increases the most with average peer ability. We interpret this as the impact on self-esteem from increasing peer ability been the highest for top ability students. In our second cost specification, equation (4.17), students compare themselves with all other students, thus their marginal effort cost depends on their ranking within the peer group. In this case, those that suffer the most from a larger cohort of students are those towards the bottom of the ability distribution. This is because a large fraction of the additional students have ability equal or higher than them. High ability students, on the other hand, do not suffer much as only a small fraction of the new students will be of higher ability than them.

$$f(\bar{w}) \equiv \bar{w}^{\sigma} \tag{4.15}$$

$$C_1(w, \bar{w}) \equiv a_1 - a_2 \frac{w}{\bar{w}} \tag{4.16}$$

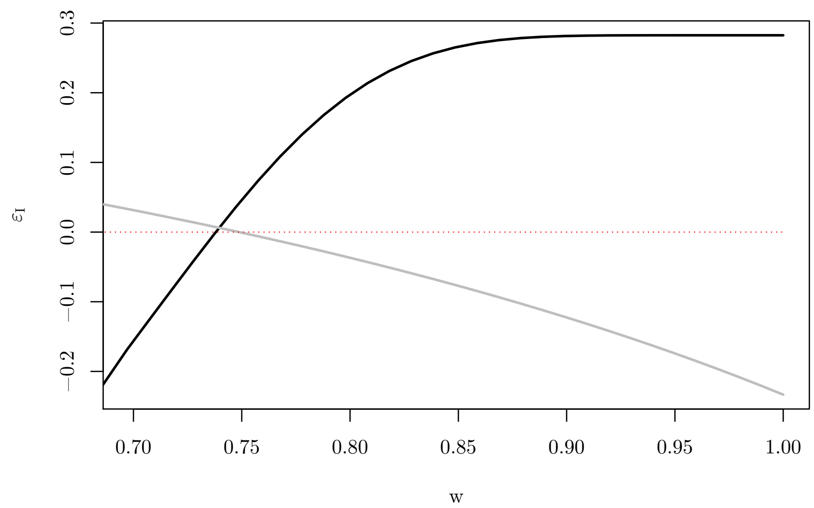
$$C_2(w, \bar{w}) \equiv b_1 - b_2 G(w) \tag{4.17}$$

Given our assumptions, the peer effect on effort productivity is constant across the ability distribution. However, the effect on the marginal effort cost varies across ability. These movements across the ability distribution have the potential to make peer effect heterogeneous, as we illustrate it in figure (4.1). Using a beta distribution that resembles the empirical distribution of ability and assuming an exponential supply of students,[22] we simulate the effect of increasing the cohort of international students on the effort provision of students in a selective university. The two cost functions in equations (4.16) and (4.17) produce two different peer effects. The linear on rank marginal effort cost, produces a peer effect that is negative for low ability students and positive for high ability. This is because increasing the average ability of peers has the same effort productivity effect on all students, but those students at the bottom suffer large self-esteem costs derived from being pushed further towards the bottom by international students. If instead costs are a linear function of student ability relative to average peer ability, we obtain the opposite behaviour. The intuition, however, is similar. The effort productivity effect is still constant, but, under cost specification (4.16), those students that suffer the most from international students pushing up the average ability of peers are those at the top of the ability distribution. This creates a positive effect for low ability students and a negative one for high ability.

These already possibly heterogeneous peer effects could be made even more flexible if we were to allow the effect on effort productivity to vary across the ability distribution. For example, imaging that own and peer ability are substitutes. In such case the effect on effort productivity from increasing average peer ability is strongest for low ability students. This creates a tension if students compare themselves with all peers instead of the average one. With this cost function the self-esteem of low ability individuals suffers the most and they receive the largest increase on marginal effort cost. However, under substitution between own and peer ability, low individual students are also the ones that benefit the most in terms of effort productivity from improved ability of peers. Which of these two dominates is, given the assumptions we have made, unknown.

[22] We assume that $\alpha_K(w) = \lambda \exp(-\lambda w)$. This can be interpreted as an approximation to logistic supplies when the number of universities is large (see Card, Cardoso, et al. 2018).

Figure 4.1: Simulated International Peer Effect
On Effort Provision



Notes: Simulated effects from increasing international cohort size on effort provision. Black line uses linear cost on student rank. Grey line uses linear cost on relative to average peer ability.

Summing up, our framework clearly delimits the effects that increasing cohort sizes of international and home student has on the number of enrolled students and their average ability. In selective universities, increasing the cohort size of international students increases the ability of the marginal enrolled student and decreases the number of enrolled home students. In non-selective universities increasing the cohort size of international students does not crowd-out home students, although it increases the average ability of enrolled students through the marginally enrolled international student having a higher ability. In supply-constrained universities that are not able to fill capacity, increasing the international cohort size has an ex-ante undetermined effect.

In contrast, the effect of increasing the cohort size of international students has a heterogeneous effect on effort provision. We show that, even with restrictive assumptions on the effects of peers on effort productivity and cots, the peer effect on effort provision can be heterogeneous across the ability distribution.

4.3 Empirical strategy

We estimate peer effects using a reduced form framework where outcome (y) is determined by a linear combination of foreign status (z), other individual characteristics (x),[23] the share of foreigners in the group defined by x ($\mathbb{E}[z|x]$) and unobserved characteristics (ϵ)

[23]Including those defining group belonging.

$$y = \beta_0 + \beta_1 \, \mathbb{E}[z|x] + x'\beta_2 + z\beta_3 + \epsilon \tag{4.18}$$

This reduced form can be obtained by following the linear peer effect model introduced by Manski (1993), where reduced form parameters are functions of individually non-identified structural parameters. Particularly, the peer effect β_1 is a combination of the contextual (individual outcomes varying with group characteristics) and endogenous (individual outcomes varying with group outcomes) effect without further capacity to distinguish one from the other (see Manski 1993).

In our set-up, reflection (Manski 1993) is not a problem since the characteristics giving the peer effect are fixed. However, correlated peer characteristics could risk causal interpretation of our estimates. Given that our interest is on the effect of foreign peers on natives, we fix z_i for all individuals (i) in our population of interest. This creates a clear cut between our subjects and the peers influencing their outcomes that helps differentiating peer effects from correlated own characteristics (Angrist 2014; Carrell, Hoekstra, et al. 2018). Although this comes at the cost of neither identifying the effect of being a foreigner on outcomes nor possible heterogeneity in peer effects for foreigners and natives.

The unobserved component (ϵ) in (4.18) includes characteristics that are hardly measurable, such as institution studying facilities. Our identification strategy relies on us controlling for these through a comprehensive set of fixed effects. For example, those UK universities that provide higher quality or a more valuable signal might be more attractive to foreign students. If university quality also affects the outcomes of native students we will have an omitted variable bias, as correlated effects induce a positive relation between natives outcomes and the share of foreign students.

In our linear set up, correlated effects given by institution or major characteristics can be easily controlled for by introducing institution and major fixed effects into (4.18). These fixed effects leave variation across cohorts within majors or institutions to identify the peer effect. However, using this variation to identify the causal effect of foreign peers would introduce rather stringent correlation conditions. For example, one could think of labour market shocks that shift returns to specific majors, therefore, modifying incentives for natives to enrol and put effort into their studies. At the same time, if immigrants plan to stay in the British labour market or the shocks are global, labour market shocks will drive the choices of foreigners and, thus, the immigrant composition of the peer group. This leads to an spurious correlation between foreign student shares and natives' outcomes that would bias our estimates. We follow the literature (see Sacerdote 2011) and use a comprehensive set of institution-major, institution-cohort and major-cohort fixed effects that difference out correlated shocks. This leads to the following estimating equation

$$y_{ihmc} = \beta_0 + \beta_1 \bar{z}_{hmc} + x_i'\beta_2 + \tau_{hm} + \upsilon_{mc} + \omega_{hc} + \epsilon_{ihmc}, \tag{4.19}$$

where h is higher education institution, m is major, and c is cohort defined by year of enrolment in the undergraduate course. It follows that identification of the peer effect relies on $\mathbb{E}[\bar{z}\epsilon|\tau, \upsilon, \omega] = 0$ and we identify peer effects from variation on foreign shares across cohorts within the same institution and major.

Thus, our strategy follows the literature on peer effects which has mainly relied on quasi-random variation across cohorts within a unit, that could be a school or a firm, to analyse a vast range of outcomes.[24] More specifically, our specification is similar to Cornelissen et al. (2017), where they include fixed effects for occupations within firms and is more flexible than some specifications in the literature that account only for institution fixed effects and institution-specific linear time trends (e.g. Geay et al. 2013; Hanushek et al. 2003).

4.4 Data and sample

We use administrative data from the Student Record provided by the Higher Education Statistical Agency (HESA). This contains information on the whole population of students that enrolled in a higher education institution in the UK in academic years 2001/02 to 2010/11.

In our analysis, we define as population of interest the sub-population of students that are enrolled in a three year undergraduate course, that at the time of entering higher education are 18-21 years old and come from high school, without any prior experience at higher education. We do this selection to avoid possible selection concerns from students that might have had some previous experience in higher education or the labour market and to limit heterogeneity in terms of the type of education students are getting. At the same time, selecting young students also eliminates possible heterogeneity due to different educational motives between young and mature students.

To construct the peer group we use the whole population of students enrolled in a first year undergraduate course, without any age restriction,[25] and define peer groups as all students that are enrolled in the same higher education institution and major.

For each peer group we compute foreign exposure as the share of foreign peers over native students, where we define native status by looking at whether a student was domiciled in the UK prior to enrolment into higher education.[26] A major benefit of defining natives and migrants by place of domicile prior to higher education instead of nationality is that students who were residing in the UK before enrolling into higher education have, most likely, received secondary education in the UK. Thus, they are comparable in terms of the previous education and incentives they had for enrolling in certain majors and institutions.

[24] For a comprehensive literature review see Sacerdote (2011).

[25] We impose the 18-21 age restriction on focal students only.

[26] A popular definition for native in the literature (e.g. Dustmann, Frattini, et al. 2013; Manacorda et al. 2012) is whether the individual was born in the country. However, we have no information about country of birth.

A feature of our data is that some students participate in multiple groups, for example there are students that are simultaneously enrolled in business and economics, or philosophy and education. To construct peer groups, we count every individual in every major she participates, and compute $\bar{z}$ in (4.19) as

$$\widehat{\mathbb{E}[z|x]} = \bar{z}_{hmc} = \frac{\sum_{j \in FP_{hc}} \mathbb{1}[j \text{ participates in major } m]}{\sum_{j \in S_{hc}: j \neq i} \mathbb{1}[j \text{ participates in major } m]} \tag{4.20}$$

where S_{hc} are all students of cohort c in university h and FP_{hc} is the subset of them that are foreigners. When computing foreign shares in 4.20, we assume that students enrolled in multiple majors have equal participation in all of them and count them into each of the majors in which they participate.[27]

In the estimating sample, we select only those students that are enrolled in a single major as it is unclear which is the peer group for those that are enrolled in multiple majors.[28] In practice, including also those that are enrolled in multiple majors creating one observation per major and weighting by the relative importance of the major in the course as provided by HESA major weights does not affect our estimates.

We consider two types of outcomes: attainment at higher education and participation and performance in the labour market. For higher education outcomes we observe the universe of all undergraduate students and we investigate whether a higher share of foreign peers in the same university and major affects the probability of graduating, graduating with a first degree, graduating with a qualification lower than a first degree, transitioning from a STEM to a non-STEM major (and vice versa), and graduating with a first, upper second, lower second, or third grade.

In table 4.1 we provide descriptive statistics of higher education outcomes for cohorts 2007-2010. We do not use earlier cohorts because only from 2007 onwards we have measures of pre-higher education attainment and, in section 4.4.2, we show that one can only treat cross-cohort variation as quasi-random after controlling for pre-higher education attainment.

In our sub-population of students, 89.8% of them graduate: 86.1% graduate with a first degree (*Graduates Successfully*) and 3.7% with a lower qualification (*Graduates Lower Qual.*); 1.7% of our students fail to graduate and 8.5% drop out. We differentiate between failing and dropping out by looking at the last course in which students were enrolled. If a student drops before reaching the third year we classify her as a dropout, if she arrives to the last year but does not graduate we classify her educational outcome as a fail. Of those students that did not drop out, 14.4% graduated with a first, 52.6% with an upper second, 23.4% with a lower second, 3.7%

[27] In unreported results we computed shares assuming that participation is proportional to major weights provided by HESA that measure the relative importance of the major inside the course. This did not change our main results.

[28] Students enrolled in a single major represent around 80% of the entire population of undergraduate students.

with a third, 3.7% with a lower qualification and 4% fail to graduate[29]

Out of 326,997 students that started in a non-STEM major, 1.8% transitioned in a STEM major. In comparison, a larger proportion, 4.8%, of the 239,322 students that started in a STEM major transitioned into a non-STEM major.

Table 4.1: Higher Education Outcomes

	Observations	Mean	Std.Dev.	Min	Max
Graduates	458180	0.898	0.303	0.000	1.000
Graduates Succesfully	458180	0.861	0.346	0.000	1.000
Graduates Lower Qual.	458180	0.037	0.188	0.000	1.000
Fails (to graduate)	458180	0.017	0.129	0.000	1.000
Dropout	458180	0.085	0.279	0.000	1.000
To STEM	327000	0.018	0.131	0.000	1.000
To non-STEM	131180	0.048	0.215	0.000	1.000
Graduates with First	416030	0.144	0.351	0.000	1.000
Graduates with Upper Second	416030	0.526	0.499	0.000	1.000
Graduates with Lower Second	416030	0.234	0.424	0.000	1.000
Graduates with Third	416030	0.037	0.190	0.000	1.000
Graduates with Lower Qual.	416030	0.037	0.190	0.000	1.000
Fails to Graduate	416030	0.040	0.197	0.000	1.000

Note: Observations from 2007 onwards only. Observations rounded to last unit.

To study labour market outcomes, we link the HESA Student Record with the Destination of Leavers from Higher Education (DLHE) survey that collects information, at six months after graduation, on the destination of all natives and EU graduates in the UK that graduated between 2002 and 2012. In the whole data, the DLHE response rate is about 80% for natives and 60% for EU students, with international students not included for the period considered. In our estimating sample we select cohorts 2007-2009. As with higher education outcomes, we do not use cohorts prior to 2007 because for earlier cohorts we do not have information on ability. Furthermore, in the DLHE data we drop cohorts that entered education after 2009 because we consider three years courses and the latest graduation year available in our DLHE data is 2012.

Our native sample for labour market outcomes comprises 146,491 students with 75.41% of our native student sub-population in the student record matched in the DLHE data. Although this is a very good matching rate, we address remaining concerns by replicating higher education outcomes estimates using the DLHE matched

[29]The number of individuals for whom we observe grades is smaller than the proportion of students that graduate or didn't drop out. This is because some majors, such as medicine, do not use this grading system.

subsample. In section 4.7 we show that our results hold in the DLHE matched subsample.

The labour market outcomes can be divided into activity status: whether employed or studying; and job attributes: whether working in a professional or graduate job and earnings. Of the 146,491 graduate students that we have matched in the DLHE data 59.4% are working, 67% are working or working while studying (*Working (& work+stud.)*), 9% are unemployed, 16.8% are studying and 7.3% are doing something else. When we select only those that are participating in the labour market (*Working vs Unemployed*), 88.2% have a job. Of those with a job, 55.9% are in professional occupations and 33.7% have a graduate job.[30] Observed average (log-)earnings are 9.803, close to average imputed earnings from the Annual Survey of Hours and Earnings (*Salary imp.*).

Although our labour market outcomes are measured at six months after graduation, a relatively short-time after exiting education, they are of economic relevance. The literature has shown that labour market entry conditions largely affect individuals' working trajectories, especially among the population of graduates (Baert et al. 2013; Kahn 2010; Oreopoulos et al. 2012; Raaum and Røed 2006; Von Wachter and Bender 2006).

Table 4.2: Labour Market Outcomes

	Observations	Mean	Std.Dev.	Min	Max
Working	146490	0.594	0.491	0.000	1.000
Working (& work+stud.)	146490	0.670	0.470	0.000	1.000
Working vs Unemployed	111260	0.882	0.323	0.000	1.000
Unemployed	146490	0.090	0.286	0.000	1.000
Studying	146490	0.168	0.374	0.000	1.000
Studying Postgrad	146490	0.091	0.287	0.000	1.000
Studying Undergrad	146490	0.075	0.264	0.000	1.000
Work & Study	146490	0.075	0.264	0.000	1.000
Other	146490	0.073	0.259	0.000	1.000
Professional Occ.	86890	0.559	0.496	0.000	1.000
Graduate Job	77480	0.337	0.473	0.000	1.000
(log)Salary	40940	9.803	0.340	2.079	12.578
(log)Salary imp.	51760	9.835	0.367	2.079	10.597

Note: Observations from years 2007-2008 only. 75.41% of the student record observations are matched in the DLHE data. Observations rounded to last unit.

[30] Differences in observation numbers are due to missing items for some individuals in the DLHE data.

4.4.1 Source of variation

By inspection of equation (4.19) it follows that identification relies on variation across cohorts within a university-major, as the fixed effects absorb variation across majors, universities and cohorts within major and within university.

In our data the average foreign peer share is at 10.5% of students in a major-university, with 60 students being foreigners on average and a standard deviation of the share of 10.8 percentage points, see table 4.3. The average EU student share is 4.1% and non-EU 6.3%, with standard deviations of 4 and 8.3 percentage points. When we introduce our set of fix effects they, obviously, reduced the amount of variation available. The standard deviation for foreign peer shares goes from 10.8 to 2.3 percentage points when looking at levels and residuals from fix effects regressions. For EU student shares standard deviations go from 4 to 1.3 percentage points and for non-EU from 8.3 to 2. This reduction on variation is desirable because it implies that we only exploit within group variation. In figure 4.9, we show that international students are mostly represented (about 20% of the whole undergraduate population) in Business studies, Engineering and Economics; followed by Maths and Computer science and Technologies. For EU students the most selected subjects are Economics, Business, Engineering and Architecture, figure 4.10. The non-homogeneous distribution of foreign students across majors suggest that there may be substantial differences in terms of incentives and therefore selection across majors and universities.[31] This makes cross major-university variation not really useful in terms of identification of the effect of foreign students on natives, as we discussed in section 4.3.

In terms of magnitudes, if we look at variation within major-university between consecutive cohorts, second panel of table 4.3, the average change is at .4 percentage points for the foreign share and .2 percentage points for EU and non-EU shares, with standard deviations of 6 percentage points for changes in foreign shares, 3.5 for EU shares and 5.4 for non-EU. Furthermore, even large increases on the number of foreign students, as the one following EU enlargement towards Eastern European countries in 2004 (figure 4.3), induce changes in the share of immigrants that are much smaller than those observed when comparing shares across majors or universities, see figure 4.4.[32] Thus, from a policy perspective, we are interested in changes that are of magnitude similar of those across cohorts and much smaller than changes across majors or higher education institutions.[33]

[31] See figure 4.11.

[32] In unreported results we estimated the effect of A10 students on major-university aggregated native outcomes in first-differences comparing the year before the accession of A10 countries and the year after. These estimates return the same picture, small effects that are typically non-statistically significant.

[33] If we consider the reduced form effect as a best linear approximation, the closest we are to the point of interest the better the approximation.

Table 4.3: Variation

	Foreign	EU	Non-EU
	Fix Effects		
Standard Deviation	0.023	0.013	0.020
Min	-0.335	-0.178	-0.318
Max	0.833	0.303	0.961
	First Differences		
Standard Deviation	0.053	0.032	0.047
Min	-0.824	-0.375	-0.941
Max	0.375	0.333	0.375
Mean	0.004	0.002	0.002
	Levels		
Average Share	0.105	0.041	0.063
Average Share Standard Deviation	0.108	0.040	0.083
Average Size	59.096	22.620	36.477

Note: Observations from 2007 onwards only.

4.4.2 Validity of Identification Strategy

Our identification strategy relies on exploiting quasi-random variation in foreign peer exposure across cohorts within the same university and major. If variation across cohorts within university and major is as good as random, we should not observe any foreign peer effects on outcomes that pre-date entry into higher education. Here we use two outcomes that are determined before native students actually meet their peers: distance between home and chosen institution, and ability. Distance between home and university is chosen at the moment of applying and accepting an university offer and as such pre-dates interaction with foreign peer students and should not be influenced by them. Ability is measured using the UCAS score and as such is a summary index of pre-higher education educational achievement, therefore, it is predetermined at the entry in higher education.

In table 4.4 we provide estimates from regressing distance and tariff scores on foreign shares using the fix effect specification in (4.19). Our estimates show no significant effects on distance between home and university location. However, there are effects on tariffs and number of natives enrolled. Increasing the share of immigrants by one percentage point leads to a .177% increase in the tariff and a -1.16% decrease on the number of enrolled natives. These effects are more strongly driven by EU students. This matches what we presented in our framework in section 4.2. EU students are treated by universities as home students and, as such, they put pressure on the number and the ability of natives across all universities. On the other hand, non-EU students only compete with natives in certain institutions, those that are more selective. This can be seen in table 4.4. On average a one percentage point increase on the share of EU students lowers the number of natives by 2%. The same increase on non-EU produces a .93% decrease. However, this changes when we consider specific university groups. In selective universities such as those in the Russell Group,[34] increasing the share of EU and non-EU students has a similar effect. But in less selection universities, those in the 1994 group and all others, the effect of non-EU students gets attenuated. In the last column of table 4.4, increasing the share of EU and non-EU students produces a 1.98% and 2.53% decrease on natives enrolled in Russell Group universities, the baseline. The effect of non-EU students gets attenuated by 1.49 percentage points in the next most selective university group, i.e. 1994 Group, and by 2.04 percentage points in all other universities.

As we introduced in section 4.2, the asymmetry on the treatment of EU and non-EU students induces an asymmetry on the effect that this have on the number and ability of enrolled natives. Furthermore, this asymmetry is heterogeneous across universities depending on how selective they can be. Therefore, we interpret the effects on tariffs and course sizes as evidence showing that universities modify the distribution of student characteristics as a function of the foreign composition of their prospective students. This means that exploiting variation across cohorts within university and major cannot extract peer effects.

[34]See figure 4.12.

Table 4.4: Placebo
Pre-entry Outcomes

	(log-)Distance	(log-)Ability		(log-)Native Students	
Foreign Peers	0.002	0.177		-1.164*	
	(0.106)	(0.109)		(0.490)	
EU Peers	0.040	0.311*	0.284^{+}	-2.055***	-1.980***
	(0.196)	(0.152)	(0.154)	(0.547)	(0.546)
non-EU Peers	-0.013	0.132	-0.086	-0.923^{+}	-2.523***
	(0.126)	(0.129)	(0.145)	(0.533)	(0.667)
1994 Grp. * EU			-0.279		-0.882
			(0.229)		(0.669)
Other HEI * EU			0.151		0.210
			(0.305)		(0.929)
1994 Grp. * non-EU			-0.086		1.487^{+}
			(0.447)		(0.853)
Other HEI * non-EU			0.339		2.043*
			(0.227)		(0.916)
Observations	447,980	458,160	458,160	4,710	4,710

Note: Sample of undergraduate students enrolled in 2007/10 in all English higher eduction institutions (HEI). Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. The overall foreign peer estimated is produced from a separate regression. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Selection carried by universities at entry would imply that our estimates return the effect of foreign students given the university selection mechanism. However, we can improve upon this thanks to the quality of our data. This is because we can control for ability using the UCAS score. We show that, conditional on individual ability, cross-cohort variation on foreign peer shares is uncorrelated with a comprehensive set of predetermined characteristics that are likely predictors of educational performance.[35] Particularly, we provide estimates from regressing socio-economic background (figure 4.2), sex, age and private school (figure 4.6); ethnicity (figure 4.7) and disability (figure 4.8) on the share of EU and non-EU students. Where each of this variables is used in its own regression with our set of fixed effects and conditioning on ability score quartile.

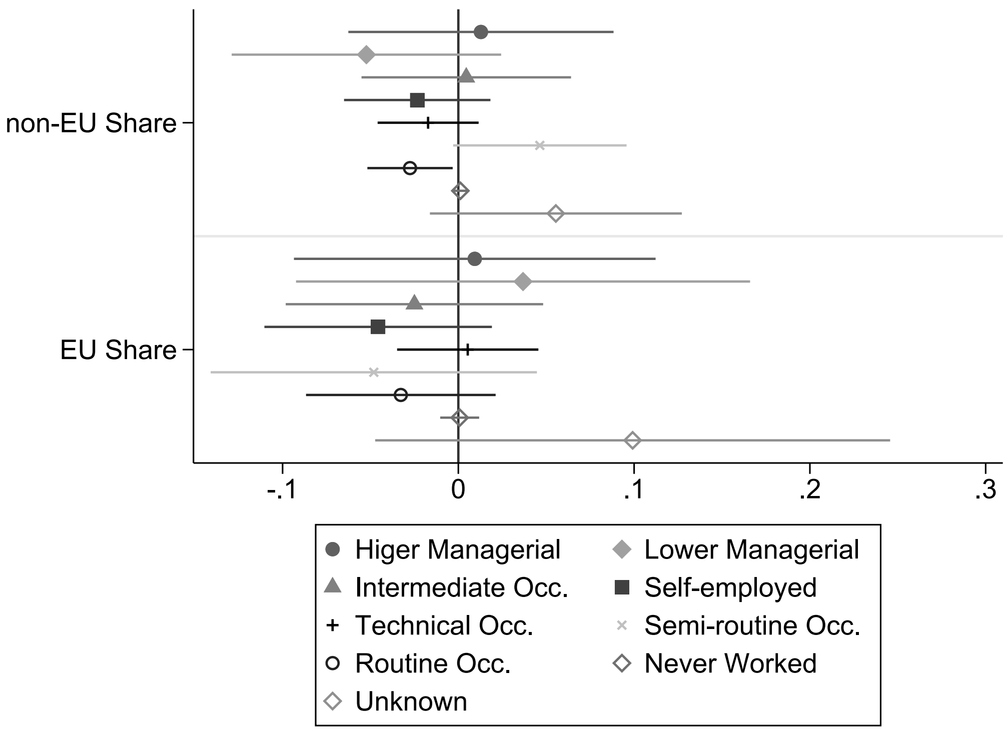
In figure 4.2, all socio-economic groups show non significant correlations, at the 95%, with both the share of EU and non-EU peers and point estimates are typically close to zero. The only exception is the correlation between routine occupations and non-EU share for which we observe a marginally significant correlation. In figures 4.6,

[35]This sort of balancing test have a long tradition on the peer effects literature (Anelli and Peri 2017; Chin et al. 2013; Cools et al. 2019; Lavy and Schlosser 2011).

4.7 and 4.8 we find similar results with most groups having non statistically significant correlations, and a few: age in months and EU share, Indian ethnicity and EU share, and mental health disability and non-EU share presenting significant correlations. To test whether these could drive our estimates, in tables 4.33 and 4.34, we replicate our baseline higher education estimates introducing controls for all these characteristics. This does not produce major changes in our estimates.

Finally, from estimates in table 4.4 one may be concerned about estimating a peer effect that is actually a composite of the foreign peer effect and course size (Angrist and Lavy 1999; Krueger 2003). This is because in table 4.4 we show that non-EU peer shares have a zero effect on the number of natives enrolled in certain universities. We address possible concerns by controlling for course size in all our regressions.

Figure 4.2: Balance Test
Socio-economic Background



Notes: Estimates from regression including year, major, university, year-major, year-university and major-university fixed effects plus controls for tariff score and class size. Balance test for age, sex, school type, ethnicity and disability in figures 4.6 to 4.8.

4.5 Foreign Students and Entry into Higher Education

Our estimates in table 4.4 support the mechanism we introduced in our framework in section 4.2. This is, treating EU and native students as home students in terms of funding constraints induces that EU students put pressure on the number of native students, particularly those of low ability, in all universities. However, non-EU students, that are not subject to funding constraints, only put pressure on natives in selective universities.

In this section we provide further evidence. For this, we estimate a simple model of major-institution choice. Keeping with the notation that we introduced earlier, an individual i from cohort c enrols in a major-institution (m, h) according to a threshold crossing rule

$$\begin{aligned} &V_{mh}(x_i) + \epsilon_{ihmc} \geq V_{m'h'}(x_i) + \epsilon_{ih'm'c} \\ &\forall\ (m', h') \in \{1, \ldots, M\} \times \{1, \ldots, H\} \end{aligned} \tag{4.21}$$

Where $V_{mh}(x_i) \equiv u_{mh}(x_i) + \tau_{hm} + \nu_{mc} + \omega_{hc}$ and x are individual characteristics including her cohort. Assume that the set of major-institution pairs can be partitioned into K disjoint subsets, $B_1, \ldots, B_K$, such that shocks ϵ are uncorrelated across subsets but may be correlated within them. Assuming that the distribution of ϵ belongs to the generalized extreme value family, it is possible to obtain closed form solutions for the probability of choosing a major and institution that are consistent with the nested correlation structure (e.g. see Train 2009, chapter 4). In particular, the probabilities have the nested logit functional form

$$P_{hm}(x) = \frac{e^{V_{mh}(x)/\lambda_k} \left(\sum_{(m',h') \in B_k} e^{V_{m'h'}(x)/\lambda_b} \right)^{\lambda_k - 1}}{\sum_{k=1}^{K} \left(\sum_{(m',h') \in B_k} e^{V_{m'h'}(x)/\lambda_b} \right)^{\lambda_k}} \tag{4.22}$$

In estimation of non-linear functions, such as (4.22), the incidental parameters problem from fixed effects is more of a concern as sufficient statistics are not always available (Gary 1980). For estimation, we exploit a convenient property of the nested logit probabilities. For two major-university pairs $(m, h), (m', h') \in B_k$ the log of the ratio of their probabilities is a linear function

$$\begin{aligned} \log \left(\frac{P_{hm}(x)}{P_{h'm'}(x)} \right) =& V_{mh}(x)/\lambda_k - V_{m'h'}(x)/\lambda_k \\ =& \left(u_{mh}(x) - u_{m'h}(x) + \tau_{hm} - \tau_{h'm'} \right. \\ & \left. + \nu_{mc(x)} - \nu_{m'c(x)} + \omega_{h'c(x)} - \omega_{hc(x)} \right) / \lambda_k \end{aligned} \tag{4.23}$$

In our application we nest major-institution pairs by major and institution groups. As in table 4.4, we classify institutions into Russell group, 1994 group and other. Russell group universities are nationally and internationally well recognized and are mainly research-focused universities. The 1994 group follows Russell in terms of being recognized great credentials and in implementing high entry requirements. Institutions in these two groups are, arguably, the most selective ones in the UK.[36] Given

[36] This can be seen in figure 4.12 where we plot the distribution of ability for natives enrolled in each of these university groups.

this nesting structure, it follows that, if $(m, h), (m', h') \in B_k$, $m = m'$, the (log-)odds ratio in (4.23) simplifies to

$$\log\left(\frac{P_{hm}(x)}{P_{h'm'}(x)}\right) = \left(u_{mh}(x) - u_{m'h}(x) + \tau_{hm} - \tau_{h'm'} + \omega_{hc(x)} - \omega_{h'c(x)}\right)/\lambda_k \qquad (4.24)$$

The probabilities on the left-hand side of (4.24) can be non-parametrically estimated from the data. For example, if x is discrete, one can use a bin estimator. We divide our sample in cells by cohort, ability quartile, region of domicile, sex and ethnicity and compute the probability of entering any major-institution pair for each cell.[37] With these probabilities we estimate the effect of foreign students on the probability of entering a major-institution as

$$\log\left(\frac{\hat{P}_{hm}(x)}{\hat{P}_{h'm'}(x)}\right) = \beta(\bar{z}_{hmc(x)} - \bar{z}_{h'm'c(x)}) + (\tilde{\tau}_{hm} - \tilde{\tau}_{h'm'}) + \tilde{\omega}_{hkc(x)} + \xi_{hmc(x)} \qquad (4.25)$$

Where we have imposed that $u_{hm}(x) = \beta\bar{z}_{hmc(x)} + g(x)$ and $\lambda_k = \lambda \ \forall \ k \in \{1, \ldots, K\}$. Moreover, given that, within a cohort, we normalize all alternatives in a nest with the same baseline alternative we write $\tilde{\omega}_{hkc(x)} \equiv \left(\omega_{hc(x)} - \omega_{h'c(x)}\right)/\lambda$.

In table 4.5, we display estimates of β in (4.25). One should be careful when interpreting these estimates as they are not the marginal effect of foreign students on the entry probabilities of natives in (4.26). This is because the scaling parameters λ_k are not identified from (4.24),[38] thus we can not estimate $\partial P_{hm}(x)/\partial V_{hm}(x)$ from estimates obtained from linear odds ratios. Nonetheless, β gives the direction of the effect and, sometimes, the scale is comparable across different sub-groups. For example, we can compare the scale of increasing the share of EU and non-EU students, but the scale across ability quartiles, institutions or any other discrete variable, is not comparable.[39]

$$\frac{dP_{hm}(x)}{d\bar{z}_{hmc(x)}} = \frac{\partial P_{hm}(x)}{\partial V_{hm}(x)}\beta \qquad (4.26)$$

The estimates in table 4.5 report bootstrap standard errors clustered by nest. We bootstrap to account for additional variation induced by estimation of the left hand side. In every bootstrap replication, therefore, we re-sample the individual level data, estimate the entry probabilities using bin estimators and regress the odds ratio on the foreign student share to estimate β in (4.25). We cluster by nest because the

[37] For ethnicity we divide students into white and non-white British.

[38] Without additional restrictions, in the linear transformation there is no variation to identify the multiplying parameter λ. For identification we will need to know the slope of a covariate or the variance of $\xi\lambda$.

[39] Imaging that we allow for different effects in a selective h and non-selective h' university. Then $\frac{dP_{hm}(x)}{d\bar{z}_{hmc(x)}} / \frac{dP_{h'm}(x)}{d\bar{z}_{h'mc(x)}} = \frac{\partial P_{hm}(x)}{\partial V_{hm}(x)}\left(\frac{\partial P_{h'm}(x)}{\partial V_{h'm}(x)}\right)^{-1}\frac{\beta_h}{\beta_{h'}}$.

odds ratio measurement error ξ is correlated for observations within the same nest by construction.[40]

Table 4.5: Probability of Entry Estimates

	All	All	Selective	Non-Selective
EU Peers	-1.125*	-0.856	-0.593	-1.291*
	(0.515)	(0.566)	(0.704)	(0.625)
non-EU Peers	-0.497	-0.585^{+}	-0.596	-0.483
	(0.310)	(0.319)	(0.475)	(0.362)
EU * $T_{(.25,.5]}$		-0.365		
		(0.319)		
EU * $T_{(.5,.75]}$		-0.600		
		(0.415)		
EU * $T_{(.75,1]}$		-0.262		
		(0.436)		
non-EU * $T_{(.25,.5]}$		-0.142		
		(0.184)		
non-EU * $T_{(.5,.75]}$		-0.359*		
		(0.143)		
non-EU * $T_{(.75,1]}$		0.688***		
		(0.155)		
Observations	125,130	125,130	41,040	84,100

Note: Sample of undergraduate students enrolled in 2007/10 in all English higher eduction institutions (HEI). Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Bootstrap standard errors from 100 replications are clustered by nest and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the first column of table 4.5 we display the average effect of increasing the share of EU and non-EU students, where both display a negative effect that, in line with estimates in table 4.4, is only statistically significant for EU students. In the second column of table 4.5, where $T_{(l,u]}$ is a dummy for whether the individual's ability is between the l and u percentile, we look at these estimates across the distribution of ability. Our estimates show that EU students have a negative effect everywhere. Although the increase on the number of parameters drives standard errors up making point estimates statistically non-significant. For non-EU we find a negative effect for everyone but those at the top of the ability distribution.

[40]To construct the odds ratio we take a baseline entry probability of a major-institution pair for each nest with which we divide all other probabilities in the nest. This induces a correlation among all observations in the nest because all ξ in the nest are a function of the measurement error of the baseline estimated probability.

In the last two columns of table 4.5 we replicate the first column but using the probabilities of entering major-institutions in the selective group, this is Russell and 1994 group, and other institutions, that we have termed non-selective. Remarkably, for selective institutions we find that the scale of the effect is homogeneous for both EU and non-EU students. On the other hand, for non-selective institutions, the effect displays an asymmetric scale between EU and non-EU. The parameter for EU is 2.67 times the estimate for non-EU. This further supports our findings in table 4.4 and goes in line with our framework in section 4.2, where we show that treating EU as home students creates an asymmetry on the effect of increasing the number of EU and non-EU students in non-selective universities.

4.6 Foreign Peers and Higher Education Outcomes

4.6.1 Baseline Estimates

In this section we present estimates of the effect of foreign peers on natives educational outcomes. We start estimating foreign peer effects on the probability of graduation and transition between STEM and non-STEM majors for native students, tables 4.6 and 4.15. In general, estimated effects are rather small, a one percentage point increase in foreign shares typically leads to non-significant changes and point estimates with magnitudes of .01 percentage points. We estimate negative, albeit small and not statistically significant, effects on the probability of graduation where this works through the probability of successful graduation, i.e. graduating with a first degree. The only graduation outcome for which we observe a larger, although still not statistically significant, effect is for the effect of EU peers on the probability of transitioning to a non-STEM major conditional on starting in a STEM, see table 4.15. For this, we estimate that a one percentage point increase in the share of EU peers leads to a .20 percentage point increase in the probability of transitioning from STEM to non-STEM majors.

Anelli, Shih, et al. (2017) report that increasing the share of international peers by one standard deviation decreases the probability of graduation in a STEM by 3 percentage points. Given that they report a foreign share standard deviation of .041, this is around a .83 percentage point decrease on the probability of graduating in STEM per percentage point increase on foreign shares. In column four of table 4.15 we replicate the result in Anelli, Shih, et al. (2017). We find that increasing the share of EU students by one percentage point lowers the probability of graduating in STEM by .237 percentage points. This is in the same direction as Anelli, Shih, et al. (2017) but the effect is smaller. Indeed, the lower 95% confidence bound of our estimate is at .591 percentage points decrease. Furthermore, we cannot reject a null of equality to zero.

Table 4.6: Higher Education Outcomes
Graduation and Transition

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	-0.004	-0.017	0.013	0.016	-0.012
	(0.027)	(0.028)	(0.018)	(0.012)	(0.021)
	EU vs Non-EU				
EU Peers	-0.002	-0.016	0.010	0.012	-0.010
	(0.039)	(0.052)	(0.038)	(0.024)	(0.044)
non-EU Peers	-0.005	-0.017	0.014	0.017	-0.012
	(0.031)	(0.034)	(0.020)	(0.013)	(0.023)
Observations	458,160	458,160	458,160	458,160	458,160

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In terms of grades, in table 4.7, we present foreign peer effect estimates on the probability of graduating with a first, upper second, lower second, third, lower qualification and failing to graduate.[41] Magnitudes of estimates are similar to those in tables 4.6 and 4.15. A one percentage point increase in foreign peer shares leads to changes in magnitudes of .01 percentage points on the probability of the different classifications. We estimate a negative, albeit not statistically significant, foreign peer effect from both EU and non-EU peers on the probability of graduating with an first and a positive non-significant effect on the probability of graduating with a lower qualification and failing to graduate.

For upper second, lower second and third the direction of the effect changes depending on whether we look at EU or non-EU peers. For EU peers, we find that a one percentage point increase in EU shares increases the probability of graduating with a third by .073 percentage points and this change is statistically significant. Still this effect is rather small, evaluated at the averages in table 4.3 a one percentage point increase in the share of EU peers maps into five additional EU students and the effect in table 4.7 leads to this five EU students increasing the number of natives graduating with a third by less than half of a student. These differences on the effects of EU and non-EU could be explained by the fact that EU and non-EU students enter HE through different channels as laid out in previous sections. This means that the average ability of these two groups of students might differ. Unfortunately, we do not have a measure of their pre-HE ability. Nevertheless, we can observe that overall EU students graduate with higher grades than native and non-EU students, see table 4.10. This provides suggestive evidence of EU students having a better academic

[41] Note that there are some majors, such as medicine, that do not use this classification scheme and therefore are not included.

preparation than the two latter groups which could explain why we observe different directions in the effects of EU and non-EU students on natives' grades.

Average foreign peer effects are, therefore, mild and typically non-significant. In the next section, as our framework in section 4.2 suggest, we explore whether this is due to heterogeneous effects across the ability distribution that compensate each other returning a zero average effect.[42]

Table 4.7: Higher Education Outcomes
Grades

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.027	-0.065	0.049	0.015	0.014	0.014
	(0.041)	(0.045)	(0.041)	(0.025)	(0.019)	(0.014)
	EU vs Non-EU					
EU Peers	-0.061	0.001	-0.035	0.073*	0.015	0.006
	(0.053)	(0.098)	(0.082)	(0.032)	(0.042)	(0.026)
non-EU Peers	-0.016	-0.088	0.079	-0.006	0.013	0.017
	(0.050)	(0.065)	(0.051)	(0.034)	(0.022)	(0.015)
Observations	416,010	416,010	416,010	416,010	416,010	416,010

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.6.2 Effects Across the Ability Distribution

In our framework we showed that heterogeneous effects across the ability distribution can appear even under fairly restrictive assumptions. This is supported by existing evidence in other peer effects. For example, Carrell, Fullerton, et al. (2009) find peer effects in college achievement that are heterogeneous across the ability distribution. We explore whether that is the case by regressing our higher education outcomes on foreign shares interacted with dummies for ability quartiles.

Our estimates in table 4.8 show significant effects on the probability of graduating and graduating successfully for native students that change across ability quantiles. For native students at the bottom of the ability distribution, an increase in the share of foreign peers has a positive effect on the probability of graduation and successful graduation, although the effect is statistically non-significant. However, the direction of the effect reverts and becomes significant for native students above the first ability quartile. For these an increase in the share of foreign students has a negative effect.

[42] In appendix 4.9.2, we investigate whether our zero effect findings masks some heterogeneity across different groups of students. For example, ethnic minorities might interact differently with foreign students compared to white British. We run our analysis separately for these two groups of native students. However, we do not find that there are significant results with respect to the ethnicity dimension. The stronger source of heterogeneity is found across the distribution of natives' ability and so we focus the discussion on this dimension in the next section.

Although it is rather small in magnitude. A one percentage point increase in the share of foreign peers decreases the probability of graduation by .014 percentage points for those in the second quartile, .02 for those in the third quartile and .027 for those at the top ability quartile. The decrease on the probability of graduation for students above the bottom ability quartile is mostly compensated by a positive effect on the probability of dropping out, that is, leaving without graduating before reaching the last year of the course.

For the effect on the probability of transitioning from non-STEM into STEM, table 4.22, our estimates are typically statistically non-significant and have magnitudes similar to those in table 4.6. If anything, non-EU students have a marginally negative effect on the probability of transitioning into non-STEM majors that changes sign for students in the second ability quartile.

Table 4.8: Graduation Effects Across Ability

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	0.040	0.053	-0.013	0.033*	-0.073*
	(0.034)	(0.038)	(0.023)	(0.015)	(0.031)
Foreign Peers * $T_{.25} - T_{.5}$	-0.054*	-0.069*	0.015	-0.023^{+}	0.077*
	(0.025)	(0.031)	(0.017)	(0.013)	(0.031)
Foreign Peers * $T_{.5} - T_{.75}$	-0.060**	-0.101***	0.041*	-0.027^{+}	0.087**
	(0.023)	(0.028)	(0.020)	(0.016)	(0.026)
Foreign Peers * $T_{.75} - T_{1}$	-0.067**	-0.116**	0.049*	-0.021^{+}	0.088**
	(0.024)	(0.036)	(0.022)	(0.012)	(0.029)
			EU vs Non-EU		
EU Peers	0.099	0.090	0.003	-0.009	-0.090
	(0.062)	(0.089)	(0.060)	(0.032)	(0.067)
EU Peers * $T_{.25} - T_{.5}$	-0.197**	-0.184*	-0.009	0.042^{+}	0.156*
	(0.068)	(0.085)	(0.056)	(0.025)	(0.072)
EU Peers * $T_{.5} - T_{.75}$	-0.046	-0.060	0.020	0.015	0.030
	(0.079)	(0.096)	(0.057)	(0.029)	(0.079)
EU Peers * $T_{.75} - T_{1}$	-0.174*	-0.178^{+}	0.007	0.039	0.135^{+}
	(0.081)	(0.095)	(0.057)	(0.035)	(0.071)
non-EU Peers	0.016	0.038	-0.019	0.050**	-0.065
	(0.040)	(0.053)	(0.027)	(0.018)	(0.041)
non-EU Peers * $T_{.25} - T_{.5}$	0.007	-0.020	0.025	-0.051**	0.044
	(0.035)	(0.050)	(0.023)	(0.019)	(0.045)
non-EU Peers * $T_{.5} - T_{.75}$	-0.060^{+}	-0.112*	0.049^{+}	-0.046*	0.106**
	(0.031)	(0.046)	(0.026)	(0.022)	(0.039)
non-EU Peers * $T_{.75} - T_{1}$	-0.028	-0.094	0.064*	-0.045*	0.072^{+}
	(0.036)	(0.058)	(0.029)	(0.019)	(0.043)
Observations	458,160	458,160	458,160	458,160	458,160

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In terms of grades we observe a negative effect of foreign peer shares on the probability of graduating with a first for native students at the bottom of the ability distribution. This effects attenuates as we move up on the ability distribution and becomes positive for natives in the third ability quartile. Furthermore, the effect is mostly driven by EU peers. A one percentage point increase in the EU share reduces the probability of graduating with a first by .162 percentage points for those at the bottom of the ability distribution and increases it by .014 percentage points for native students in the third ability quartile. In comparison the effect from non-EU peers are never greater than .09 and are always non statistically significant.

Where we find stronger effects is on the probability of graduating with an upper or lower second. On the probability of graduating with an upper second, a one percentage point increase in the share of foreign students produces a .105 increase for those at the lower ability quartile. The sign of the effect changes when we look at those right above in the ability distribution, for whom a one percentage point increase in the foreign share decreases the probability of graduating with an upper second by .004 percentage points. The effect increases further, to a .138 percentage point decrease, for those at the third quartile and even further, to a .196 percentage point decrease, for those at the top of the ability distribution. Exactly in the opposite direction, we observe that foreign peers have a negative, though non significant, effect on the probability of graduating with a lower second for natives at the bottom ability quartile. This effect changes sign as we move up on the ability distribution, leading to a one percentage point increase in foreign shares increasing the probability of graduating with a lower second by .131 percentage points for natives at the top of the ability distribution. Both effects on upper and lower second are most strongly driven by the effect of non-EU peers.

On the probability of graduating with a third we find no significant effects other than on the effect of EU peers on natives for which a one percentage point increase in the EU share increases the probability of graduating with a third by .099 percentage points. For lower qualification we find that foreign peers, driven by non-EU, increase the probability of graduation with a lower qualification for those at the third and top ability quartiles. Finally, on the probability of failing to graduate we find that increasing the share of foreign peers increases the probability of failing by .034 percentage points for each percentage point increase in the share, although the effect becomes almost zero for natives at the second ability quartile and up.

To conclude, overall, foreign peers affect natives at the higher tail of the ability distribution. These effects are generally negative, although their magnitude is modest to negligible. As already seen in the previous section, some effects are driven by EU students and others by non-EU students.

Table 4.9: Effects on Grades Across Ability

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.097	0.105^{+}	-0.041	0.024	-0.024	0.034^{+}
	(0.059)	(0.054)	(0.054)	(0.027)	(0.028)	(0.017)
Foreign Peers * $T_{.25} - T_{.5}$	0.071	-0.109*	0.047	-0.010	0.026	-0.025^{+}
	(0.072)	(0.052)	(0.033)	(0.012)	(0.022)	(0.014)
Foreign Peers * $T_{.5} - T_{.75}$	0.111*	-0.243***	0.124*	-0.017	0.055*	-0.030^{+}
	(0.046)	(0.042)	(0.056)	(0.012)	(0.027)	(0.016)
Foreign Peers * $T_{.75} - T_{1}$	0.096^{+}	-0.301***	0.172***	-0.008	0.065*	-0.023^{+}
	(0.053)	(0.055)	(0.051)	(0.015)	(0.030)	(0.012)
			EU vs Non-EU			
EU Peers	-0.162^{+}	0.071	-0.001	0.099*	0.014	-0.021
	(0.090)	(0.132)	(0.119)	(0.045)	(0.073)	(0.035)
EU Peers * $T_{.25} - T_{.5}$	0.045	-0.010	-0.009	-0.064	-0.014	0.051^{+}
	(0.106)	(0.123)	(0.111)	(0.041)	(0.068)	(0.029)
EU Peers * $T_{.5} - T_{.75}$	0.176^{+}	-0.015	-0.131	-0.052	0.003	0.019
	(0.102)	(0.131)	(0.134)	(0.042)	(0.070)	(0.033)
EU Peers * $T_{.75} - T_{1}$	0.170	-0.183	-0.041	0.011	-0.001	0.044
	(0.130)	(0.160)	(0.135)	(0.044)	(0.071)	(0.040)
non-EU Peers	-0.076	0.123	-0.058	-0.006	-0.039	0.055**
	(0.076)	(0.079)	(0.077)	(0.039)	(0.035)	(0.020)
non-EU Peers * $T_{.25} - T_{.5}$	0.082	-0.152^{+}	0.072	0.012	0.043	-0.056**
	(0.099)	(0.078)	(0.052)	(0.020)	(0.031)	(0.020)
non-EU Peers * $T_{.5} - T_{.75}$	0.086	-0.332***	0.223**	-0.003	0.076*	-0.050*
	(0.063)	(0.057)	(0.076)	(0.022)	(0.037)	(0.023)
non-EU Peers * $T_{.75} - T_{1}$	0.070	-0.354***	0.256***	-0.012	0.090*	-0.050*
	(0.072)	(0.072)	(0.068)	(0.027)	(0.039)	(0.019)
Observations	416,010	416,010	416,010	416,010	416,010	416,010

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the appendix, tables 4.23 to 4.31, we present further heterogeneity results by university groups. In Russell group universities increasing the share of EU students has a negative effect on the probability of graduating for students in the third and forth quartile. This works through decreasing the probability of successful graduation and increasing the probability of failing. In 1994 universities, however, we only observe a statistically significant effect for top ability students and this works through an increase on the probability of dropping out, rather than failing. Finally, in the rest of universities we observe that increasing the share of foreign peers decreases the probability of graduating and successful graduation for all native students, but most strongly those at the top of the ability distribution. These effects are compensated by positive foreign peer effects on the probability of graduating with a lower qualification and dropping out before graduation.

The effect on grades is also heterogeneous across universities. For example, in Russell Group universities increasing the foreign share has very small effects on the probability of obtaining a first. Although this hides substantial heterogeneity from

EU and non-EU. EU peers decrease the probability of obtaining a first for natives at the third and fourth quartile. Non-EU peers, on the other hand, increase the probability of graduating with a first for the same group of native students. In 1994 universities and other universities we find no statistically signifiant effects on the probability of graduating with a first, other than a marginally significant negative effect of EU peers on top ability natives in 1994 universities. Finally, the foreign peer effect pushing high ability natives from upper to lower second is mostly driven by the effect of foreign peers on natives in other universities.

4.6.3 Grading on a Curve

Our results in section 4.6.2 show that foreign peers displace native students across grades and that the effect depends on the position of the student in the ability distribution. This could be explained by two main channels, through interaction among natives and foreign students and through grading on a curve. The first channel embodies several mechanisms laid out in the introduction and in section 4.2. Grading on a curve, on the other hand, consists on teachers grading with a curve in mind, leading to mechanical foreign peer effects on grades that are given by differences between foreign and native student quality alone. It is important to disentangle these two channels as the first one means that foreign students affect human capital accumulation of natives, while the second, grading on a curve, is purely mechanical.

We do not have any information on the type of interaction that students have, but we can formally test the grading on the curve mechanism. This allows us to make inference on the effect of foreign students in the actual acquisition of human capital of natives. Let us explore this point by introducing a simple model. In a group with a foreign share π, true grades (X) follow a distribution $F_X(X) = F_X(X|Native)(1-\pi) + F_X(X|Foreign)\pi$. Teachers have some target in mind of how the distribution of grades should look (F_Y) and instead of reporting true grades they report a transformation that has the desired distribution, i.e. $Y \equiv F_Y(F_X^{-1}(X))$.

A straight forward implication of the grading on a curve mechanism is that changes in the distribution of true grades do not modify the distribution of reported grades. However, if we condition on being a native, grading on a curve leads to a mechanical effect on reported grades of natives. This implies that one can easily test whether the effect on grades is driven by a grading on a curve mechanism by testing if the effect of foreign shares in the whole population, including foreign students, is zero.[43] However, we cannot produce this test because we lack information on pre-higher education ability for foreigners and there is selection at entry in terms of ability (see section 4.4.2) and heterogeneous peer effects on grades across the ability distribution (see section 4.6.2). Instead, we explore the grading on a curve mechanism to provide us with further implications that we can test with our data.

[43] A stronger test will seek rejection of the null of no effects on multiple moments, not just the first one.

Let's build from what we have. In the data, we only observe a categorization of the grade; a student is awarded a first if her reported grade is above 70%, an upper second if between 60-70%, a lower second if between 50-60%, a third if in 40-50%, a lower qualification if below 40% but not too far away and fails to graduate otherwise. Then we can easily recast the implications of grading on a curve in terms of probabilities of grades that we can estimate from the data. For example, the probability of been awarded a first for a native at the top ability quartile, $T_{(.75,1]}$, is

$$\begin{aligned}\Pr(First|Native, T_{(.75,1]}) &= \int_{.7} dF_Y(y|Native, T_{(.75,1]}) \\ &= 1 - F_X\left(F_X^{-1}(F_Y(.7))|Native, T_{(.75,1]}\right)\end{aligned}$$

implying that the effect of increasing the share of foreigners when there is grading on a curve is

$$\frac{\partial \Pr(First|Native, T_{(.75,1]})}{\partial \pi} = -f_X\left(F_X^{-1}(F_Y(.7))|Native, T_{(.75,1]}\right)\frac{\partial F_X^{-1}(F_Y(.7))}{\partial \pi}$$

$$\frac{\partial F_X^{-1}(F_Y(.7))}{\partial \pi} = -\frac{F_X\left(F_X^{-1}(F_Y(.7))|Foreign\right) - F_X\left(F_X^{-1}(F_Y(.7))|Native\right)}{f_X\left(F_X^{-1}(F_Y(.7))\right)} \tag{4.27}$$

At first sight this expression may seem helpless in terms of testable implications, as we do not know or can directly infer the distribution of true grades. However, closer inspection reveals that actually we can learn the direction of the effect from our data. The direction is determined by the difference in cumulative distributions between foreigners and natives at the transformed threshold, or more simply, the difference between the observed probability of graduating with a first for native and foreign students

$$\begin{aligned}\frac{\partial \Pr(First|Native, T_{(.75,1]})}{\partial \pi} =& \frac{f_X\left(F_X^{-1}(F_Y(.7))|Native, T_{(.75,1]}\right)}{f_X\left(F_X^{-1}(F_Y(.7))\right)} \\ &\times [\Pr(First|Native) - \Pr(First|Foreign)]\end{aligned} \tag{4.28}$$

This determination of the sign of the effect by the differences in observed probabilities happens for the top (first) and lower (fail) classifications. For classifications in between, e.g. upper second, the direction of the effect also depends on the density of true grades and there can be changes on the direction across ability quartiles[44]

[44]Or any other conditioning.

$$
\begin{aligned}
\frac{\partial \Pr(UpperSecond|Native, T_{(.75,1]})}{\partial \pi} = &- \frac{f_X\left(F_X^{-1}(F_Y(.7))|Native, T_{(.75,1]}\right)}{f_X\left(F_X^{-1}(F_Y(.7))\right)} \\
&\times \left[\Pr(First|Native) - \Pr(First|Foreign)\right] \\
&+ \frac{f_X\left(F_X^{-1}(F_Y(.6))|Native, T_{(.75,1]}\right)}{f_X\left(F_X^{-1}(F_Y(.6))\right)} \\
&\times \left[F_X(F_X^{-1}(F_Y(.6))|Foreign) - F_X(F_X^{-1}(F_Y(.6))|Native)\right]
\end{aligned}
$$

Given that we cannot directly learn the distribution of true grades, foreign effects on probabilities of grades between first and failure are uninformative in terms of whether the mechanism is grading on a curve. However, we can test the implications of the grading on a curve mechanism by looking at foreign peer effects on the probabilities of obtaining a first and failing to pass the course.

In table 4.10, we present the distribution of grades for natives, foreigners and foreigners differentiating between EU and non-EU. Foreign students are over represented at the top of the grade distribution, 14.4% of natives graduate with a first while the probability of graduation with a first for natives is 16.4%, rising to 20.3% if we consider EU alone and falling just below the native probability if we consider non-EU only. This over-representation of foreign students at the top of the grading distribution holds for both STEM and non-STEM majors, although non-EU are only over-represented in non-STEM. These differences are all statistically significant, although, in the case of the average probability of graduating at the top, the difference with non-EU is only significant at the 90% level.

On average, foreigners have a lower probability of failing, but this changes if we look at STEM and non-STEM majors. For STEM majors, foreigners, both EU and non-EU, have a lower probability of failure than natives and the differences are statistically significant. For non-STEM the probability of failing is higher for foreigners, although differences are only significant, at 90%, for non-EU.

Table 4.10: Distribution of Grades

	Distribution				P-value Null of Equality		
	UK	Foreign	EU	Non-EU	Foreign	EU	Non-EU
				All			
First	0.144	0.164	0.203	0.140	0.000	0.000	0.097
Upper Second	0.526	0.449	0.476	0.432	0.000	0.000	0.000
Lower Second	0.234	0.269	0.220	0.299	0.000	0.000	0.000
Third	0.037	0.068	0.047	0.081	0.000	0.000	0.000
Lower Qual	0.040	0.033	0.039	0.030	0.000	0.320	0.000
Fail	0.019	0.017	0.016	0.018	0.044	0.019	0.437
				Non-STEM			
First	0.160	0.186	0.221	0.169	0.000	0.000	0.057
Upper Second	0.477	0.378	0.406	0.363	0.000	0.000	0.000
Lower Second	0.246	0.283	0.244	0.303	0.000	0.791	0.000
Third	0.047	0.089	0.063	0.103	0.000	0.000	0.000
Lower Qual	0.046	0.038	0.042	0.036	0.000	0.224	0.000
Fail	0.023	0.025	0.024	0.026	0.101	0.745	0.070
				STEM			
First	0.137	0.155	0.197	0.127	0.000	0.000	0.000
Upper Second	0.545	0.478	0.499	0.463	0.000	0.000	0.000
Lower Second	0.230	0.263	0.211	0.298	0.000	0.000	0.000
Third	0.034	0.059	0.041	0.071	0.000	0.000	0.000
Lower Qual.	0.038	0.032	0.038	0.028	0.000	0.830	0.000
Fail	0.017	0.014	0.013	0.014	0.000	0.005	0.009

Note: Observations from 2007 onwards only.

Using our previous discussion and the evidence in 4.10, it follows that, if the effect is only given by grading on a curve, then we should see that foreign students lower the probability of graduating with a first for natives in both STEM and non-STEM degrees and that the effect is stronger for EU than non-EU. For the probability of failing we should observe a positive effect in STEM and a negative effect in non-STEM that should be larger for non-EU.

In table 4.11 we find a positive effect of foreign peers on the probability of failing to graduate that is strongest for those at the bottom of the ability distribution. Furthermore, this positive effect is driven by non-EU peers alone: a one percentage point increase in the non-EU foreign peer share increases the probability of failing to graduate by .06 percentage points for native students at the bottom of the ability distribution, .003 percentage points for those in the second ability quartile, .012 percentage points for those in the third quartile and .007 percentage points for those at the top of the ability distribution. This positive effect is inconsistent with the grading on a curve mechanism, non-EU students in non-STEM majors have a larger probability of failure than natives and, therefore, should decrease the probability of failing for native students, not increase it.

Additionally, if we look at the effects of EU peers on the probability of graduating with a first in non-STEM majors, we observe that EU students have a negative effect at the bottom of the ability distribution that decreases for students at the second and top ability quartile and becomes positive for native students in the third quartile. This crossing in terms of direction of the effect is not consistent with the grading on a curve mechanism. If we inspect (4.28) it is clear that the magnitude of the effect can change across different groups of natives but the direction must be always the same.

Table 4.11: Effects on Grades
Non-STEM

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.101	0.092	-0.036	0.025	-0.020	0.040*
	(0.069)	(0.063)	(0.067)	(0.032)	(0.031)	(0.017)
Foreign Peers * $T_{.25} - T_{.5}$	0.056	-0.078	0.050	-0.015	0.015	-0.028*
	(0.075)	(0.049)	(0.038)	(0.013)	(0.026)	(0.013)
Foreign Peers * $T_{.5} - T_{.75}$	0.103*	-0.222***	0.121^{+}	-0.019	0.049^{+}	-0.032^{+}
	(0.049)	(0.049)	(0.066)	(0.014)	(0.029)	(0.016)
Foreign Peers * $T_{.75} - T_{1}$	0.098^{+}	-0.266***	0.153*	-0.013	0.056^{+}	-0.029**
	(0.056)	(0.057)	(0.061)	(0.015)	(0.031)	(0.011)
	EU vs Non-EU					
EU Peers	-0.168^{+}	0.090	-0.063	0.087^{+}	0.070	-0.016
	(0.094)	(0.152)	(0.151)	(0.047)	(0.074)	(0.036)
EU Peers * $T_{.25} - T_{.5}$	0.082	-0.036	0.002	-0.043	-0.053	0.049
	(0.106)	(0.131)	(0.126)	(0.044)	(0.077)	(0.032)
EU Peers * $T_{.5} - T_{.75}$	0.197^{+}	0.011	-0.110	-0.048	-0.062	0.013
	(0.111)	(0.142)	(0.155)	(0.044)	(0.078)	(0.035)
EU Peers * $T_{.75} - T_{1}$	0.230	-0.145	-0.042	-0.003	-0.081	0.042
	(0.144)	(0.180)	(0.151)	(0.053)	(0.075)	(0.042)
non-EU Peers	-0.075	0.098	-0.031	0.000	-0.056	0.064**
	(0.089)	(0.087)	(0.097)	(0.047)	(0.042)	(0.021)
non-EU Peers * $T_{.25} - T_{.5}$	0.044	-0.099	0.073	-0.003	0.046	-0.061**
	(0.102)	(0.072)	(0.065)	(0.023)	(0.038)	(0.020)
non-EU Peers * $T_{.5} - T_{.75}$	0.065	-0.312***	0.211*	-0.008	0.095*	-0.052*
	(0.068)	(0.065)	(0.097)	(0.026)	(0.038)	(0.023)
non-EU Peers * $T_{.75} - T_{1}$	0.049	-0.319***	0.232*	-0.014	0.110**	-0.057**
	(0.076)	(0.082)	(0.089)	(0.031)	(0.042)	(0.018)
Observations	297,840	297,840	297,840	297,840	297,840	297,840

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Inspection of the foreign effects on grade probabilities for those in STEM majors (table 4.12) produces no evidence to reject the grading on a curve hypothesis. Given the probabilities of graduating with a first and failing in table 4.10, we should observe that foreign peers decrease the probability of graduating with a first, with the negative effect driven by EU students while non-EU students have an effect on the opposite direction. Furthermore, both non-EU and EU students should increase the probability of failing. Instead, we observe non-significant effects for both the probability of graduating with a first and failing across all ability quartiles and foreign student region of origin. This is not inconsistent with the grading on a curve mechanism as it could be the case that the density ratio in (4.28) multiplying the difference in probabilities attenuates the magnitude of the effect leading to differences that are not strong enough to be picked by the significance test.

For students in STEM majors, we observe a positive effect on the probability of graduating with an upper second for natives at the bottom of the ability distribution that attenuates as we move up on the ability distribution and becomes negative for top ability students. For the probability of graduating with a lower second we observe the opposite pattern: a negative effect for those at the bottom of the ability distribution that attenuates as we move up on the ability distribution and becomes positive for top ability students. Given the information we have, these effects at the middle of the grade distribution do not produce evidence to reject the grading on a curve hypothesis. Form (4.6.3) it is evident that without knowing the density of true grades the direction of the effect at the middle of the grade distribution is ex-ante unclear.

In addition, the evidence in table 4.26 is also inconsistent with the grading on a curve mechanism. Our estimates in table 4.26 show that increasing the share of foreign peers has a negative effect on the probability of failing for low ability native students, for whom a one percentage point increase on the share of foreign peers decreases the probability of failing by .039 percentage points. However, for all other ability groups a one percentage point increase on the share of foreign students increases the probability of failing by around .032 percentage points. This crossing on the direction of the effect is inconsistent with the grading on a curve mechanism.

Therefore, the evidence we provide rules out that grading on a curve is the only mechanism behind the effects we estimate. However, this is not to say that there is no grading on a curve. Our evidence would be consistent with a situation where grading on a curve and foreign peers affecting the human capital accumulation of natives coexist.

Table 4.12: Effects on Grades
STEM

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.058	0.428**	-0.309**	-0.018	-0.064	0.022
	(0.104)	(0.159)	(0.102)	(0.046)	(0.056)	(0.045)
Foreign Peers * $T_{.25} - T_{.5}$	0.158	-0.352***	0.104	0.018	0.073*	-0.001
	(0.099)	(0.078)	(0.071)	(0.036)	(0.031)	(0.025)
Foreign Peers * $T_{.5} - T_{.75}$	0.148	-0.413***	0.218**	-0.006	0.065^{+}	-0.013
	(0.091)	(0.100)	(0.067)	(0.045)	(0.036)	(0.024)
Foreign Peers * $T_{.75} - T_{1}$	0.104	-0.562***	0.358***	0.001	0.090*	0.009
	(0.090)	(0.108)	(0.076)	(0.040)	(0.040)	(0.025)
			EU vs Non-EU			
EU Peers	0.073	0.471^{+}	-0.195	0.077	-0.393***	-0.033
	(0.221)	(0.268)	(0.246)	(0.098)	(0.115)	(0.100)
EU Peers * $T_{.25} - T_{.5}$	-0.130	-0.144	0.090	-0.146^{+}	0.234*	0.096
	(0.202)	(0.184)	(0.203)	(0.084)	(0.112)	(0.074)
EU Peers * $T_{.5} - T_{.75}$	0.092	-0.264	-0.160	-0.062	0.334**	0.061
	(0.201)	(0.252)	(0.245)	(0.098)	(0.113)	(0.081)
EU Peers * $T_{.75} - T_{1}$	0.027	-0.572*	0.071	0.027	0.383**	0.065
	(0.234)	(0.281)	(0.302)	(0.092)	(0.125)	(0.089)
non-EU Peers	-0.089	0.416*	-0.328**	-0.042	0.009	0.034
	(0.116)	(0.187)	(0.123)	(0.049)	(0.052)	(0.046)
non-EU Peers * $T_{.25} - T_{.5}$	0.232	-0.407***	0.103	0.061	0.037	-0.025
	(0.140)	(0.092)	(0.075)	(0.039)	(0.028)	(0.030)
non-EU Peers * $T_{.5} - T_{.75}$	0.155	-0.451***	0.324***	0.005	-0.002	-0.031
	(0.138)	(0.125)	(0.084)	(0.051)	(0.040)	(0.037)
non-EU Peers * $T_{.75} - T_{1}$	0.118	-0.557***	0.437***	-0.012	0.017	-0.004
	(0.131)	(0.121)	(0.081)	(0.044)	(0.037)	(0.039)
Observations	118,170	118,170	118,170	118,170	118,170	118,170

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.7 Foreign Peers and Labour Market Outcomes

Here we present estimates of foreign peer effects on early labour market outcomes obtained from the destination of leaver from higher education (DLHE) survey. The DLHE survey contacts students six months after graduation and asks a set of questions about current status, including labour market participation and characteristics of job if the student is working.

Even though the survey does quite well in terms of response, with 80% response rate on average, 75% in our selected sub-population; one may be concerned about the representativeness of the sample. We address possible representativeness concerns by producing estimates of the effect of foreign peers on the distribution of grades using the DLHE matched and non-matched sub-samples.[45] The estimates we obtain from

[45] As the DLHE survey is about graduates only, our estimates are conditional on having successfully graduated from an undergraduate degree.

both sub-samples, tables 4.16 and 4.17, display similar coefficients with significant effects always in the same direction and similar magnitudes. We take this as supporting evidence for the representativeness of the matched sub-sample.

In terms of labour market outcomes, in table 4.13 we display estimates of foreign peer effects on the probability of labour market participation. In general the magnitude of the estimated effects is similar to the ones we find for higher education outcomes. For example, there is a negative effect on the probability of working after graduation that becomes positive for native students at the third and top ability quartile, implying that a one percentage point increase in foreign shares increases the probability of working for students at the top of the ability quartile by .064 percentage points. These effects are most strongly driven by EU students for those natives at the middle of the ability distribution and by non-EU peers for natives at the top of the ability distribution. Furthermore, the effect works mostly through decreasing the probability of continuing studying rather than the probability of employment. This is clear from the third column in table 4.13 where we estimate the effect on the probability of working conditional on entering the labour market and the effect dissipates. Or even clearer in table 4.18, where we estimate that a one percentage point increase in foreign shares decreases the probability of studying by .044 percentage points for those at the top of the ability distribution and the effect works by reducing the probability of going into postgraduate studies.

These results are consistent with those on educational outcomes. Foreign peers negatively affect the performance of natives in their degree and this negatively affects the probability of natives of keep studying in postgraduate courses and induce them to enter the labour market. Also, those native students at the middle-top of the ability distribution are those driving the results, again consistent with the findings on higher eduction outcomes.

Table 4.13: Labour Market Participation

	Working	Working (& work.+stud.)	Working vs. Unemployed	Unemployed
Foreign Peers	-0.044	-0.085	0.015	-0.003
	(0.114)	(0.108)	(0.085)	(0.067)
Foreign Peers * $T_{.25} - T_{.5}$	0.045	0.013	0.020	-0.015
	(0.036)	(0.060)	(0.030)	(0.023)
Foreign Peers * $T_{.5} - T_{.75}$	0.102*	0.085	-0.001	0.015
	(0.044)	(0.064)	(0.034)	(0.025)
Foreign Peers * $T_{.75} - T_{1}$	0.084	0.097	0.034	-0.018
	(0.062)	(0.103)	(0.045)	(0.029)
	EU vs Non-EU			
EU Peers	-0.322	-0.267	-0.125	0.116
	(0.219)	(0.233)	(0.185)	(0.144)
EU Peers * $T_{.25} - T_{.5}$	0.242^{+}	0.293*	0.180^{+}	-0.122
	(0.128)	(0.127)	(0.107)	(0.086)
EU Peers * $T_{.5} - T_{.75}$	0.459**	0.367*	0.122	-0.063
	(0.152)	(0.165)	(0.137)	(0.113)
EU Peers * $T_{.75} - T_{1}$	0.347^{+}	0.257	0.125	-0.084
	(0.184)	(0.161)	(0.147)	(0.121)
non-EU Peers	0.061	-0.002	0.069	-0.047
	(0.133)	(0.142)	(0.102)	(0.079)
non-EU Peers * $T_{.25} - T_{.5}$	-0.042	-0.105	-0.047	0.030
	(0.060)	(0.085)	(0.039)	(0.031)
non-EU Peers * $T_{.5} - T_{.75}$	-0.040	-0.031	-0.054	0.048
	(0.061)	(0.088)	(0.056)	(0.042)
non-EU Peers * $T_{.75} - T_{1}$	-0.026	0.019	-0.008	0.011
	(0.088)	(0.132)	(0.070)	(0.049)
Observations	146,460	146,460	111,230	146,460

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, in table 4.14, we present estimates of the effect of foreign peers on job characteristics of those that work. We observe that, conditional on working, increasing the share of foreign students by one percentage point leads to .28 percentage point increase in the probability of working in a professional occupations.[46] With this effect mostly driven by non-EU students: a one percentage point increase in the non-EU share leads to a .352 percentage point increase in the probability of working in a professional occupation.

Across the other job attributes we consider, i.e. graduate job and earnings, we do not find any significant effect of foreign peers. Only by looking at EU and non-EU peers separately we find that a one percentage point increase in non-EU share has a marginally significant effect on the probability of working in a graduate job, and EU shares increase log wages by, around, .17% per percentage point increase for those at the third and top ability quartiles, although the effect is only significant for those at

[46] The effect increases for those above the lower quantile but the increase is not statistically significant.

the third quartile. Thus, as in the case of higher education outcomes, there are only mild effects of foreign students on native labour market outcomes.

Table 4.14: Job Attributes

	Professional occ.	Graduate job	(log)Salary	(log)Salary imp.
Foreign Peers	0.273	0.166	-0.188	-0.015
	(0.174)	(0.147)	(0.114)	(0.143)
Foreign Peers * $T_{.25} - T_{.5}$	0.038	-0.051	-0.010	0.047
	(0.096)	(0.074)	(0.043)	(0.039)
Foreign Peers * $T_{.5} - T_{.75}$	0.103	-0.054	0.064	0.100
	(0.092)	(0.104)	(0.064)	(0.074)
Foreign Peers * $T_{.75} - T_{1}$	0.102	0.052	0.113^{+}	0.145^{+}
	(0.088)	(0.108)	(0.063)	(0.079)
	EU vs Non-EU			
EU Peers	-0.283	-0.270	-0.356	-0.077
	(0.276)	(0.295)	(0.243)	(0.244)
EU Peers * $T_{.25} - T_{.5}$	0.089	0.114	-0.116	0.151
	(0.174)	(0.162)	(0.122)	(0.143)
EU Peers * $T_{.5} - T_{.75}$	0.644**	0.271	0.226	0.207
	(0.226)	(0.224)	(0.137)	(0.164)
EU Peers * $T_{.75} - T_{1}$	0.500*	0.308	0.376*	0.355^{+}
	(0.235)	(0.273)	(0.173)	(0.195)
non-EU Peers	0.438*	0.299^{+}	-0.151	0.012
	(0.182)	(0.154)	(0.123)	(0.157)
non-EU Peers * $T_{.25} - T_{.5}$	0.009	-0.120	0.029	0.001
	(0.128)	(0.106)	(0.066)	(0.070)
non-EU Peers * $T_{.5} - T_{.75}$	-0.103	-0.177	0.004	0.053
	(0.122)	(0.135)	(0.086)	(0.118)
non-EU Peers * $T_{.75} - T_{1}$	-0.054	-0.047	0.023	0.067
	(0.125)	(0.159)	(0.090)	(0.121)
Observations	86,860	77,430	40,880	51,710

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis.Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.8 Conclusion

With the number of international students steadily rising, understanding what is the effect of foreign peers on native students is key to get the full picture of immigrant effects.

We use data on the universe of English higher education and exploit cross-cohort variation within university and major to estimate the effect of foreign peers on educational and early labour market outcomes of native students. We show that this cross-cohort variation on foreign shares correlates positively with average ability of incoming native students, with the effect driven by EU students. This is because of both native and EU students being subjected to the same funding constraint caps in higher eduction.[47] Furthermore, we show that this asymmetry in the treatment of international students causes that increasing the share of non-EU shares rises the total number of students while increasing the share of EU does not modify group size. However, this changes for selective universities in which both EU and non-EU students displace natives. We show that this follows from imposing asymmetric funding constraints on EU and non-EU students. In selective universities, both EU and non-EU students display natives because these universities select the best British, EU and non-EU students. Non-selective universities, on the other hand, use non-EU students to fill capacity once their funding constraint for EU and British binds. Thus increasing the number of non-EU does not displace native students.

Student selection by universities in terms of ability as a function of the number of international students creates a challenge for identification of the effect of foreign peers on native students. However, we are able to eliminate the selection effect by controlling for pre-higher education ability. For this, we use the UCAS tariff score, which is what universities observe in students UCAS application and the main characteristic that they use when selecting among the pool of prospective students. To support that controlling for ability gives us quasi-random variation, we show that cross-cohort variation on peer-shares is uncorrelated with a comprehensive set of predetermined individual characteristics after controlling for ability.

In general, we find mild effects. A one percentage point increase in foreign shares, typically, doesnt change the probability of an outcome by more than half of a percentage point. Even more, for the average student we find no significant effects on any of our higher education outcomes: graduating, dropping out, changing between STEM and non STEM majors, graduating with a first, graduating with an upper second, graduating with a lower second, graduating with a third, graduating with a lower qualification or failing to graduate.

When we look at foreign effects across the distribution of ability, we find that foreign peers decrease the probability of graduating for native students that are not at the bottom of the ability distribution. For these students, a one percentage point increase in the share of foreign peers decreases the probability of graduating by .014-.027 percentage points, depending on where in the ability distribution the student is. This change on the probability of graduating works mostly through foreigners increasing the probability of dropping out for natives above the bottom quartile of ability. For these, a one percentage point increase in foreign shares increases the probability of dropout by .004-.015 percentage points.

[47] While non-EU students are only subjected to visa and university capacity restrictions.

Moreover, we also estimate a heterogeneous effect across the ability distribution on grades, where the strongest effects are on the probabilities of graduating with an upper or lower second. Increasing the share of foreign students increases the probability of graduating with a lower second and decreases the probability of graduating with an upper second for native students at the second ability quartile and up. By comparing the grade distribution for native and foreign students in STEM and non-STEM majors, we show that our estimated effects are not consistent with a grading on a curve mechanism. This is of policy relevance as, under grading on a curve, differences in the quality of foreign and native students mechanically displace native grades. Therefore, rejection of the grading on a curve hypothesis suggest that foreign students cause mild changes on the human capital acquisition of native students.

Finally, we explore whether these changes translate into effects on labour market outcomes six months after graduation. We find mild effects. Increasing the foreign share by a one percentage point increases the probability of working by .064 percentage points for native students at the top of the ability distribution. Mostly by driving them out of postgraduate education. In terms of job attributes, we find that foreign peers increase the probability of working in a professional occupation for all native students, have no effect on the probability of holding a graduate job and increase earnings for natives at the third and top ability quantiles by around .17%. Although the effect is only statistically significant for those at the third ability quartile.

Overall, we do not find that migration in the context of higher eduction importantly affects the educational outcomes of natives. Similar to some effects of immigrants on the labour market, we find that there are only effects for certain groups of natives, which usually are those sharing similar characteristics to migrants (Dustmann, Frattini, et al. 2013; Peri and Sparber 2009). Indeed, in our case, foreign students, that show higher educational performances, affect mainly natives at the top of the ability distribution.

Our paper is the first, as we are aware of, in shedding light on foreign peer effects in higher education by considering the whole universe of undergraduate students. Other papers (Anelli, Shih, et al. 2017; Chevalier et al. 2019) focus on single universities and courses. Despite providing richer information compared to what we can do in terms of mechanisms underlying the results, there may be external validity concerns. We add to them as we provide the overall effect of foreign students on higher education.

4.9 Appendix

4.9.1 EU A10 Accession

Figure 4.3: A10 Students Evolution

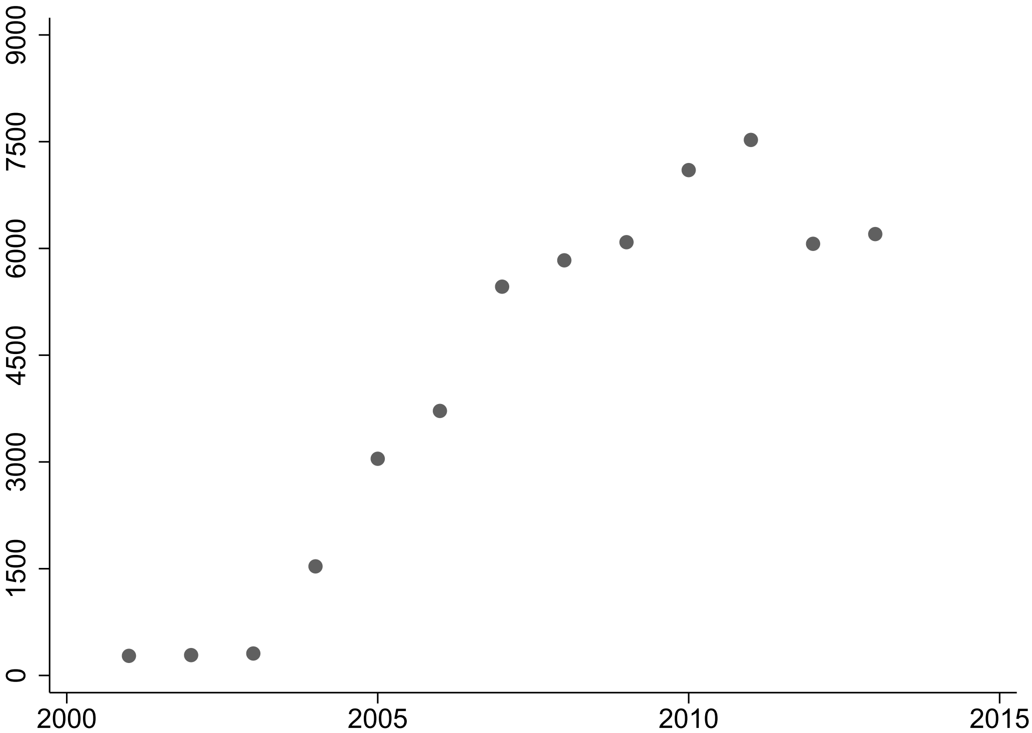
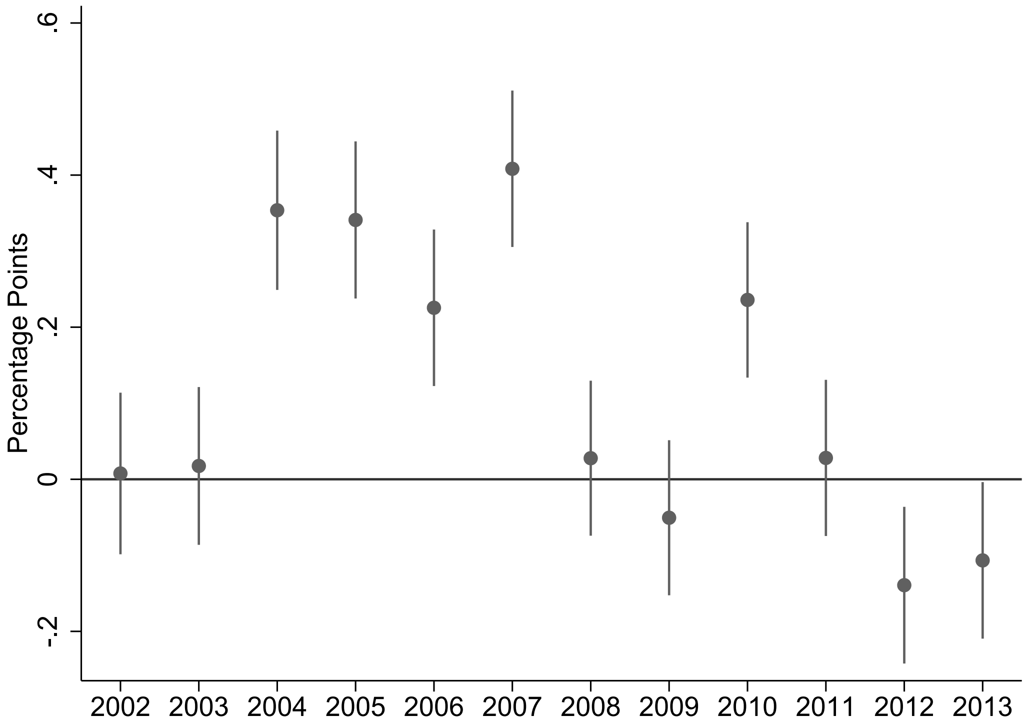


Figure 4.4: A10 Students Evolution
Between Cohorts Changes



4.9.2 Additional Tables

Table 4.15: STEM and non-STEM Transitions

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	0.006	-0.009	-0.007	0.021
	(0.029)	(0.041)	(0.045)	(0.072)
	EU vs Non-EU			
EU Peers	-0.028	0.037	0.201	-0.237
	(0.053)	(0.069)	(0.136)	(0.181)
non-EU Peers	0.018	-0.025	-0.061	0.087
	(0.032)	(0.047)	(0.051)	(0.073)
Observations	326,990	326,990	131,170	131,170

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.16: Grades
(*DLHE Matched Sub-Sample*)

	First	Upper Second	Lower Second	Third	Lower Qual.
Foreign Peers	-0.189^{+}	0.183	-0.019	0.057	-0.032
	(0.108)	(0.132)	(0.131)	(0.061)	(0.037)
Foreign Peers * $T_{.25} - T_{.5}$	0.053	-0.176**	0.122^{+}	-0.028	0.029
	(0.072)	(0.056)	(0.066)	(0.037)	(0.025)
Foreign Peers * $T_{.5} - T_{.75}$	0.124	-0.383***	0.245^{+}	-0.022	0.035
	(0.081)	(0.101)	(0.130)	(0.023)	(0.027)
Foreign Peers * $T_{.75} - T_{1}$	0.085	-0.381***	0.280**	-0.022	0.039
	(0.089)	(0.097)	(0.106)	(0.035)	(0.025)
			EU vs Non-EU		
EU Peers	-0.550**	0.385	0.174	0.070	-0.078
	(0.196)	(0.280)	(0.232)	(0.091)	(0.103)
EU Peers * $T_{.25} - T_{.5}$	0.122	0.233	-0.251	-0.093	-0.012
	(0.113)	(0.179)	(0.189)	(0.079)	(0.083)
EU Peers * $T_{.5} - T_{.75}$	0.311*	-0.036	-0.261	-0.009	-0.005
	(0.157)	(0.250)	(0.253)	(0.070)	(0.082)
EU Peers * $T_{.75} - T_{1}$	0.261	-0.166	-0.168	0.059	0.013
	(0.202)	(0.269)	(0.259)	(0.081)	(0.078)
non-EU Peers	-0.075	0.164	-0.113	0.047	-0.023
	(0.132)	(0.152)	(0.175)	(0.074)	(0.046)
non-EU Peers * $T_{.25} - T_{.5}$	0.022	-0.349***	0.284**	-0.002	0.046
	(0.103)	(0.099)	(0.103)	(0.058)	(0.033)
non-EU Peers * $T_{.5} - T_{.75}$	0.051	-0.527**	0.451*	-0.025	0.051
	(0.107)	(0.157)	(0.179)	(0.044)	(0.033)
non-EU Peers * $T_{.75} - T_{1}$	0.018	-0.487**	0.465**	-0.046	0.051
	(0.117)	(0.145)	(0.153)	(0.053)	(0.033)
Observations	145,450	145,450	145,450	145,450	145,450

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007-2008. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.17: Grades
(*DLHE Non-Matched Sub-Sample*)

	First	Upper Second	Lower Second	Third	Lower Qual.
Foreign Peers	0.103	-0.126	-0.020	0.141	-0.098
	(0.099)	(0.208)	(0.135)	(0.117)	(0.121)
Foreign Peers * $T_{.25} - T_{.5}$	0.008	-0.204**	0.161	-0.034	0.069
	(0.064)	(0.062)	(0.111)	(0.078)	(0.063)
Foreign Peers * $T_{.5} - T_{.75}$	0.101^{+}	-0.197***	0.108	-0.122	0.110
	(0.057)	(0.055)	(0.066)	(0.078)	(0.067)
Foreign Peers * $T_{.75} - T_{1}$	0.083	-0.362***	0.225***	-0.091	0.144^{+}
	(0.059)	(0.065)	(0.065)	(0.082)	(0.082)
			EU vs Non-EU		
EU Peers	0.048	-0.374	0.440	0.034	-0.148
	(0.255)	(0.329)	(0.348)	(0.201)	(0.243)
EU Peers * $T_{.25} - T_{.5}$	-0.080	-0.114	0.409*	-0.065	-0.150
	(0.145)	(0.242)	(0.202)	(0.131)	(0.202)
EU Peers * $T_{.5} - T_{.75}$	0.022	0.530*	-0.140	-0.225	-0.187
	(0.219)	(0.254)	(0.291)	(0.168)	(0.196)
EU Peers * $T_{.75} - T_{1}$	0.282	0.081	-0.322	-0.073	0.032
	(0.272)	(0.319)	(0.264)	(0.153)	(0.205)
non-EU Peers	0.108	-0.023	-0.137	0.164	-0.111
	(0.098)	(0.225)	(0.162)	(0.149)	(0.147)
non-EU Peers * $T_{.25} - T_{.5}$	0.047	-0.251*	0.057	-0.020	0.167*
	(0.121)	(0.119)	(0.168)	(0.116)	(0.082)
non-EU Peers * $T_{.5} - T_{.75}$	0.130	-0.459***	0.189	-0.086	0.226*
	(0.088)	(0.122)	(0.124)	(0.119)	(0.090)
non-EU Peers * $T_{.75} - T_{1}$	0.031	-0.535***	0.389**	-0.090	0.205*
	(0.098)	(0.133)	(0.116)	(0.119)	(0.103)
Observations	46,780	46,780	46,780	46,780	46,780

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007-2008. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.18: Studying after Graduation

	Studying	Studying postgrad.	Studying undergrad.	Work. & Stud.	Other
Foreign Peers	0.079	-0.019	0.104*	-0.041	0.008
	(0.054)	(0.050)	(0.047)	(0.078)	(0.062)
Foreign Peers * $T_{.25} - T_{.5}$	-0.042	-0.007	-0.044^{+}	-0.032	0.044
	(0.028)	(0.021)	(0.023)	(0.037)	(0.062)
Foreign Peers * $T_{.5} - T_{.75}$	-0.093**	-0.035	-0.060***	-0.017	-0.007
	(0.030)	(0.030)	(0.016)	(0.036)	(0.029)
Foreign Peers * $T_{.75} - T_{1}$	-0.123**	-0.037	-0.085***	0.012	0.044
	(0.045)	(0.035)	(0.021)	(0.053)	(0.050)
			EU vs Non-EU		
EU Peers	0.006	-0.025	0.043	0.055	0.145
	(0.159)	(0.119)	(0.104)	(0.146)	(0.121)
EU Peers * $T_{.25} - T_{.5}$	-0.080	-0.069	-0.008	0.051	-0.090
	(0.085)	(0.066)	(0.067)	(0.074)	(0.078)
EU Peers * $T_{.5} - T_{.75}$	-0.224*	-0.141	-0.078	-0.091	-0.081
	(0.111)	(0.091)	(0.075)	(0.065)	(0.086)
EU Peers * $T_{.75} - T_{1}$	-0.120	-0.020	-0.109	-0.090	-0.053
	(0.105)	(0.095)	(0.081)	(0.099)	(0.093)
non-EU Peers	0.092	-0.027	0.124*	-0.063	-0.042
	(0.074)	(0.062)	(0.053)	(0.085)	(0.070)
non-EU Peers * $T_{.25} - T_{.5}$	-0.025	0.021	-0.057	-0.063	0.100
	(0.051)	(0.040)	(0.036)	(0.055)	(0.085)
non-EU Peers * $T_{.5} - T_{.75}$	-0.044	0.005	-0.054^{+}	0.009	0.027
	(0.050)	(0.042)	(0.029)	(0.054)	(0.045)
non-EU Peers * $T_{.75} - T_{1}$	-0.115^{+}	-0.034	-0.078*	0.045	0.085
	(0.065)	(0.055)	(0.038)	(0.078)	(0.063)
Observations	146,460	146,460	146,460	146,460	146,460

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.19: Labour Market Participation
Controlling for Grades

	Working	Working (& work.+stud.)	Working vs. Unemployed	Unemployed
Foreign Peers	-0.051	-0.092	0.027	-0.014
	(0.119)	(0.110)	(0.083)	(0.065)
Foreign Peers * $T_{.25} - T_{.5}$	0.047	0.016	0.020	-0.016
	(0.036)	(0.059)	(0.028)	(0.023)
Foreign Peers * $T_{.5} - T_{.75}$	0.107*	0.091	0.005	0.009
	(0.045)	(0.063)	(0.027)	(0.021)
Foreign Peers * $T_{.75} - T_{1}$	0.084	0.098	0.041	-0.026
	(0.063)	(0.101)	(0.038)	(0.024)
	EU vs Non-EU			
EU Peers	-0.335	-0.284	-0.082	0.077
	(0.218)	(0.237)	(0.189)	(0.145)
EU Peers * $T_{.25} - T_{.5}$	0.253*	0.301*	0.160	-0.100
	(0.126)	(0.126)	(0.105)	(0.084)
EU Peers * $T_{.5} - T_{.75}$	0.465**	0.370*	0.095	-0.040
	(0.157)	(0.167)	(0.134)	(0.111)
EU Peers * $T_{.75} - T_{1}$	0.348^{+}	0.259	0.119	-0.076
	(0.182)	(0.161)	(0.145)	(0.119)
non-EU Peers	0.055	-0.007	0.071	-0.048
	(0.141)	(0.145)	(0.103)	(0.080)
non-EU Peers * $T_{.25} - T_{.5}$	-0.043	-0.105	-0.038	0.019
	(0.059)	(0.083)	(0.035)	(0.030)
non-EU Peers * $T_{.5} - T_{.75}$	-0.035	-0.023	-0.034	0.031
	(0.061)	(0.086)	(0.048)	(0.037)
non-EU Peers * $T_{.75} - T_{1}$	-0.026	0.021	0.006	-0.004
	(0.087)	(0.129)	(0.062)	(0.044)
Observations	145,450	145,450	110,460	145,450

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.20: Studying After Graduation
Controlling for Grades

	Studying	Studying postgrad.	Studying undergrad.	Work. & Stud.	Other
Foreign Peers	0.089	-0.009	0.102*	-0.041	0.017
	(0.059)	(0.055)	(0.047)	(0.078)	(0.062)
Foreign Peers * $T_{.25} - T_{.5}$	-0.041	-0.007	-0.042	-0.032	0.041
	(0.027)	(0.024)	(0.026)	(0.035)	(0.062)
Foreign Peers * $T_{.5} - T_{.75}$	-0.091*	-0.036	-0.057***	-0.016	-0.009
	(0.035)	(0.036)	(0.017)	(0.033)	(0.028)
Foreign Peers * $T_{.75} - T_{1}$	-0.116*	-0.034	-0.081***	0.014	0.043
	(0.049)	(0.040)	(0.020)	(0.050)	(0.050)
			EU vs Non-EU		
EU Peers	0.050	0.007	0.053	0.051	0.157
	(0.154)	(0.116)	(0.105)	(0.144)	(0.121)
EU Peers * $T_{.25} - T_{.5}$	-0.106	-0.087	-0.017	0.048	-0.095
	(0.083)	(0.063)	(0.068)	(0.073)	(0.076)
EU Peers * $T_{.5} - T_{.75}$	-0.251*	-0.163^{+}	-0.083	-0.095	-0.079
	(0.114)	(0.094)	(0.075)	(0.060)	(0.085)
EU Peers * $T_{.75} - T_{1}$	-0.125	-0.025	-0.109	-0.089	-0.058
	(0.100)	(0.093)	(0.082)	(0.095)	(0.093)
non-EU Peers	0.089	-0.025	0.118*	-0.062	-0.034
	(0.077)	(0.066)	(0.052)	(0.086)	(0.070)
non-EU Peers * $T_{.25} - T_{.5}$	-0.012	0.028	-0.052	-0.062	0.098
	(0.048)	(0.043)	(0.038)	(0.052)	(0.085)
non-EU Peers * $T_{.5} - T_{.75}$	-0.031	0.013	-0.048	0.012	0.023
	(0.051)	(0.048)	(0.029)	(0.050)	(0.045)
non-EU Peers * $T_{.75} - T_{1}$	-0.102	-0.027	-0.072*	0.047	0.085
	(0.066)	(0.059)	(0.036)	(0.075)	(0.063)
Observations	145,450	145,450	145,450	145,450	145,450

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.21: Job Attributes
Controlling for Grades

	Professional occ.	Graduate job	(log)Salary	(log)Salary imp.
Foreign Peers	0.278	0.168	-0.178	-0.021
	(0.179)	(0.145)	(0.123)	(0.146)
Foreign Peers * $T_{.25} - T_{.5}$	0.049	-0.048	-0.014	0.039
	(0.080)	(0.059)	(0.046)	(0.039)
Foreign Peers * $T_{.5} - T_{.75}$	0.122^{+}	-0.043	0.064	0.095
	(0.073)	(0.081)	(0.074)	(0.077)
Foreign Peers * $T_{.75} - T_{1}$	0.122^{+}	0.066	0.111	0.139^{+}
	(0.070)	(0.092)	(0.071)	(0.082)
	EU vs Non-EU			
EU Peers	-0.222	-0.202	-0.300	-0.066
	(0.275)	(0.287)	(0.261)	(0.239)
EU Peers * $T_{.25} - T_{.5}$	0.053	0.061	-0.145	0.122
	(0.161)	(0.170)	(0.119)	(0.141)
EU Peers * $T_{.5} - T_{.75}$	0.600**	0.201	0.181	0.194
	(0.217)	(0.214)	(0.145)	(0.156)
EU Peers * $T_{.75} - T_{1}$	0.459*	0.261	0.339^{+}	0.347^{+}
	(0.220)	(0.265)	(0.178)	(0.184)
non-EU Peers	0.422*	0.277^{+}	-0.156	-0.001
	(0.189)	(0.150)	(0.137)	(0.162)
non-EU Peers * $T_{.25} - T_{.5}$	0.040	-0.094	0.035	0.001
	(0.109)	(0.093)	(0.068)	(0.071)
non-EU Peers * $T_{.5} - T_{.75}$	-0.058	-0.134	0.021	0.052
	(0.103)	(0.110)	(0.097)	(0.121)
non-EU Peers * $T_{.75} - T_{1}$	-0.010	-0.009	0.036	0.063
	(0.107)	(0.145)	(0.101)	(0.123)
Observations	86,230	76,870	40,500	51,300

Note: Sample of undergraduate students enrolled in 2007/8 in all English higher eduction institutions (HEI). Controls: tariff score and size of course. Fixed effects: HEI, year (of enrolment), major, HEI-year, HEI-major, major-year. Observations are weighted with analytic weights to account for the different contribution of course size in our estimates. Standard errors are clustered at HEI level and reported in parenthesis. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.22: Higher Education Outcomes
STEM transitions

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	0.015	0.031	-0.064	0.065
	(0.031)	(0.045)	(0.070)	(0.108)
Foreign Peers * $T_{.25} - T_{.5}$	-0.014*	-0.045^{+}	0.089	-0.074
	(0.007)	(0.024)	(0.060)	(0.064)
Foreign Peers * $T_{.5} - T_{.75}$	-0.005	-0.056*	0.098	-0.108
	(0.008)	(0.023)	(0.088)	(0.099)
Foreign Peers * $T_{.75} - T_{1}$	-0.019	-0.060*	0.069	-0.028
	(0.012)	(0.027)	(0.092)	(0.107)
	EU vs Non-EU			
EU Peers	-0.016	0.120	0.142	-0.151
	(0.059)	(0.086)	(0.208)	(0.274)
EU Peers * $T_{.25} - T_{.5}$	-0.036	-0.171*	-0.037	-0.041
	(0.027)	(0.075)	(0.168)	(0.220)
EU Peers * $T_{.5} - T_{.75}$	-0.015	0.006	0.083	-0.183
	(0.024)	(0.080)	(0.241)	(0.292)
EU Peers * $T_{.75} - T_{1}$	0.005	-0.178^{+}	0.145	-0.108
	(0.034)	(0.093)	(0.248)	(0.299)
non-EU Peers	0.025	-0.003	-0.116^{+}	0.119
	(0.035)	(0.054)	(0.063)	(0.095)
non-EU Peers * $T_{.25} - T_{.5}$	-0.005	0.010	0.120*	-0.080
	(0.011)	(0.035)	(0.050)	(0.064)
non-EU Peers * $T_{.5} - T_{.75}$	-0.001	-0.074*	0.095	-0.082
	(0.014)	(0.034)	(0.093)	(0.081)
non-EU Peers * $T_{.75} - T_{1}$	-0.025	-0.019	0.042	-0.000
	(0.019)	(0.043)	(0.087)	(0.093)
Observations	326,990	326,990	131,170	131,170

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimates by University Group

Table 4.23: Graduation Estimates
Russell Group

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	0.034	-0.040	0.073	-0.031^{+}	-0.003
	(0.131)	(0.091)	(0.064)	(0.015)	(0.130)
Foreign Peers * $T_{.25} - T_{.5}$	-0.013	0.076	-0.088	0.060*	-0.047
	(0.125)	(0.079)	(0.066)	(0.023)	(0.129)
Foreign Peers * $T_{.5} - T_{.75}$	-0.106	-0.038	-0.067	0.064**	0.042
	(0.126)	(0.080)	(0.060)	(0.018)	(0.129)
Foreign Peers * $T_{.75} - T_{1}$	-0.080	-0.014	-0.065	0.061**	0.020
	(0.116)	(0.073)	(0.066)	(0.019)	(0.121)
			EU vs Non-EU		
EU Peers	0.667^{+}	0.768*	-0.120	-0.288^{+}	-0.379
	(0.343)	(0.328)	(0.210)	(0.158)	(0.314)
EU Peers * $T_{.25} - T_{.5}$	-0.659**	-0.665*	0.023	0.245	0.415^{+}
	(0.198)	(0.240)	(0.163)	(0.185)	(0.213)
EU Peers * $T_{.5} - T_{.75}$	-0.690^{+}	-0.782*	0.111	0.322^{+}	0.368
	(0.342)	(0.328)	(0.173)	(0.154)	(0.328)
EU Peers * $T_{.75} - T_{1}$	-0.718*	-0.817*	0.116	0.302^{+}	0.416
	(0.343)	(0.321)	(0.181)	(0.154)	(0.322)
non-EU Peers	-0.111	-0.229	0.120*	0.030	0.081
	(0.138)	(0.149)	(0.044)	(0.042)	(0.130)
non-EU Peers * $T_{.25} - T_{.5}$	0.131	0.241^{+}	-0.113*	0.020	-0.150
	(0.145)	(0.119)	(0.053)	(0.042)	(0.147)
non-EU Peers * $T_{.5} - T_{.75}$	0.026	0.131	-0.107^{+}	0.005	-0.031
	(0.126)	(0.127)	(0.052)	(0.029)	(0.130)
non-EU Peers * $T_{.75} - T_{1}$	0.065	0.169	-0.107^{+}	0.006	-0.070
	(0.114)	(0.120)	(0.055)	(0.031)	(0.118)
Observations	133,650	133,650	133,650	133,650	133,650

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.24: Graduation Estimates
1994 Group

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	0.045	0.070	-0.026	0.004	-0.048
	(0.070)	(0.081)	(0.038)	(0.034)	(0.052)
Foreign Peers * $T_{.25} - T_{.5}$	-0.010	-0.023	0.014	-0.035	0.045
	(0.051)	(0.048)	(0.030)	(0.034)	(0.027)
Foreign Peers * $T_{.5} - T_{.75}$	-0.020	-0.024	0.006	-0.029	0.049
	(0.053)	(0.042)	(0.026)	(0.033)	(0.034)
Foreign Peers * $T_{.75} - T_1$	-0.006	-0.018	0.012	-0.033	0.039
	(0.064)	(0.060)	(0.032)	(0.032)	(0.035)
	EU vs Non-EU				
EU Peers	0.151	0.199	-0.053	0.003	-0.153
	(0.188)	(0.306)	(0.140)	(0.109)	(0.142)
EU Peers * $T_{.25} - T_{.5}$	-0.215	-0.210	-0.001	0.012	0.203
	(0.186)	(0.327)	(0.145)	(0.110)	(0.141)
EU Peers * $T_{.5} - T_{.75}$	-0.153	-0.168	0.019	-0.018	0.171
	(0.135)	(0.246)	(0.118)	(0.106)	(0.101)
EU Peers * $T_{.75} - T_1$	-0.296^{+}	-0.326	0.031	0.050	0.247*
	(0.142)	(0.270)	(0.147)	(0.092)	(0.089)
non-EU Peers	0.016	0.025	-0.007	-0.004	-0.011
	(0.086)	(0.083)	(0.036)	(0.062)	(0.033)
non-EU Peers * $T_{.25} - T_{.5}$	0.081	0.061	0.021	-0.056	-0.025
	(0.140)	(0.172)	(0.040)	(0.071)	(0.087)
non-EU Peers * $T_{.5} - T_{.75}$	0.043	0.044	-0.000	-0.036	-0.008
	(0.122)	(0.143)	(0.025)	(0.062)	(0.085)
non-EU Peers * $T_{.75} - T_1$	0.127	0.123	0.004	-0.072	-0.055
	(0.113)	(0.137)	(0.029)	(0.063)	(0.067)
Observations	36,720	36,720	36,720	36,720	36,720

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.25: Graduation Estimates
Other Universities

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	-0.021	0.041	-0.063	0.053*	-0.033
	(0.049)	(0.068)	(0.044)	(0.025)	(0.048)
Foreign Peers * $T_{.25} - T_{.5}$	-0.071*	-0.095*	0.025	-0.032*	0.103*
	(0.033)	(0.044)	(0.021)	(0.014)	(0.040)
Foreign Peers * $T_{.5} - T_{.75}$	-0.079*	-0.140***	0.061**	-0.038	0.117**
	(0.031)	(0.037)	(0.020)	(0.023)	(0.035)
Foreign Peers * $T_{.75} - T_{1}$	-0.095*	-0.177***	0.082***	-0.028^{+}	0.123**
	(0.035)	(0.048)	(0.023)	(0.014)	(0.040)
	EU vs Non-EU				
EU Peers	-0.098	0.077	-0.181	0.025	0.073
	(0.191)	(0.236)	(0.129)	(0.075)	(0.195)
EU Peers * $T_{.25} - T_{.5}$	-0.127	-0.150	0.023	-0.008	0.135
	(0.181)	(0.224)	(0.126)	(0.047)	(0.183)
EU Peers * $T_{.5} - T_{.75}$	0.077	0.068	0.024	0.004	-0.081
	(0.232)	(0.250)	(0.110)	(0.070)	(0.226)
EU Peers * $T_{.75} - T_{1}$	0.052	0.039	0.014	-0.018	-0.034
	(0.249)	(0.268)	(0.123)	(0.099)	(0.218)
non-EU Peers	-0.005	0.042	-0.047	0.062*	-0.057
	(0.044)	(0.074)	(0.043)	(0.026)	(0.051)
non-EU Peers * $T_{.25} - T_{.5}$	-0.049	-0.074	0.025	-0.041	0.090
	(0.058)	(0.086)	(0.041)	(0.025)	(0.072)
non-EU Peers * $T_{.5} - T_{.75}$	-0.134^{+}	-0.212**	0.073*	-0.053	0.186*
	(0.068)	(0.076)	(0.035)	(0.040)	(0.075)
non-EU Peers * $T_{.75} - T_{1}$	-0.144	-0.249*	0.105*	-0.031	0.175^{+}
	(0.087)	(0.105)	(0.043)	(0.036)	(0.093)
Observations	123,220	123,220	123,220	123,220	123,220

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.26: Grade Effects
Russell Group

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	0.018	-0.144	-0.061	0.141^{+}	0.085	-0.039*
	(0.107)	(0.171)	(0.083)	(0.078)	(0.066)	(0.016)
Foreign Peers * $T_{.25} - T_{.5}$	0.066	0.313**	-0.226*	-0.122	-0.102	0.071*
	(0.098)	(0.098)	(0.089)	(0.073)	(0.070)	(0.025)
Foreign Peers * $T_{.5} - T_{.75}$	0.026	0.136	-0.020	-0.140^{+}	-0.078	0.075**
	(0.076)	(0.142)	(0.084)	(0.073)	(0.062)	(0.020)
Foreign Peers * $T_{.75} - T_{1}$	0.018	0.087	0.035	-0.133	-0.077	0.071**
	(0.074)	(0.169)	(0.083)	(0.079)	(0.069)	(0.021)
			EU vs Non-EU			
EU Peers	0.749	0.730	-0.577	-0.406*	-0.145	-0.350^{+}
	(0.499)	(0.585)	(0.626)	(0.178)	(0.228)	(0.180)
EU Peers * $T_{.25} - T_{.5}$	-0.761	-0.503	0.623	0.304	0.040	0.296
	(0.520)	(0.573)	(0.536)	(0.256)	(0.179)	(0.206)
EU Peers * $T_{.5} - T_{.75}$	-1.063^{+}	-0.429	0.573	0.405^{+}	0.134	0.381*
	(0.552)	(0.561)	(0.649)	(0.203)	(0.190)	(0.176)
EU Peers * $T_{.75} - T_{1}$	-1.055^{+}	-0.425	0.502	0.479*	0.141	0.358^{+}
	(0.509)	(0.577)	(0.657)	(0.181)	(0.198)	(0.178)
non-EU Peers	-0.098	-0.396	0.053	0.263**	0.142**	0.036
	(0.164)	(0.242)	(0.129)	(0.084)	(0.049)	(0.047)
non-EU Peers * $T_{.25} - T_{.5}$	0.254^{+}	0.496**	-0.417***	-0.219*	-0.134*	0.020
	(0.138)	(0.160)	(0.102)	(0.082)	(0.056)	(0.046)
non-EU Peers * $T_{.5} - T_{.75}$	0.279^{+}	0.262	-0.153	-0.265**	-0.127*	0.005
	(0.136)	(0.210)	(0.131)	(0.087)	(0.057)	(0.033)
non-EU Peers * $T_{.75} - T_{1}$	0.268^{+}	0.199	-0.066	-0.277**	-0.128*	0.004
	(0.135)	(0.233)	(0.139)	(0.093)	(0.059)	(0.035)
Observations	127,050	127,050	127,050	127,050	127,050	127,050

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.27: Grade Effects
1994 Group

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	0.195$^+$	-0.405*	0.086	0.147	-0.029	0.006
	(0.089)	(0.150)	(0.099)	(0.080)	(0.043)	(0.039)
Foreign Peers * $T_{.25} - T_{.5}$	-0.022	0.359$^+$	-0.206**	-0.108	0.017	-0.039
	(0.064)	(0.174)	(0.060)	(0.093)	(0.034)	(0.038)
Foreign Peers * $T_{.5} - T_{.75}$	0.006	0.239$^+$	-0.094$^+$	-0.125	0.007	-0.033
	(0.043)	(0.123)	(0.043)	(0.084)	(0.030)	(0.037)
Foreign Peers * $T_{.75} - T_{1}$	-0.025	0.246	-0.081*	-0.117	0.015	-0.038
	(0.082)	(0.145)	(0.035)	(0.085)	(0.036)	(0.037)
			EU vs Non-EU			
EU Peers	0.297	-1.090*	0.815*	0.044	-0.060	-0.007
	(0.192)	(0.368)	(0.267)	(0.247)	(0.164)	(0.124)
EU Peers * $T_{.25} - T_{.5}$	-0.044	1.256***	-1.121***	-0.118	0.003	0.024
	(0.218)	(0.236)	(0.226)	(0.245)	(0.167)	(0.125)
EU Peers * $T_{.5} - T_{.75}$	-0.077	1.317**	-1.154**	-0.103	0.024	-0.008
	(0.156)	(0.285)	(0.246)	(0.250)	(0.139)	(0.121)
EU Peers * $T_{.75} - T_{1}$	-0.216	1.085***	-0.967**	-0.008	0.043	0.063
	(0.148)	(0.198)	(0.243)	(0.221)	(0.170)	(0.107)
non-EU Peers	0.141$^+$	-0.147	-0.199	0.211*	-0.009	0.002
	(0.074)	(0.118)	(0.127)	(0.085)	(0.040)	(0.071)
non-EU Peers * $T_{.25} - T_{.5}$	-0.010	-0.044	0.203	-0.105$^+$	0.023	-0.066
	(0.103)	(0.140)	(0.125)	(0.054)	(0.048)	(0.082)
non-EU Peers * $T_{.5} - T_{.75}$	0.046	-0.249*	0.385**	-0.136$^+$	-0.000	-0.046
	(0.084)	(0.098)	(0.104)	(0.072)	(0.033)	(0.072)
non-EU Peers * $T_{.75} - T_{1}$	0.062	-0.143	0.327*	-0.165$^+$	0.003	-0.085
	(0.129)	(0.159)	(0.118)	(0.088)	(0.036)	(0.072)
Observations	34,680	34,680	34,680	34,680	34,680	34,680

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^+$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.28: Grade Effects
Other Universities

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.038	-0.049	0.008	0.097*	-0.078	0.059*
	(0.101)	(0.099)	(0.093)	(0.040)	(0.052)	(0.029)
Foreign Peers * $T_{.25} - T_{.5}$	-0.035	-0.050	0.083^{+}	-0.009	0.044	-0.033*
	(0.050)	(0.046)	(0.047)	(0.018)	(0.026)	(0.014)
Foreign Peers * $T_{.5} - T_{.75}$	0.051	-0.263***	0.168*	-0.005	0.088**	-0.039
	(0.050)	(0.044)	(0.072)	(0.020)	(0.026)	(0.024)
Foreign Peers * $T_{.75} - T_{1}$	0.055	-0.391***	0.238***	0.013	0.112***	-0.028^{+}
	(0.074)	(0.059)	(0.039)	(0.017)	(0.028)	(0.014)
			EU vs Non-EU			
EU Peers	-0.059	0.107	-0.102	0.208**	-0.178	0.025
	(0.173)	(0.242)	(0.239)	(0.072)	(0.165)	(0.087)
EU Peers * $T_{.25} - T_{.5}$	-0.036	0.158	-0.071	-0.065	0.024	-0.009
	(0.144)	(0.191)	(0.210)	(0.083)	(0.158)	(0.054)
EU Peers * $T_{.5} - T_{.75}$	0.169	0.116	-0.259	-0.021	-0.005	0.001
	(0.212)	(0.248)	(0.253)	(0.070)	(0.145)	(0.080)
EU Peers * $T_{.75} - T_{1}$	0.371	-0.524	0.148	0.026	0.003	-0.025
	(0.330)	(0.383)	(0.240)	(0.082)	(0.157)	(0.110)
non-EU Peers	-0.027	-0.038	-0.008	0.075^{+}	-0.070	0.069*
	(0.110)	(0.096)	(0.101)	(0.041)	(0.057)	(0.027)
non-EU Peers * $T_{.25} - T_{.5}$	-0.034	-0.123	0.137*	0.011	0.051	-0.042
	(0.071)	(0.079)	(0.061)	(0.026)	(0.055)	(0.025)
non-EU Peers * $T_{.5} - T_{.75}$	0.011	-0.394***	0.316**	0.002	0.119*	-0.054
	(0.076)	(0.090)	(0.113)	(0.032)	(0.049)	(0.041)
non-EU Peers * $T_{.75} - T_{1}$	-0.051	-0.349**	0.271**	0.010	0.149*	-0.029
	(0.095)	(0.101)	(0.081)	(0.029)	(0.058)	(0.038)
Observations	109,360	109,360	109,360	109,360	109,360	109,360

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.29: STEM transitions
Russell Group

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	0.037	0.004	0.040	-0.041
	(0.034)	(0.127)	(0.121)	(0.210)
Foreign Peers * $T_{.25} - T_{.5}$	0.019	0.022	0.094	-0.163
	(0.029)	(0.155)	(0.058)	(0.180)
Foreign Peers * $T_{.5} - T_{.75}$	-0.018	-0.069	0.043	-0.124
	(0.022)	(0.132)	(0.072)	(0.185)
Foreign Peers * $T_{.75} - T_{1}$	-0.044^{+}	-0.032	-0.022	0.003
	(0.023)	(0.116)	(0.078)	(0.203)
	EU vs Non-EU			
EU Peers	0.012	1.099*	-0.311	0.361
	(0.104)	(0.427)	(0.232)	(0.677)
EU Peers * $T_{.25} - T_{.5}$	0.000	-1.042**	0.063	-0.088
	(0.107)	(0.337)	(0.164)	(0.578)
EU Peers * $T_{.5} - T_{.75}$	-0.053	-0.997*	0.463*	-0.603
	(0.101)	(0.401)	(0.219)	(0.639)
EU Peers * $T_{.75} - T_{1}$	-0.056	-1.044*	0.532**	-0.587
	(0.107)	(0.421)	(0.171)	(0.611)
non-EU Peers	0.049	-0.290	0.093	-0.105
	(0.039)	(0.185)	(0.141)	(0.197)
non-EU Peers * $T_{.25} - T_{.5}$	0.024	0.295	0.107	-0.185
	(0.034)	(0.210)	(0.065)	(0.212)
non-EU Peers * $T_{.5} - T_{.75}$	-0.009	0.171	-0.041	-0.026
	(0.023)	(0.166)	(0.085)	(0.155)
non-EU Peers * $T_{.75} - T_{1}$	-0.041	0.229	-0.140	0.130
	(0.027)	(0.148)	(0.097)	(0.182)
Observations	90,110	90,110	43,540	43,540

Note: Notes: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.30: STEM transitions
1994 Group

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	0.047	-0.030	0.253^{+}	0.168
	(0.034)	(0.069)	(0.111)	(0.249)
Foreign Peers * $T_{.25} - T_{.5}$	-0.035^{+}	0.012	-0.299*	-0.008
	(0.019)	(0.051)	(0.129)	(0.278)
Foreign Peers * $T_{.5} - T_{.75}$	-0.026	-0.008	-0.435*	0.182
	(0.021)	(0.056)	(0.160)	(0.224)
Foreign Peers * $T_{.75} - T_{1}$	-0.027	-0.010	-0.376^{+}	0.351
	(0.018)	(0.065)	(0.189)	(0.223)
	EU vs Non-EU			
EU Peers	-0.059	0.088	0.706	-0.112
	(0.108)	(0.245)	(0.462)	(0.581)
EU Peers * $T_{.25} - T_{.5}$	0.073	-0.205	-0.255	-0.278
	(0.101)	(0.297)	(0.338)	(0.379)
EU Peers * $T_{.5} - T_{.75}$	0.106	-0.133	-0.403	-0.014
	(0.099)	(0.233)	(0.478)	(0.307)
EU Peers * $T_{.75} - T_{1}$	0.091	-0.320	-0.276	0.018
	(0.108)	(0.225)	(0.517)	(0.374)
non-EU Peers	0.091	-0.060	-0.035	0.350
	(0.059)	(0.136)	(0.182)	(0.322)
non-EU Peers * $T_{.25} - T_{.5}$	-0.082^{+}	0.106	-0.316*	0.140
	(0.045)	(0.186)	(0.114)	(0.332)
non-EU Peers * $T_{.5} - T_{.75}$	-0.084	0.050	-0.452*	0.301
	(0.047)	(0.171)	(0.160)	(0.254)
non-EU Peers * $T_{.75} - T_{1}$	-0.078	0.128	-0.422*	0.529^{+}
	(0.046)	(0.157)	(0.157)	(0.275)
Observations	25,290	25,290	11,430	11,430

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.31: STEM transitions
Other Universities

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	-0.062	0.062	-0.178	0.125
	(0.086)	(0.112)	(0.168)	(0.241)
Foreign Peers * $T_{.25} - T_{.5}$	-0.018	-0.052^{+}	0.086	-0.045
	(0.012)	(0.030)	(0.127)	(0.143)
Foreign Peers * $T_{.5} - T_{.75}$	-0.007	-0.069*	0.125	-0.149
	(0.010)	(0.030)	(0.198)	(0.258)
Foreign Peers * $T_{.75} - T_{1}$	0.008	-0.091*	0.171	-0.182
	(0.015)	(0.041)	(0.253)	(0.278)
	EU vs Non-EU			
EU Peers	-0.192	0.171	-0.116	-0.253
	(0.246)	(0.305)	(0.491)	(0.611)
EU Peers * $T_{.25} - T_{.5}$	-0.088	-0.081	0.280	-0.118
	(0.070)	(0.189)	(0.384)	(0.471)
EU Peers * $T_{.5} - T_{.75}$	-0.115^{+}	0.208	0.612	-0.550
	(0.058)	(0.219)	(0.543)	(0.737)
EU Peers * $T_{.75} - T_{1}$	0.128	-0.041	0.492	-0.325
	(0.116)	(0.284)	(0.659)	(0.768)
non-EU Peers	-0.053	0.058	-0.164	0.170
	(0.069)	(0.088)	(0.147)	(0.190)
non-EU Peers * $T_{.25} - T_{.5}$	0.006	-0.041	0.015	-0.016
	(0.017)	(0.060)	(0.081)	(0.142)
non-EU Peers * $T_{.5} - T_{.75}$	0.029	-0.162*	-0.063	0.012
	(0.020)	(0.070)	(0.111)	(0.128)
non-EU Peers * $T_{.75} - T_{1}$	-0.031	-0.106	0.037	-0.111
	(0.034)	(0.099)	(0.135)	(0.174)
Observations	89,770	89,770	33,450	33,450

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.32: Grade Distribution
By University Group

	Distribution				P-value Null of Equality		
	UK	Foreign	EU	Non-EU	Foreign	EU	Non-EU
				Russell Group			
First	0.186	0.184	0.248	0.160	0.529	0.000	0.000
Upper Second	0.621	0.500	0.544	0.484	0.000	0.000	0.000
Lower Second	0.147	0.239	0.160	0.268	0.000	0.020	0.000
Third	0.020	0.054	0.026	0.065	0.000	0.002	0.000
Lower Qual	0.012	0.011	0.008	0.012	0.097	0.017	0.573
Fail	0.014	0.012	0.014	0.012	0.067	0.700	0.052
				1994 Group			
First	0.154	0.155	0.205	0.117	0.932	0.000	0.000
Upper Second	0.593	0.440	0.492	0.401	0.000	0.000	0.000
Lower Second	0.197	0.299	0.244	0.339	0.000	0.000	0.000
Third	0.023	0.074	0.036	0.102	0.000	0.000	0.000
Lower Qual	0.015	0.016	0.011	0.020	0.496	0.105	0.023
Fail	0.017	0.017	0.012	0.020	0.663	0.082	0.380
				Other Universities			
First	0.127	0.149	0.189	0.118	0.000	0.000	0.083
Upper Second	0.471	0.417	0.467	0.380	0.000	0.601	0.000
Lower Second	0.280	0.291	0.231	0.336	0.033	0.000	0.000
Third	0.044	0.078	0.053	0.097	0.000	0.013	0.000
Lower Qual.	0.055	0.046	0.042	0.048	0.000	0.001	0.037
Fail	0.022	0.019	0.018	0.021	0.117	0.091	0.520

Note: Observations from 2007 onwards only.

Higher Education Outcomes Estimates with Full Set of Controls

Table 4.33: Graduation Estimates
Controls for Individual Characteristics

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	0.041	0.055	-0.014	0.032*	-0.072*
	(0.034)	(0.038)	(0.022)	(0.015)	(0.032)
Foreign Peers * $T_{.25} - T_{.5}$	-0.051*	-0.067*	0.015	-0.023^{+}	0.074*
	(0.025)	(0.031)	(0.017)	(0.013)	(0.031)
Foreign Peers * $T_{.5} - T_{.75}$	-0.059*	-0.101***	0.042*	-0.026	0.086**
	(0.023)	(0.029)	(0.019)	(0.016)	(0.028)
Foreign Peers * $T_{.75} - T_{1}$	-0.064*	-0.115**	0.051*	-0.020	0.084**
	(0.025)	(0.037)	(0.022)	(0.012)	(0.030)
	EU vs Non-EU				
EU Peers	0.096	0.089	0.001	-0.011	-0.085
	(0.061)	(0.086)	(0.059)	(0.032)	(0.066)
EU Peers * $T_{.25} - T_{.5}$	-0.191**	-0.179*	-0.008	0.043^{+}	0.148*
	(0.066)	(0.082)	(0.056)	(0.025)	(0.070)
EU Peers * $T_{.5} - T_{.75}$	-0.037	-0.053	0.023	0.018	0.019
	(0.077)	(0.093)	(0.058)	(0.028)	(0.077)
EU Peers * $T_{.75} - T_{1}$	-0.160*	-0.166^{+}	0.010	0.042	0.119^{+}
	(0.078)	(0.091)	(0.057)	(0.035)	(0.068)
non-EU Peers	0.018	0.041	-0.020	0.049**	-0.067
	(0.040)	(0.053)	(0.027)	(0.018)	(0.041)
non-EU Peers * $T_{.25} - T_{.5}$	0.008	-0.020	0.025	-0.051**	0.043
	(0.035)	(0.049)	(0.022)	(0.019)	(0.045)
non-EU Peers * $T_{.5} - T_{.75}$	-0.063*	-0.115*	0.050^{+}	-0.045*	0.108**
	(0.031)	(0.046)	(0.026)	(0.022)	(0.040)
non-EU Peers * $T_{.75} - T_{1}$	-0.029	-0.097^{+}	0.066*	-0.044*	0.073^{+}
	(0.036)	(0.057)	(0.028)	(0.019)	(0.044)
Observations	458,160	458,160	458,160	458,160	458,160

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.34: Grade Effects
Controls for Individual Characteristics

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.089	0.114*	-0.051	0.020	-0.025	0.032^{+}
	(0.057)	(0.054)	(0.051)	(0.027)	(0.027)	(0.017)
Foreign Peers * $T_{.25} - T_{.5}$	0.066	-0.118*	0.056^{+}	-0.007	0.027	-0.024^{+}
	(0.067)	(0.055)	(0.030)	(0.012)	(0.022)	(0.014)
Foreign Peers * $T_{.5} - T_{.75}$	0.104*	-0.253***	0.134**	-0.013	0.056*	-0.029^{+}
	(0.041)	(0.041)	(0.050)	(0.012)	(0.026)	(0.016)
Foreign Peers * $T_{.75} - T_{1}$	0.084^{+}	-0.314***	0.188***	-0.003	0.066*	-0.022^{+}
	(0.047)	(0.054)	(0.044)	(0.015)	(0.029)	(0.012)
			EU vs Non-EU			
EU Peers	-0.140	0.087	-0.027	0.089*	0.012	-0.022
	(0.090)	(0.131)	(0.119)	(0.045)	(0.072)	(0.035)
EU Peers * $T_{.25} - T_{.5}$	0.029	-0.022	0.011	-0.057	-0.012	0.052^{+}
	(0.102)	(0.126)	(0.108)	(0.041)	(0.068)	(0.029)
EU Peers * $T_{.5} - T_{.75}$	0.151	-0.034	-0.100	-0.041	0.004	0.020
	(0.099)	(0.130)	(0.127)	(0.042)	(0.071)	(0.033)
EU Peers * $T_{.75} - T_{1}$	0.133	-0.208	0.003	0.027	-0.000	0.046
	(0.129)	(0.161)	(0.130)	(0.044)	(0.070)	(0.040)
non-EU Peers	-0.073	0.129	-0.063	-0.008	-0.040	0.054**
	(0.072)	(0.081)	(0.074)	(0.040)	(0.035)	(0.020)
non-EU Peers * $T_{.25} - T_{.5}$	0.081	-0.158^{+}	0.076	0.014	0.043	-0.056**
	(0.093)	(0.084)	(0.049)	(0.021)	(0.030)	(0.020)
non-EU Peers * $T_{.5} - T_{.75}$	0.087	-0.338***	0.225**	-0.001	0.077*	-0.050*
	(0.056)	(0.057)	(0.069)	(0.023)	(0.036)	(0.023)
non-EU Peers * $T_{.75} - T_{1}$	0.067	-0.361***	0.262***	-0.010	0.092*	-0.049*
	(0.063)	(0.074)	(0.061)	(0.028)	(0.038)	(0.020)
Observations	416,010	416,010	416,010	416,010	416,010	416,010

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.35: STEM Transitions and Ability

	To STEM	Graduates non-STEM	To non-STEM	Graduates STEM
Foreign Peers	0.015	0.031	-0.064	0.065
	(0.031)	(0.045)	(0.070)	(0.108)
Foreign Peers * $T_{.25} - T_{.5}$	-0.014*	-0.045^{+}	0.089	-0.074
	(0.007)	(0.024)	(0.060)	(0.064)
Foreign Peers * $T_{.5} - T_{.75}$	-0.005	-0.056*	0.098	-0.108
	(0.008)	(0.023)	(0.088)	(0.099)
Foreign Peers * $T_{.75} - T_{1}$	-0.019	-0.060*	0.069	-0.028
	(0.012)	(0.027)	(0.092)	(0.107)
	EU vs Non-EU			
EU Peers	-0.016	0.120	0.142	-0.151
	(0.059)	(0.086)	(0.208)	(0.274)
EU Peers * $T_{.25} - T_{.5}$	-0.036	-0.171*	-0.037	-0.041
	(0.027)	(0.075)	(0.168)	(0.220)
EU Peers * $T_{.5} - T_{.75}$	-0.015	0.006	0.083	-0.183
	(0.024)	(0.080)	(0.241)	(0.292)
EU Peers * $T_{.75} - T_{1}$	0.005	-0.178^{+}	0.145	-0.108
	(0.034)	(0.093)	(0.248)	(0.299)
non-EU Peers	0.025	-0.003	-0.116^{+}	0.119
	(0.035)	(0.054)	(0.063)	(0.095)
non-EU Peers * $T_{.25} - T_{.5}$	-0.005	0.010	0.120*	-0.080
	(0.011)	(0.035)	(0.050)	(0.064)
non-EU Peers * $T_{.5} - T_{.75}$	-0.001	-0.074*	0.095	-0.082
	(0.014)	(0.034)	(0.093)	(0.081)
non-EU Peers * $T_{.75} - T_{1}$	-0.025	-0.019	0.042	-0.000
	(0.019)	(0.043)	(0.087)	(0.093)
Observations	326,990	326,990	131,170	131,170

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Heterogeneity Across Ethnic Groups

Driven by evidence showing that different ethnic groups perform differently in the UK (Dustmann, Machin, et al. 2010), we explore whether there are heterogeneous foreign peer effects across ethnic groups. In Tables 4.36 to 4.39, we replicate the foreign peer effects on graduation and STEM transitions in Table 4.6 for white and non-white British. As the average effect estimates reported in the main paper, all estimates in Tables 4.36 to 4.39 have moderate magnitudes and are typically non statistically significant. The largest effect is that of EU peers on the probability of transitioning from STEM into non-STEM majors for white British students. Where a one percentage point increase in EU peer shares leads to a .21 percentage point increase in the probability of STEM to non-STEM transition. Although, it is not statistically significant. We only find statistically significant effects on the probability of failing for white British. For whom a one percentage point increase on the foreign peer share increases the probability of failing by .027 percentage points and the effect is driven by both EU and non-EU peers. For non-white British, we find, marginally, significant effects on the probability of graduating and dropping out. Particularly, a one percentage point increase on the foreign student share increases the probability of graduation by .096 percentage points and decreases the probability of dropping out by .086 percentage points.

Moreover, in tables 4.40 and 4.41, we report estimates of the effect of foreign peers on grades of white and non-white British students. As with the other outcomes we do not find much of an effect. If anything, increasing the share of foreign students increases the probability of failing to graduate for white British, as already reported in table 4.36, and increases the probability of graduating with a lower second for non-white British.

Therefore, across ethnic groups, foreign peer effects are similar to the average peer effects in section 4.6.1, both in terms of magnitudes and significance.

Table 4.36: Graduation
White British

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	-0.038	-0.053	0.015	0.027*	0.011
	(0.028)	(0.033)	(0.018)	(0.013)	(0.023)
			EU vs Non-EU		
EU Peers	-0.026	-0.023	-0.003	0.037^{+}	-0.011
	(0.053)	(0.051)	(0.041)	(0.019)	(0.051)
non-EU Peers	-0.042	-0.063	0.021	0.024^{+}	0.018
	(0.030)	(0.038)	(0.020)	(0.014)	(0.025)
Observations	364,950	364,950	364,950	364,950	364,950

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.37: Transitions
White British

	To STEM	To non-STEM
Foreign Peers	0.004	0.002
	(0.035)	(0.044)
	EU vs Non-EU	
EU Peers	-0.040	0.210
	(0.061)	(0.128)
non-EU Peers	0.019	-0.045
	(0.042)	(0.057)
Observations	259,980	104,970

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.38: Graduation
Non-White British

	Graduates	Success	Lower Qual.	Failure	Dropout
Foreign Peers	0.096^{+}	0.087	0.011	-0.010	-0.086^{+}
	(0.057)	(0.056)	(0.035)	(0.024)	(0.047)
	EU vs Non-EU				
EU Peers	0.082	0.060	0.013	-0.027	-0.055
	(0.075)	(0.101)	(0.061)	(0.045)	(0.082)
non-EU Peers	0.103	0.098	0.010	-0.002	-0.100^{+}
	(0.074)	(0.070)	(0.040)	(0.026)	(0.058)
Observations	93,140	93,140	93,140	93,140	93,140

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.39: Transitions
Non-White British

	To STEM	To non-STEM
Foreign Peers	0.018	-0.077
	(0.034)	(0.122)
	EU vs Non-EU	
EU Peers	-0.015	0.159
	(0.077)	(0.212)
non-EU Peers	0.032	-0.168
	(0.040)	(0.129)
Observations	66,980	26,160

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.40: Grades
White British

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	0.000	-0.038	-0.003	-0.004	0.016	0.029*
	(0.047)	(0.054)	(0.050)	(0.025)	(0.020)	(0.014)
	EU vs Non-EU					
EU Peers	-0.067	0.071	-0.090	0.059*	-0.007	0.034
	(0.059)	(0.099)	(0.076)	(0.029)	(0.043)	(0.022)
non-EU Peers	0.022	-0.074	0.026	-0.025	0.024	0.028^{+}
	(0.061)	(0.073)	(0.062)	(0.035)	(0.022)	(0.016)
Observations	332,730	332,730	332,730	332,732	332,730	332,730

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.41: Grades
Non-White British

	First	Upper Second	Lower Second	Third	Lower Qual.	Fail
Foreign Peers	-0.077	-0.115	0.156^{+}	0.049	0.006	-0.019
	(0.064)	(0.085)	(0.084)	(0.053)	(0.039)	(0.029)
			EU vs Non-EU			
EU Peers	0.004	-0.144	0.062	0.091	0.030	-0.043
	(0.094)	(0.174)	(0.180)	(0.071)	(0.073)	(0.050)
non-EU Peers	-0.112	-0.103	0.197*	0.030	-0.004	-0.009
	(0.078)	(0.107)	(0.099)	(0.068)	(0.043)	(0.032)
Observations	83,210	83,210	83,210	83,210	83,210	83,210

Note: Higher education institution clustered standard errors in parenthesis. All include tariff controls and fix effects. Observations from 2007 onwards only. Observations rounded to last unit. $^{+}$ $p < .1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.9.3 Plots outside main body

Figure 4.6: Sex, Age and School

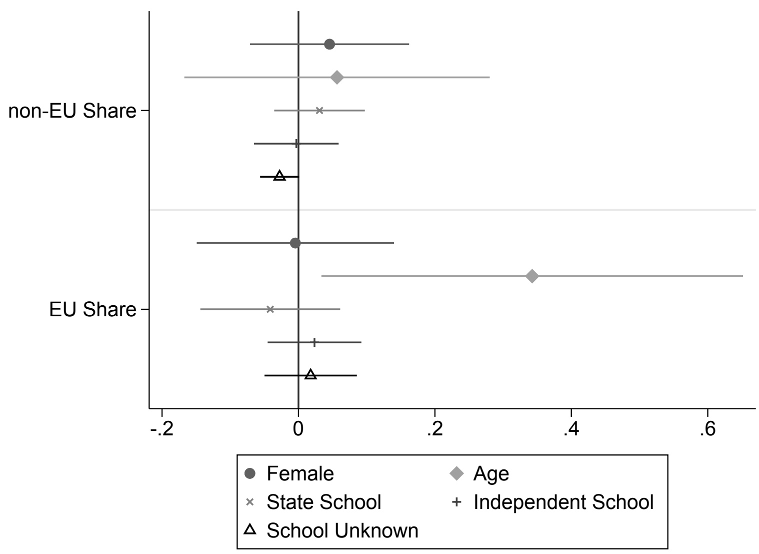


Figure 4.5: Simulated Ability Distribution

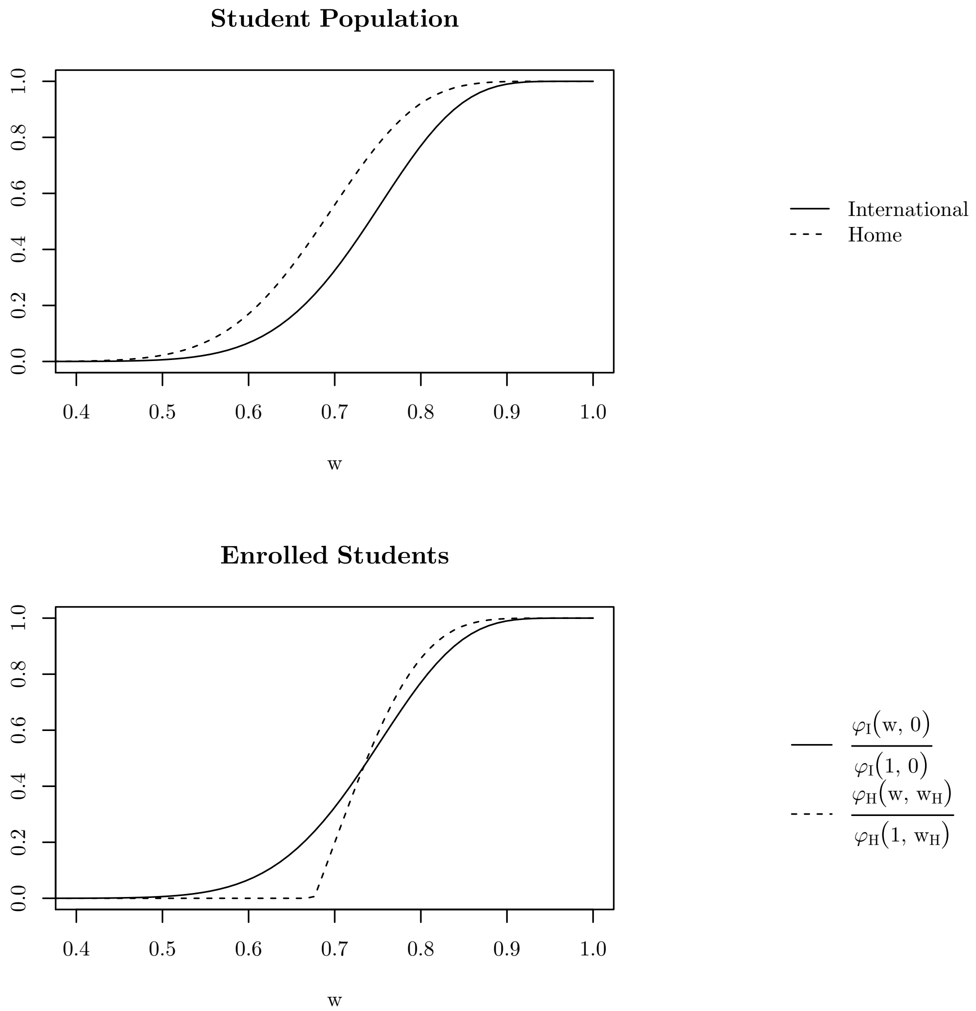


Figure 4.7: Ethnicity

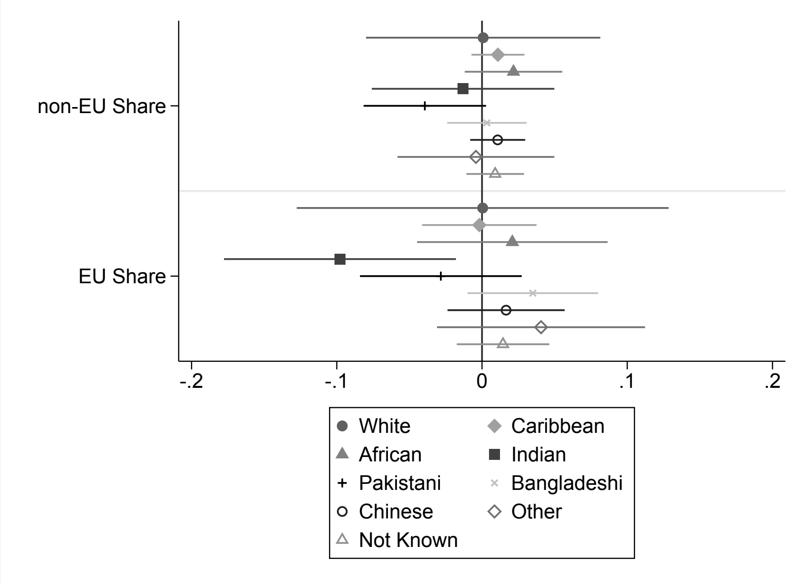


Figure 4.8: Disability

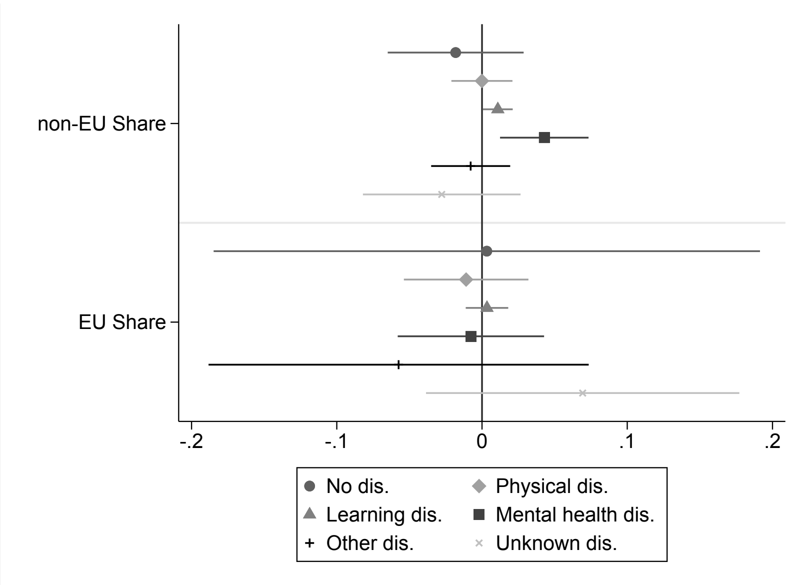


Figure 4.9: Share of international students by major, 2001/2010

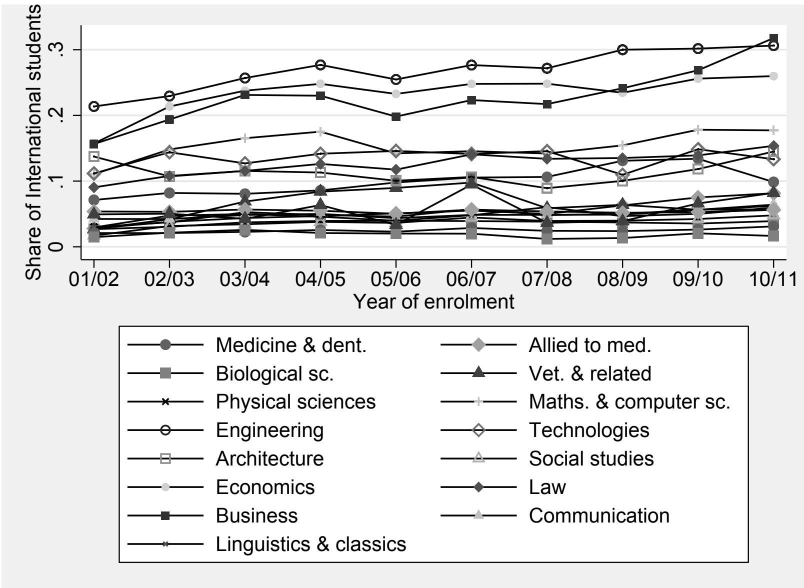


Figure 4.10: Share of EU students by major, 2001/2010

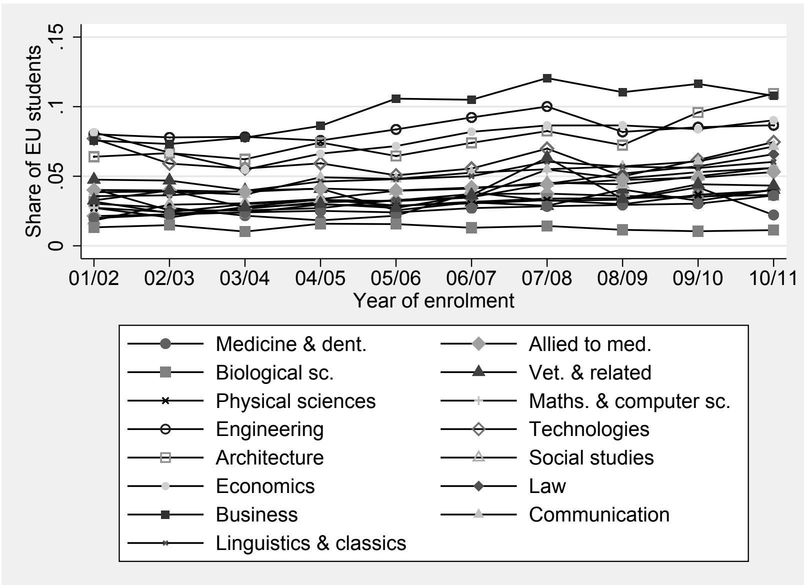


Figure 4.11: Share of Foreign students by type of HEI

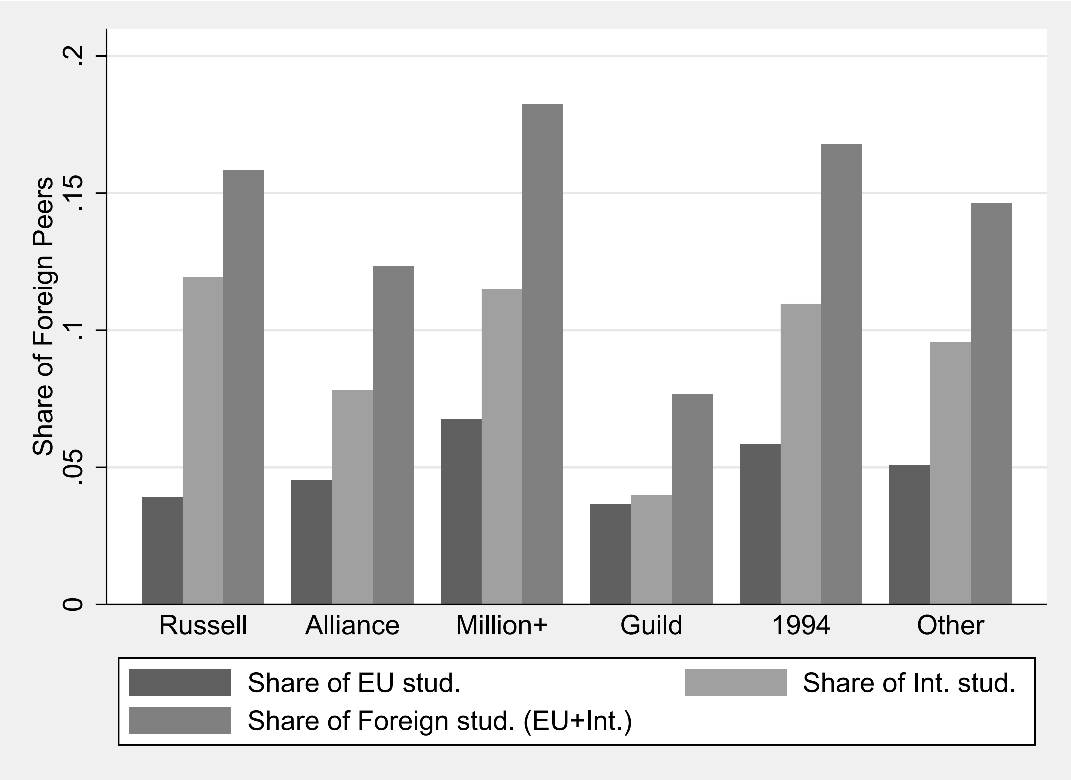


Figure 4.12: Ability Density
Enrolled Students

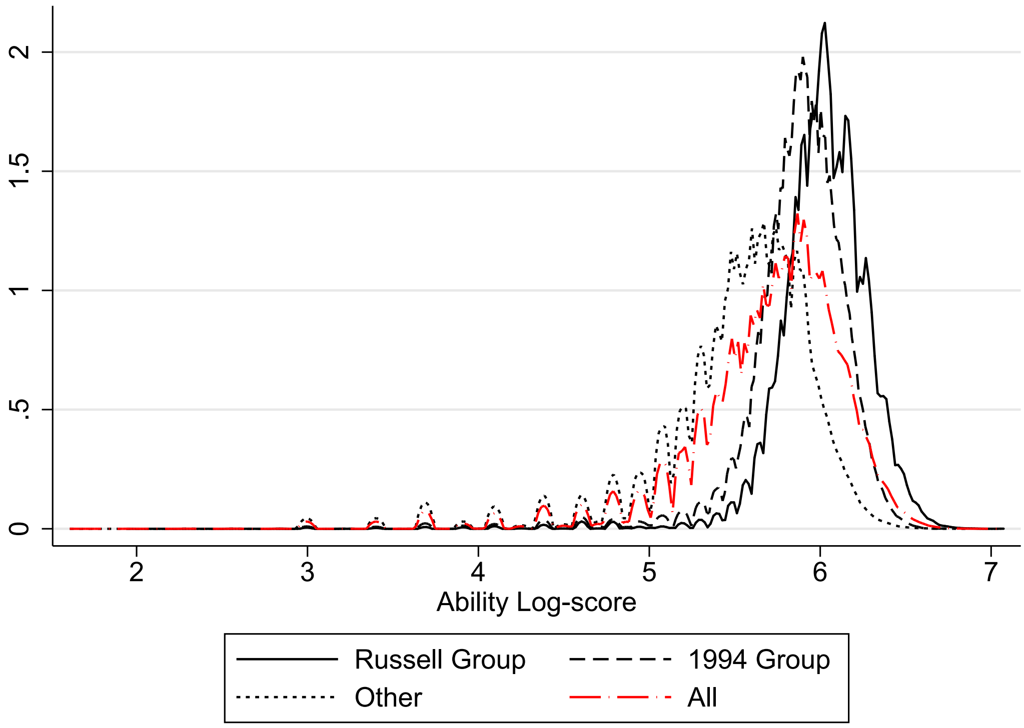
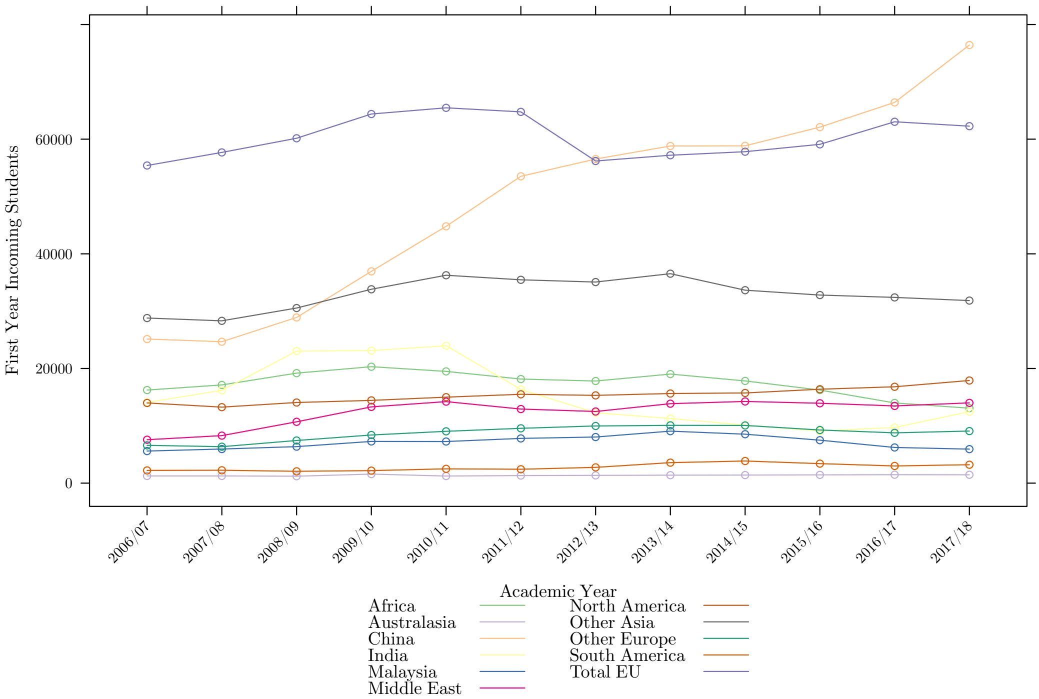


Figure 4.13: Evolution of Foreign Students Inflows



Bibliography

Abramitzky, R. & Boustan, L. (2017). Immigration in American economic history. *Journal of Economic Literature*, *55*(4), 1311–45.

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics*, *113*(4), 1055–1089.

Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies*, *69*(4), 781–809.

Acemoglu, D. & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier.

Aleksynska, M. & Tritah, A. (2009). Immigration and host countries income and productivity: A channel accounting approach. *Journal of Economic Literature*.

Altonji, J. G. & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market* (pp. 201–234). University of Chicago Press.

Ambrosini, J. W. & Peri, G. (2012). The determinants and the selection of Mexico–US migrants. *The World Economy*, *35*(2), 111–151.

Anelli, M. & Peri, G. (2017). The effects of high school peers' gender on college major, college performance and income. *The Economic Journal*, *129*(618), 553–602.

Anelli, M., Shih, K., & Williams, K. (2017). Foreign Peer Effects and STEM Major Choice. *IZA Discussion Paper*, (10743).

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, *30*, 98–108.

Angrist, J. D. & Lang, K. (2004). Does school integration generate peer effects? evidence from boston's metco program. *American Economic Review*, *94*(5), 1613–1634.

Angrist, J. D. & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*(2), 533–575.

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

Arnold, B. C., Beaver, R. J., Groeneveld, R. A., & Meeker, W. Q. (1993). The non-truncated marginal of a truncated bivariate normal distribution. *Psychometrika*, *58*(3), 471–488.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 47–57.

Autor, D., Katz, L. F., & Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, *90*(2), 300–323.

Aydemir, A. & Borjas, G. J. (2011). Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics*, *29*(1), 69–112.

Aydemir, A. & Skuterud, M. (2008). The immigrant wage differential within and across establishments. *ILR Review*, *61*(3), 334–352.

Baert, S., Cockx, B., & Verhaest, D. (2013). Overeducation at the start of the career: Stepping stone or trap? *Labour Economics*, *25*, 123–140.

Ballatore, R. M., Fort, M., & Ichino, A. (2018). Tower of babel in the classroom: Immigrants and natives in italian schools. *Journal of Labor Economics*, *36*(4), 885–921.

Bauer, T., Epstein, G. S., & Gang, I. N. (2007). The influence of stocks and flows on migrants' location choices. In *Aspects of worker well-being* (pp. 199–229). Emerald Group Publishing Limited.

BEIS. (2017). *Industrial Strategy: building a Britain fit for the future*. Department for Business Energy & Industrial Strategy.

Belfield, C., Britton, J., Buscha, F., Dearden, L., Dickson, M., van der Erve, L., ... Zhu, Y. (2018). *The relative labour market returns to different degrees*. Institute for Fiscal Studies. London, UK.

Bell, B., Fasani, F., & Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, *21*(3), 1278–1290.

Belot, M. V. & Hatton, T. J. (2012). Immigrant selection in the OECD*. *The Scandinavian Journal of Economics*, *114*(4), 1105–1128.

Betts, J. R. (1998). Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities? In D. S. Hamermesh & F. D. Bean (Eds.), *In Help or Hindrance? The Economic Implications of Immigration for African-Americans*. New York: Russell Sage Foundation.

Betts, J. R. & Fairlie, R. W. (2003). Does immigration induce "native flight" from public schools into private schools? *Journal of Public Economics*, *87*(5-6), 987–1012.

Bhattacharya, D., Kanaya, S., & Stevens, M. (2017). Are university admissions academically fair? *Review of Economics and Statistics*, *99*(3), 449–464.

Bishop, K. (2012). A dynamic model of location choice and hedonic valuation. *Unpublished, Washington University in St. Louis.*

Blundell, R., Crawford, C., & Jin, W. (2014). What can wages and employment tell us about the uk's productivity puzzle? *The Economic Journal*, *124*(576), 377–407.

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

Borjas, G. J. (1991). Immigration and self-selection. In J. M. Abowd & R. B. Freeman (Eds.), *Immigration, trade, and the labor market* (pp. 29–76). University of Chicago Press.

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Borjas, G. J., Freeman, R. B., & Katz, L. F. (1992). On the labor market effects of immigration and trade. In *Immigration and the workforce: Economic consequences for the united states and source countries* (pp. 213–244). University of Chicago Press.

Borjas, G. J., Bronars, S. G., & Trejo, S. J. (1992). Self-selection and internal migration in the united states. *Journal of Urban Economics*, *32*(2), 159–185.

Borjas, G. J., Kauppinen, I., & Poutvaara, P. (2018). Self-selection of emigrants: Theory and evidence on stochastic dominance in observable and unobservable characteristics. *The Economic Journal*, *129*(617), 143–171.

Borjas, G. J. & Tienda, M. (1987). The economic consequences of immigration. *Science*, *235*(4789), 645–651.

Boubtane, E., Dumont, J.-C., & Rault, C. (2016). Immigration and economic growth in the oecd countries 1986–2006. *Oxford Economic Papers*, *68*(2), 340–360.

Boucher, A. & Cerna, L. (2014). Current policy trends in skilled immigration policy. *International Migration*, *52*(3), 21–25.

Braakmann, N. & McDonald, S. (2018). Student exposure to socio-economic diversity and students' university outcomes–evidence from english administrative data. *Working Paper*.

Brunello, G. & Rocco, L. (2013). The effect of immigration on the school performance of natives: Cross country evidence using pisa test scores. *Economics of Education Review*, *32*, 234–246.

Butcher, K. F. & Card, D. (1991). Immigration and wages: Evidence from the 1980's. *The American Economic Review*, *81*(2), 292–296.

Cadena, B. C. & Kovak, B. K. (2016). Immigrants equilibrate local labor markets: Evidence from the great recession. *American Economic Journal: Applied Economics*, *8*(1), 257–90.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review*, *43*(2), 245–257.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Card, D. (2009). Immigration and inequality. *The American Economic Review*, *99*(2), 1–21. Retrieved from http://www.jstor.org/stable/25592368

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, *36*(S1), S13–S70.

Card, D. & Peri, G. (2016). Immigration economics by George J. Borjas: A review essay. *Journal of Economic Literature*, *54*(4), 1333–49.

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, *27*(3), 439–464.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, *108*(11), 3377–3415.

Carrell, S. E., Malmstrom, F. V., & West, J. E. (2008). Peer effects in academic cheating. *Journal of Human Resources*, *43*(1), 173–207.

Chevalier, A., Isphording, I. E., & Lisauskaite, E. (2019). *Peer diversity, college performance and educational choices*. Institute for the Study of Labor (IZA).

Chin, A., Daysal, N. M., & Imberman, S. A. (2013). Impact of bilingual education programs on limited english proficient students and their peers: Regression discontinuity evidence from texas. *Journal of Public Economics*, *107*, 63–78.

Chiquiar, D. & Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy*, *113*, 239–281.

Chiswick, B. R. (1978). The effect of americanization on the earnings of foreign-born men. *Journal of political Economy*, *86*(5), 897–921.

Clemens, M. A., Lewis, E. G., & Postel, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the Mexican bracero exclusion. *American Economic Review*, *108*(6), 1468–87.

Cools, A., Fernández, R., & Patacchini, E. (2019). *Girls, boys, and high achievers*. National Bureau of Economic Research.

Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, *107*(2), 425–56.

Dahl, G. B. (2002). Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica*, *70*(6), 2367–2420.

Dancygier, R. & Margalit, Y. (2019). The evolution of the immigration debate: Evidence from a new dataset of party positions over the last half-century. *Comparative Political Studies*, 0010414019858936.

Davidson, R. & Duclos, J.-Y. (2013). Testing for restricted stochastic dominance. *Econometric Reviews*, *32*(1), 84–125.

Dustmann, C., Fabbri, F., & Preston, I. P. (2005). The impact of immigration on the british labour market. *The Economic Journal*, *115*(507).

Dustmann, C., Frattini, T., & Preston, I. P. (2013). The effect of immigration along the distribution of wages. *The Review of Economic Studies*, *80*(1), 145–173.

Dustmann, C., Glitz, A., & Vogel, T. (2010). Employment, wages, and the economic cycle: Differences between immigrants and natives. *European Economic Review*, *54*(1), 1–17.

Dustmann, C. & Görlach, J.-S. (2015). Selective out-migration and the estimation of immigrants' earnings profiles. In *Handbook of the economics of international migration* (Vol. 1, pp. 489–533). Elsevier.

Dustmann, C., Machin, S., & Schönberg, U. (2010). Ethnicity and educational achievement in compulsory schooling. *The Economic Journal*, *120*(546), F272–F297.

Dustmann, C. & Preston, I. P. (2007). Racial and economic factors in attitudes to immigration. *The BE Journal of Economic Analysis & Policy*, *7*(1).

Dustmann, C. & Preston, I. P. (2012). Comment: Estimating the effect of immigration on wages. *Journal of the European Economic Association*, *10*(1), 216–223.

Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, *30*(4), 31–56.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, *132*(1), 435–483.

Eckstein, Z. & Weiss, Y. (2004). On the wage growth of immigrants: Israel, 1990–2000. *Journal of the European Economic Association*, *2*(4), 665–695.

Ehrenberg, R. G. (2004). Econometric studies of higher education. *Journal of Econometrics*, *121*(1-2), 19–37.

Elsner, B. & Isphording, I. E. (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics*, *35*(3), 787–828.

Faber, C. S. (2000). *Geographical mobility, population characteristics 1997-1998*. Current Population Reports.

Facchini, G., Frattini, T., & Signorotto, C. (2016). Mind what your voters read: Media exposure and international economic policy making.

Feliciano, C. (2005). Educational selectivity in us immigration: How do immigrants compare to those left behind? *Demography*, *42*(1), 131–152.

Feng, A. & Graetz, G. (2017). A question of degree: The effects of degree class on labor market outcomes. *Economics of Education Review*, *61*, 140–161.

Fernandez-Huertas Moraga, J. (2011). New evidence on emigrant selection. *The Review of Economics and Statistics*, *93*(1), 72–96.

Fernandez-Huertas Moraga, J. (2013). Understanding different migrant selection patterns in rural and urban Mexico. *Journal of Development Economics*, *103*, 182–201.

Finney, N. & Simpson, L. (2009). *Sleepwalking to segregation?: Challenging myths about race and migration*. Policy Press.

Flood, S., King, M., Ruggles, S., & Warren, J. R. (2015). *Integrated public use microdata series, current population survey: Version 4.0. [machine-readable database]. minneapolis: University of minnesota.*

French, M. T., Homer, J. F., Popovici, I., & Robins, P. K. (2015). What you do in high school matters: High school gpa, educational attainment, and labor market earnings as a young adult. *Eastern Economic Journal*, *41*(3), 370–386.

Gabriel, P. E. & Schmitz, S. (1995). Favorable self-selection and the internal migration of young white males in the united states. *Journal of Human Resources*, 460–471.

Gary, C. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies*, *47*(1), 225.

Geay, C., McNally, S., & Telhaj, S. (2013). Non-native speakers of english in the classroom: What are the effects on pupil performance? *The Economic Journal*, *123*(570), F281–F307.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2017). Bartik instruments: What, when, why and how. *Unpublished Working Paper*.

Goodwin, M. & Milazzo, C. (2017). Taking back control? Investigating the role of immigration in the 2016 vote for Brexit. *The British Journal of Politics and International Relations*, *19*(3), 450–464.

Gould, E. D., Lavy, V., & Daniele Paserman, M. (2009). Does immigration affect the long-term educational outcomes of natives? quasi-experimental evidence. *The Economic Journal*, *119*(540), 1243–1269.

Grogger, J. & Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, *95*(1), 42–57.

Grossman, J. B. (1982). The substitutability of natives and immigrants in production. *The Review of Economics and Statistics*, 596–603.

Ham, J. C., Li, X., & Reagan, P. B. (2011). Matching and semi-parametric iv estimation, a distance-based measure of migration, and the wages of young men. *Journal of Econometrics*, *161*(2), 208–227.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of applied econometrics*, *18*(5), 527–544.

Hatton, T. J. & Tani, M. (2005). Immigration and inter-regional mobility in the UK, 1982–2000. *The Economic Journal*, *115*(507), F342–F358.

Home Office. (2018). *The UK's future skills-based immigration system*. Home Office, UK Visas, and Immigration.

Hornbeck, R. & Naidu, S. (2014). When the levee breaks: black migration and economic development in the American South. *American Economic Review*, *104*(3), 963–90.

Hornung, E. (2014). Immigration and the diffusion of technology: The huguenot diaspora in prussia. *The American Economic Review*, *104*(1), 84–122.

Horowitz, J. L. (2011). Applied nonparametric instrumental variables estimation. *Econometrica*, *79*(2), 347–394.

Hunt, J. (1992). The impact of the 1962 repatriates from algeria on the french labor market. *ILR Review*, *45*(3), 556–572.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources*, *52*(4), 1060–1118.

Hunt, J. & Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, *2*(2), 31–56.

Ibarraran, P. & Lubotsky, D. (2007). Mexican immigration and self-selection: New evidence from the 2000 Mexican census. In G. J. Borjas (Ed.), *Mexican immigration to the united states* (pp. 159–192). University of Chicago Press.

Iranzo, S., Schivardi, F., & Tosetti, E. (2008). Skill dispersion and firm productivity: An analysis with employer-employee matched data. *Journal of Labor Economics*, *26*(2), 247–285.

Izquierdo, M., Lacuesta, A., & Vegas, R. (2009). Assimilation of immigrants in Spain: A longitudinal analysis. *Labour Economics*, *16*(6), 669–678.

Jaeger, D. A. (2007). Green cards and the location choices of immigrants in the united states, 1971–2000. *Immigration (Research in Labor Economics)*, *27*, 131–183.

Jaeger, D. A. & Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, 733–740.

Jaeger, D. A., Ruist, J., & Stuhler, J. (2018). *Shift-share instruments and the impact of immigration*. National Bureau of Economic Research.

Jasso, G., Massey, D. S., Rosenzweig, M. R., & Smith, J. P. (2004). Immigrant Health–Selectivity and Acculturation. In N. B. Anderson, R. A. Bulatao, & B. Cohen (Eds.), *Critical perspectives on racial and ethnic differences in health in late life* (pp. 213–244). National Academy Press.

Johnson, G. E. (1980). The labor market effects of immigration. *ILR Review*, *33*(3), 331–341.

Jones, E. B. & Jackson, J. D. (1990). College grades and labor market rewards. *The Journal of Human Resources*, *25*(2), 253.

Kaestner, R. & Malamud, O. (2014). Self-selection and international migration: New evidence from Mexico. *Review of Economics and Statistics*, *96*(1), 78–91.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, *17*(2), 303–316.

Kangasniemi, M., Mas, M., Robinson, C., & Serrano, L. (2012). The economic impact of migration: Productivity analysis for spain and the uk. *Journal of Productivity Analysis*, 1–11.

Kaplan, G. & Schulhofer-Wohl, S. (2012). Interstate migration has fallen less than you think: Consequences of hot deck imputation in the current population survey. *Demography*, *49*(3), 1061–1074.

Kaplan, G. & Schulhofer-Wohl, S. (2017). Understanding the long-run decline in interstate migration. *International Economic Review*, *58*(1), 57–94.

Kennan, J. & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, *79*(1), 211–251.

Kennedy, S., Kidd, M. P., McDonald, J. T., & Biddle, N. (2015). The healthy immigrant effect: Patterns and evidence from four countries. *Journal of International Migration and Integration*, *16*(2), 317–332.

Kleibergen, F. & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126.

Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal*, *113*(485), F34–F63.

LaLonde, R. J. & Topel, R. H. (1991). Immigrants in the american labor market: Quality, assimilation, and distributional effects. *The American Economic Review*, *81*(2), 297–302.

Lavy, V. & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*(2), 1–33.

Llull, J. (2017). Immigration, wages, and education: A labour market equilibrium structural model. *The Review of Economic Studies*, *85*(3), 1852–1896.

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *The Review of Economics and Statistics*, *89*(2), 289–299.

MAC. (2018). *Impact of international students in the UK*. Migration Advisory Committee.

Machin, S. & Murphy, R. (2017). Paying out and crowding out? the globalization of higher education. *Journal of Economic Geography*, *17*(5), 1075–1110.

Manacorda, M., Manning, A., & Wadsworth, J. (2012). The impact of immigration on the structure of wages: Theory and evidence from britain. *Journal of the European Economic Association*, *10*(1), 120–151.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542.

Mastrobuoni, G. & Pinotti, P. (2015). Legal status and the criminal activity of immigrants. *American Economic Journal: Applied Economics*, *7*(2), 175–206.

McKenzie, D., Stillman, S., & Gibson, J. (2010). How important is selection? experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association*, *8*(4), 913–945.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, *112*(4), 725–753.

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the united states. *The Journal of Economic Perspectives*, *25*(3), 173–196. Retrieved from http://0-www.jstor.org.serlib0.essex.ac.uk/stable/23049428

Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review*, *94*(3), 656–690.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, *118*(2), 549–599.

Murphy, R. & Weinhardt, F. (2018). *Top of the class: The importance of ordinal rank*. National Bureau of Economic Research.

Ohinata, A. & Van Ours, J. C. (2013). How immigrant children affect the academic achievement of native dutch children. *The Economic Journal*, *123*(570), F308–F331.

Oreopoulos, P., Von Wachter, T., & Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, *4*(1), 1–29.

Orrenius, P. M. & Zavodny, M. (2005). Self-selection among undocumented immigrants from Mexico. *Journal of Development Economics*, *78*(1), 215–240.

Orrenius, P. M. & Zavodny, M. (2015). Does immigration affect whether us natives major in science and engineering? *Journal of Labor Economics*, *33*(S1), S79–S108.

Ortega, F. & Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of International Economics*, *92*(2), 231–251.

Ottaviano, G. I. & Peri, G. (2006). The economic value of cultural diversity: Evidence from US cities. *Journal of Economic Geography*, *6*(1), 9–44.

Ottaviano, G. I. & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, *10*(1), 152–197.

Ottaviano, G. I., Peri, G., & Wright, G. C. (2018). Immigration, trade and productivity in services: Evidence from UK firms. *Journal of International Economics*, *112*, 88–108.

Ottaviano, G. I. & Puga, D. (1998). Agglomeration in the global economy: A survey of the "new economic geography". *World Economy*, *21*(6), 707–731.

Parey, M., Ruhose, J., Waldinger, F., & Netz, N. (2017). The selection of high-skilled emigrants. *Review of Economics and Statistics*, *99*(5), 776–792.

Paserman, M. D. (2013). Do high-skill immigrants raise productivity? evidence from Israeli manufacturing firms, 1990-1999. *IZA Journal of Migration*, *2*(1), 6.

Peri, G. (2012). The effect of immigration on productivity: Evidence from us states. *Review of Economics and Statistics*, *94*(1), 348–358.

Peri, G., Shih, K., & Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics*, *33*(S1), S225–S255.

Peri, G. & Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, *1*(3), 135–69.

Pessoa, J. P. & Van Reenen, J. (2014). The UK productivity and jobs puzzle: Does the answer lie in wage flexibility? *The Economic Journal*, *124*(576), 433–452.

Postel-Vinay, F. & Robin, J.-M. (2002). The distribution of earnings in an equilibrium search model with state-dependent offers and counteroffers. *International Economic Review*, *43*(4), 989–1016.

Preston, I. P. (2014). The effect of immigration on public finances. *The Economic Journal*, *124*(580), F569–F592.

Raaum, O. & Røed, K. (2006). Do business cycle conditions at the time of labor market entry affect future employment prospects? *The Review of Economics and Statistics*, *88*(2), 193–210.

Ransom, T. & Winters, J. V. (2016). Do foreigners crowd natives out of STEM Degrees and occupations? Evidence from the US Immigration Act of 1990. *IZA Discussion Paper*, (9920).

Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London*, *48*(2), 167–235.

Reder, M. W. (1963). The economic consequences of increased immigration. *The Review of Economics and Statistics*, 221–230.

Rojas-Romagosa, H. & Bollen, J. (2018). *Estimating migration changes from the eu's free movement of people principle* (tech. rep. No. 385).

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

Sá, F. (2014). Immigration and House Prices in the UK. *The Economic Journal*, *125*(587), 1393–1424.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics*, *116*(2), 681–704.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the economics of education* (Vol. 3, pp. 249–277). Elsevier.

Spengler, J. J. (1958). Effects produced in receiving countries by pre-1939 immigration. In *Economics of International Migration* (pp. 17–51). Springer.

Spitzer, Y. & Zimran, A. (2018). Migrant self-selection: Anthropometric evidence from the mass migration of italians to the united states, 1907–1925. *Journal of Development Economics*, *134*, 226–247.

Tallis, G. M. (1961). The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, 223–229.

Taylor, M. & Rampino, T. (2014). Educational aspirations and attitudes over the business cycle. *Economica*, *81*(324), 649–673.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531–543.

Thistle, P. D. (1993). Negative moments, risk aversion, and stochastic dominance. *Journal of Financial and Quantitative Analysis*, *28*(02), 301–311.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

U.S. Department of Education, I. A. O. (2008). Organization of U.S. education: The school level. Retrieved from http://www2.ed.gov/about/offices/list/ous/international/usnei/us/edlite-org-us.html

Universities UK. (2018). *International facts and figures*. Universities UK International. London, UK.

Viskanic, M. (2017). Fear and loathing on the campaign trail: did immigration cause Brexit? *Working Paper*.

Von Wachter, T. & Bender, S. (2006). In the right place at the wrong time: The role of firms and luck in young workers' careers. *American Economic Review*, *96*(5), 1679–1705.

Wadsworth, J. (2010). The UK labour market and immigration. *National Institute Economic Review*, *213*(1), R35–R42.

Wadsworth, J. (2013). Mustn't grumble: immigration, health and health service use in the UK and Germany. *Fiscal Studies*, *34*(1), 55–82.

Walker, I. & Zhu, Y. (2008). The college wage premium and the expansion of higher education in the UK. *Scandinavian Journal of Economics*, *110*(4), 695–709.

Walker, I. & Zhu, Y. (2018). University selectivity and the relative returns to higher education: Evidence from the UK. *Labour Economics.*

Westat & Mathematica Policy Research, I. (2001). *Survey of income and program participation users' guide* (third). U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau. Retrieved from https://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/SIPP_USERS_Guide_Third_Edition_2001.pdf

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420–445.