Workplace segregation and the labour market performance of immigrants*

Sébastien Willis†

July 2021

Abstract

Segregation of immigrants across workplaces has been widely documented, however the consequences of such segregation remain subject to conjecture. I use survey data matched to administrative records to study the effect of segregation in an immigrant's first job on her subsequent labour market outcomes. I argue that controlling for the wealth of pre-migration characteristics recorded in the survey data, not typically available in studies of immigrant outcomes, is sufficient to account for selection into high-conational firms. OLS and semi-parametric estimates indicate that a one percentage point increase in the share of conationals in an immigrant's first job is associated with 0.14–0.16 percentage point lower employment rates in the medium- to longer-term, while there is no clear evidence of a long-term earnings effect. Formal tests show that the results are robust to selection on unobservables. Differences in human capital acquisition do not appear to explain the employment effect, while there is some evidence that it is explained by differences in the quality of social network induced by differences in the initial workplace.

Keywords: Employment, segregation, coworker networks, immigrant earnings dynamics
JEL codes: J61, J64, J31

*An earlier version of this paper was the second chapter of my dissertation at Universitat Pompeu Fabra. I would like to thank my advisors, Albrecht Glitz and Ruben Enikolopov, as well as Caterina Alacevich, Lídia Farré, Flavio Hafner, Joan Monras, Matti Sarvimäki, and seminar participants at the UPF applied seminar, CREI international lunch, and UIL workshop on migration and integration for helpful feedback. Thanks also to Philipp vom Berge and the staff at the Research Data Centre of the Institut für Arbeitsmarkt- und Berufsforschung in Nürnberg for their support in accessing the data and to Parvati Trübswetter and Lucas Guichard for generating the data on conational shares. All results have been reviewed to ensure no sensitive information has been disclosed.

†Department of Economics, Uppsala University. sebastien.willis@nek.uu.se

1 Introduction

A growing body of evidence has documented substantial segregation across workplaces by country of origin in developed economies (Andersson et al., 2014; Åslund et al., 2014; Glitz, 2014; Hellerstein and Neumark, 2008). Not only do immigrants tend to be segregated from natives, they also tend to be segregated from other immigrant groups. However, evidence on whether workplace segregation might contribute to persistent wage and employment gaps between immigrants and natives (documented in Borjas, 1985; Chiswick, 1978; Lubotsky, 2007; Sarvimäki, 2011) is much scarcer. In particular, while previous work has noted a negative association between the degree of segregation of an ethnic group and average labour market outcomes of that group, there is still little evidence on the direct effect of the composition of a worker's current or past workplace on subsequent outcomes.

In this paper, I set out to address this gap in our understanding by studying how the ethnic composition of the set of coworkers in the first job held by an immigrant affects the immigrant's subsequent labour market outcomes. This question fits in a tradition of economists studying how the initial conditions upon an immigrant's arrival affect an immigrant's career path. Typically, prior research has focused on the initial place of residence and the relationship between the size of an immigrant's ethnic group in the initial location of residence and the immigrant's subsequent labour market outcomes (Battisti et al., 2018; Beaman, 2012; Damm, 2009; Edin et al., 2003; Munshi, 2003). The switch of focus, to the initial place of work and composition of the set of coworkers, is novel. It is motivated by recent evidence that coworker networks are a more important determinant of an individual's labour market outcomes than residential networks (Eliason et al., 2019). It also builds on evidence that an immigrant's firm identity accounts for as much as 40 per cent of the immigrant-native wage gap (Aydemir and Skuterud, 2008; Barth et al., 2012), mirroring the more general finding that firm identity explains a substantial portion of workers' wages and wage inequality (Abowd et al., 1999; Card et al., 2013, 2018; Song et al., 2019).[1] Here I study whether a particular time-varying characteristic of the initial firm, the conational share, has persistent long-term effects on an immigrant's labour market outcomes.

Previous work has shown that more segregated groups have worse labour market outcomes on average (Åslund and Skans, 2010; Glitz, 2014) and that higher conational shares in the first job are negatively associated with individual outcomes (Ansala et al., 2021). The first contribution of this paper is to go beyond associations and provide individual-level evidence on the effect of the conational share in the first job on subsequent outcomes. To study the effect of the initial conational share on later outcomes, I use a survey of immigrants in Germany, the migrant supplement of the German Socio-Economic Panel (SOEP), that has been linked to the respondents' social security records from the Institute for Employment Research (IAB). The IAB-SOEP dataset is unique in combining systematic information on employment histories in Germany, available in the administrative data, with a wealth of survey information on immigrants' pre-migration characteristics, including pre-migration employment status, German proficiency, or the presence

[1] Though Bonhomme et al. (2020) have recently argued that the contribution of firms to wage inequality is less than previously thought.

of social networks in Germany at migration. The central identification claim in this paper will be that the available set of pre-migration characteristics, the information on the major decisions an immigrant makes before finding a job, such as when and where to migrate, and information on the initial job-search process are jointly sufficient to explain selection into first jobs with either a high or a low share of conationals.

I find that starting out in a firm with a higher conational share has a negative effect on an immigrant's probability of being employed in the longer term. A one-percentage-point increase in the initial conational share leads, on average, to a 0.14–0.16-percentage-point lower employment rate six or more years after the first job. Importantly, the employment effect is specific to the conational share, and does not exist for immigrants who do not share the immigrant's nationality, suggesting that the underlying mechanism must be specific to the conational share. In contrast, there is at best only a short-term, positive association between the initial conational share and wages, and no longer-term effect of the composition of the initial place of work on wages. A formal test of selection on unobservables (Oster, 2019), confirms that the estimated employment effect is unlikely to be explained by selection on unobservables. Semi-parametric estimates of the effect, using variable selection methods for treatment effects in the presence of many interactions and transformations of the control variables (Belloni et al., 2012, 2014; Chernozhukov et al., 2015), also confirm that the results are robust to alternative assumptions on the functional form.

I review the evidence for different mechanisms that might explain my finding. I find that differential human capital accumulation, specifically learning German, is unlikely to explain the observed effect on employment rates. On the other hand, there is some evidence that differences in the quality of social network induced by working with more conationls may explain the observed findings. The conational share is persistent across time, suggesting that conationals in the first job may be an important source of information for finding subsequent jobs. Furthermore, the association between the conational share and subsequent employment rates varies with the local employment rate of conationals, which I take to be a proxy for the quality of one's conational coworkers.

Previous research on the earnings assimilation of migrants has shown that systematic selection into return migration with respect to realised earnings can bias estimates of the time profile of immigrant earnings in the host country (Borjas, 1985; Lubotsky, 2007). In this paper, I am interested in the related, but distinct, question of how the time profile of employment and earnings varies with the conational share in the first job. The second contribution of this paper is to formally show how systematic selection into outmigration, potentially caused by the variable of interest, in my case the initial conational share, can bias estimates of the effect of that variable on labour market outcomes. Selection into return migration with respect to the variable of interest is a potential source of bias in all studies of how initial conditions at migration affect subsequent outcomes (e.g. Azlor et al., 2020; Battisti et al., 2018; Beaman, 2012; Damm, 2009; Edin et al., 2003; Munshi, 2003). Importantly, this form of selection is independent of the more traditional selection into the treatment on unobservables that most previous research designs were intended to minimise.

I show that the sign of the selection bias will depend on the sign of both (i) the effect of the outcome of interest, say employment rates, on return migration; and (ii) the gross effect, i.e. without netting out any effect mediated by the outcome, of the variable of interest, the initial conational share, on return migration. Limitations in the data make it difficult to empirically assess the effect of the conational share on outmigration in my sample. However, under admittedly restrictive assumptions, I provide evidence that selection into outmigration on the initial conational share may imply that my results underestimate the full negative effect of the initial conational share on subsequent employment.

The paper proceeds as follows. In the following section I review predictions derived from different theories of wage determination about the effect of the composition of the set of coworkers in an immigrant's first job on an immigrant's later labour market outcomes, to structure the empirical analysis. In Section 3 I discuss the data used for this project. In Section 4 I present evidence on the association between initial workplace composition and subsequent employment rates and wages. In Section 5 I discuss different possible sources of bias: selection into the initial conational share on pre-employment characteristics, selection into return migration on the initial conational share, and model misspecification. In Section 6 I assess different possible mechanisms that could explain my result. Finally, Section 7 concludes.

2 Review of relevant theoretical predictions

In this section I review relevant theories that could explain any observed association between the initial conational share and labour market outcomes. I classify theories into two groups: (i) theories that predict there will be an association between the conational share when a worker starts a job and the starting wage in that job; and (ii) theories that predict there will be an association between the conational share when a worker starts a job and outcomes in later periods, such as wages and turnover—whether in the same job or in subsequent jobs—or unemployment. I call the former set of predictions contemporaneous associations and the latter long-term associations. I pay attention to whether a theory predicts that the conational share will cause an outcome, be caused by it, or simply be associated with it by sharing a common cause. I also pay particular attention to whether theories make different predictions for the conational share—the share of coworkers who are themselves immigrants from the same country of origin—and the other immigrant share—the share of coworkers who are also immigrants, but from other countries of origin.

2.1 Contemporaneous associations of workplace segregation

At a basic level, the conational share is likely to be associated with other firm characteristics. For example, it will be positively correlated with the total immigrant share and, given shares are capped at one, it is likely to be negatively correlated with the other immigrant share. It may also be negatively correlated with firm size, if immigrants tend to work in smaller, family-run firms. Inasmuch as these other firm characteristics are directly or indirectly associated with firm productivity, the conational share will also be associated with firm productivity. In models where workers' wages are an increasing function of firm productivity (e.g. Card et al., 2018; Manning, 2011), the conational and other immigrant share will then be associated with the starting wage of an immigrant.

Outside of more mechanistic relationships, the simplest form of contemporaneous association arises from a model of compensating differentials (Rosen, 1986; Sorkin, 2018). Immigrants might value the opportunity to work with conationals, and may as a consequence accept a lower wage to work in a firm where they get to work with relatively more conationals. Such compensating differentials, however, are presumably not present when working with immigrants from other countries—since immigrants may not feel much closer to immigrants from other countries than they do to natives—or at best will be strongly attenuated. In a static model of labour supply, compensating differentials will lower the wage an immigrant needs to be paid to work for a firm when the conational share is higher, although it is arguably the preference for working with conationals, not the conational share per se, that has a causal effect on wages. If a preference for working with conationals correlates negatively with unobservable individual productivity, the negative association between the conational share and the wage will likely be reinforced.

Firms might also take advantage of the networks of their employees in the hiring process to overcome information frictions. In a model where this is the case, a higher conational share has been shown to be a proxy for a newly hired immigrant having received a referral from another conational at the firm, which gives employers a more precise signal about the productivity of a match with a worker than hiring workers on the open market (Dustmann et al., 2016). Similarly, Åslund et al. (2014) argue that immigrant managers disproportionately hire other immigrants, relative to native managers, exploiting their superior information about immigrant workers.[2] In both cases, the higher conational share will be associated with more productive matches, raising the offered wage, although the association is arguably non-causal, as in the case of compensating differentials; it is the use of a referral that has a direct causal effect on wages. A higher other immigrant share is presumably not a proxy for having received a referral, or is at least a much worse proxy, so the other immigrant share will have at best a much smaller effect on wages via the use of referrals.

The conational share may also have a direct effect on the contemporaneous productivity of workers. Lazear (1999) has noted that mixed teams likely suffer from higher communication costs, either directly, due to the absence of a common language, or indirectly, due to the absence of a shared work culture, which creates friction or misunderstandings in the workplace, lowering productivity. Empirical work has also documented specific settings where diverse workplaces are less productive, including French supermarkets (Glover et al., 2017) and Kenyan flower factories (Hjort, 2014). Lazear (1999) argues that the existence of mixed teams implies that there must therefore be countervailing productivity gains to forming teams of workers from different cultural backgrounds. Evidence for complementarities in production between workers from different countries has been presented for specific industries, such as sport (Kahane et al., 2013), and at the aggregate level (Peri and Sparber, 2009; Ottaviano and Peri, 2012), however

[2] Åslund et al. (2014) do not distinguish between conational and non-conational immigrants, though the point is presumably most relevant for conationals.

there is as yet no evidence that all industries enjoy productivity benefits from workforce diversity at the firm-level.[3]

The net direct effect of the conational share on worker productivity is therefore ambiguous, however in models where workers' wages are an increasing function of their marginal product, any direct effect of the conational share on productivity will pass through to observed wages. Other immigrants are similar to natives in this model; they impose communications costs, but may be net complements. Increasing the other immigrant share, holding the conational share constant, therefore has similarly ambiguous effects on the productivity of an immigrant. However, one might expect that the effect of the conational share will be of the opposite sign to the effect of the other immigrant share since; if communications costs dominate the gains from complementarity, the conational share will have a positive effect and the other immigrant share will have a negative effect.[4]

2.2 Long-term associations of workplace segregation

The existence of compensating differentials implies that turnover will be lower in jobs where the conational share is higher, since a higher conational share raises the reservation wage for accepting another job offer. However, compensating differentials do not imply an effect of the conational share on involuntary (from the worker's perspective) separations, and hence unemployment. Furthermore, the wage effect of compensation differentials should be constant throughout the job, the preference for working with conationals being a fixed characteristic.

If firms learn about workers' productivity on the job, then the use of referrals implies that turnover from involuntary separations (from the worker's perspective) will be lower when the conational share is higher, since employers are less likely to receive negative news about a worker's productivity (Dustmann et al., 2016; Glitz and Vejlin, 2020). By decreasing the probability of an involuntary separation, the conational share will also be negatively associated with medium-term unemployment, assuming workers spend time searching after a separation. Furthermore, the referral-induced wage and turnover effects of the conational share will fade with tenure, since after a time workers hired in low conational-share firms, i.e. without a referral, will only stay in the job if the firm receives relatively good news about their productivity, and adjusts the wage accordingly.

By affecting the starting wage, the initial conational share also affects the starting position of the individual on the job ladder (Burdett and Mortensen, 1998). If either the conational share or the other immigrant share are associated with a lower starting wage, for any of the reasons discussed above, they will increase turnover in the short- to medium-run, as the worker moves up the job ladder. If job offers arrive at random and are drawn from the same distribution for

[3] In particular, it is not clear whether the documented aggregate complementarities derive from complementarities within production units such as firms, or through specialisation across firms.

[4] One might argue that immigrants are more likely to be substitutes rather than complements, even when they don't share a country of origin; in this case there are no productivity benefits to working with more non-conational immigrants, only communications costs, so the effect of the other immigrant share will be negative.

all workers, then the initial effect of a lower starting wage will fade out over time. However, if past wages affect subsequent wages, say because of wage bargaining where the current wage is the worker's threat point, then the effect of the starting wage may not fade out over the course of a career. A higher conational share could therefore lead to persistently lower wages if it lowers the starting wage, or higher wages if it raises the starting wage.

The use of social networks as a source of either information about job openings (Calvó-Armengol and Jackson, 2004; Boucher and Goussé, 2019) or referrals when applying for jobs (Montgomery, 1991; Galenianos, 2013; Dustmann et al., 2016) will also affect workers' wages and employment rates. Eliason et al. (2019) show that coworker networks are a particularly important determinant of labour market outcomes, more so than residential networks. Having a greater fraction of unemployed former coworkers has been shown to lower the rate of arrival of job offers for unemployed workers (Cingano and Rosolia, 2012; Glitz, 2017). It will also lower the probability of receiving a referral, since only employed workers can provide referrals, likely lowering the offered wage.

It is well-documented that immigrants have lower wages and are less likely to be employed than natives (e.g. Lubotsky, 2007; Sarvimäki, 2011). The initial conational share may, therefore, through its effects on network quality, lower the offer rate, leading to persistent differences in employment rates, and the distribution of offered wages, independently of whether the conational share affects the wage in the first job. The effect is likely to be heterogeneous by nationality; immigrants from groups with worse employment outcomes on average will be more negatively affected by starting out in a high-conational share firm. The other immigrant share will also have a negative effect on the employment and wages, although the size of the effect will depend on whether other immigrant groups have on average better employment outcomes than the worker's own group. If they have worse outcomes, other immigrants will be a worse source of information and referrals than the own group, and the negative effect of the other immigrant group will be larger in absolute value than the effect of the conational share.

Immigrants may also interact more intensively with their conational coworkers than with other types of workers, given the well-documented tendency towards homophily in the constitution of social networks (Mcpherson et al., 2001). In the terminology of Granovetter (1995), conationals might, therefore, be classified as strong ties and other workers as weak ties. Montgomery (1992) shows that if the offer rate from weak ties is higher, or the wage distribution of those offers stochastically dominates that of offers from strong ties, then increasing the share of weak ties in an individual's network will raise their reservation wage. A larger conational coworker share would therefore lower an individual's reservation wage. In particular, this effect is likely to be specific to the conational share, not the other immigrant share, since immigrants may be no more likely to interact with non-conational immigrants than with natives.[5]

The initial conational share might also affect subsequent outcomes through more traditional

[5]Empirical evidence on the value of weak ties is more mixed; strong ties appear more productive in the sense that an individual is more likely to end up working with a given strong tie than a given weak tie (Gee et al., 2017b,a), however this does not imply that having more strong ties leads to higher or lower wage offers on average, as predicted by the theory. The result is also subject to selection bias, since it relies on accepted jobs, not on all job offers.

human capital accumulation channels. Acquiring host country-specific human capital has been shown to account for a substantial portion of the convergence of immigrant wages to native wages over time (Eckstein and Weiss, 2010). Furthermore, Battisti et al. (2018) show that a higher share of conationals in the district of residence lowers the acquisition of host country-specific human capital in the longer run. They argue that this is because a larger share of conational co-residents makes job-finding easier, lowering the benefit from acquiring host country-specific human capital, though it is possible that it also raises the cost, e.g. of learnings the host country's language. A higher conational coworker share may also lower the benefit of acquiring host country-specific human capital, though the effect is likely to be attenuated relative to the co-resident conational share, since it concerns individuals who have already found a job. However, it clearly raises the cost of acquiring the host country's language, which could have a negative effect on long-term outcomes. The other immigrant share likewise probably only weakly affects the benefit of acquiring human capital, but, holding the conational share constant, it will probably raise the cost of learning the host country's language, since the worker interacts less with native speakers. This implies that the other immigrant share should also have a negative effect on long-term outcomes.

Finally, other characteristics of the initial firm may also influence the longer-term labour market outcomes of the worker. For example, starting one's career in a large firm has been argued to improve longer-term labour market outcomes (Arellano-Bover, 2020), perhaps because these firms provide more or better on-the-job training. If either the conational or other immigrant share is associated with these characteristics, they will be associated with the long-term outcomes of the worker, if these characteristics are not controlled for.

In sum, the starting conational share will have ambiguous effects on both initial wages and longer-term wages, although job-ladder models suggest both effects will be of the same sign. The other immigrant share will likely have a weaker effect, and may be of the opposite sign if there is no complementarity between immigrants of different origins. The effect of the conational share on subsequent employment, while also ambiguous, is a little easier to sign. By lowering the quality of job-finding network and the proportion of weak ties, and reducing incentives and increasing the cost of acquiring host-country specific human capital, a higher conational share is likely to increase the unemployment rate in the medium to long term. This is particularly likely to be true if one accounts for the method of finding the first job, removing any negative effect of the initial conational share, via the use of referrals, on job separations.

3 Data

This project uses the IAB-SOEP Migration Sample linked to administrative data of the Institute for Employment Research (officially, the IAB-SOEP-MIG-ADIAB), which is described in detail in Brücker et al. (2013). The IAB-SOEP Migration Sample is an annual survey of individuals in Germany with a migration background (i.e. immigrants or descendants of immigrants), conducted as a supplement to the German Socio-Economic Panel (SOEP). It contains much richer information about the survey respondents than is typically available in social security data. Particularly relevant to this project, individuals who emigrated to Germany are asked about when and under what circumstances they moved to Germany, their situation before moving to Germany, their language capacity and prior knowledge of people in Germany, and how they found their first job. The survey data are then, conditional on the consent of the respondents, linked to their social security records by the Institute for Employment Research (IAB).

The construction of the dataset from the SOEP surveys and its linking to the administrative data imply an important caveat when working with the data. The only waves of the SOEP linked to social security data and publicly released at the time of writing are from 2013 and 2014. The social security data are filled in retrospectively, from 1975 to 2014. This implies that survivors, those immigrants who do not return to the home country, will be disproportionately selected into my sample. Return migrants are generally negatively selected on ability or earnings (Borjas, 1985; Lubotsky, 2007; Sarvimäki, 2011), which implies that the individuals in my sample will tend to be positively selected on unobserved labour market ability or integration potential relative to the general population of immigrants. While this type of survivor bias is common to studies of immigrants, it is nevertheless important to note that this dataset is not exempt. I will discuss the possible effects of different forms of survivor bias when interpreting my results below.

The social security data cover all periods of benefit receipt, participation in job training programmes, and employment in a job covered by the social security system. This last condition means that the self-employed and civil servants are not covered; breaks in the social security data could be indicative of unemployment or employment in one of these categories. The data are reported as notifications, which record employment or benefit receipt spells to the day. I transform the data into an annual panel, starting from the immigrant's first year of social security-covered employment. In particular, I record the fraction of days worked in the calendar year, which I refer to as an individual employment rate, the total wage earnings from social security-covered jobs in the course of the calendar year, and a dummy variable for whether the individual was employed on June 30 of the given year. Employment notifications are associated with a unique establishment identifier. Establishments correspond to all production sites of a single employer in the same municipality in the same narrowly defined industry class. I follow standard practice when working with IAB data in referring to an establishment as a firm. All establishment-level variables in the IAB data are calculated at June 30.

I restrict my attention to the subset of individuals in the linked IAB-SOEP data who were born in a foreign country with a foreign nationality and who arrived in Germany between the ages of 15 and 64. Furthermore, individuals surveyed in the SOEP but who have never worked in a social security-covered job in Germany are by default excluded.

The final sample contains 863 individuals. I report summary statistics on the data in Table 1. All share variables are measured on $[0, 1]$, and wage and earnings variables are deflated to 2010 Euros. In Panel A I report time-varying information during the time following the first job in Germany. The individuals in the sample are employed a relatively high fraction of the time, particularly for immigrants, on average 74 per cent of the year. This can be attributed to positive selection into the sample, since individuals who never work a social security-covered job do not make it into the sample. In panel B I report some pre-migration time-invariant statistics. Half the sample are women and they are relatively educated on average. The average immigrant was 29 on arrival, and had a probability of 0.71 of being employed in the year before migrating; two-thirds of immigrants had support from someone in Germany at the time of migration, mostly family. In panel C I report characteristics of the first job held and the firm where it was held. The average first firm is large, at 470 workers, though the distribution (not shown) is highly skewed, with a long right tail. The social security data do not include hourly wages or hours worked, distinguishing only between full- and part-time jobs. However, notwithstanding the sample being positively selected, daily wages in the first job are on average substantially lower (43 Euros) than median daily wages in the firm (74 Euros). Just over half of my sample found their first job through contacts and they took on average 3.3 years to find that job after migrating. Finally, Table A.1 shows the frequency of the main nationalities in my sample. The individuals in my sample are more likely to come from more recent sending countries in Eastern Europe, such as Russia, Romania, and Poland, than former guestworker-sending countries such as Turkey, Italy, and Greece.

In my results I will focus on the long-term effects of the initial conational share. The conational share is defined as the share of coworkers on 30 June who share the same nationality as the worker. The initial conational share is the conational share in the year of the first job subject to social security an individual holds. The average initial conational share is 0.07, while the average initial other immigrant share is 0.17. However, the distribution of the initial conational share, shown in Figure A.1, is rather skewed. Around 55 per cent of immigrants in my sample do not have any conational coworkers at the start of their first job, while around 5 per cent of my sample start out working in a firm where more than 50 per cent of their coworkers are immigrants.

4 OLS analysis

4.1 Overview and identifying assumption

In this section I present evidence on the association between the initial conational coworker share and immigrants' subsequent labour market outcomes. I will regress an outcome of interest t years after the start of i's first job, Y_{it}, on the initial conational share s_i^{own}. For now I assume the outcome follows some nonparametric time trend, $f_2(t)$, and the effect of interest, $f_1(t)$, is likewise non-constant over time, and I include a quadratic in age as relevant control variables in X_{1it}.

$$Y_{it} = f_1(t) \times s_i^{own} + f_2(t) + \Gamma_1 X_{1it} + \epsilon_{it}. \tag{1}$$

The main threat to identifying the true causal effect of the initial conational share on subsequent outcomes is the possible existence of factors that (i) are pre-determined with respect to the initial conational share; and (ii) affect both the initial conational share and subsequent outcomes of interest. Obvious examples include individual preferences, such as a taste for working with conationals, and fixed characteristics, such as employability in Germany, as well as more

Table 1: Summary statistics

	Mean	St. dev.	N
Panel A			
Employment rate	0.74	0.38	10061
$P(Y > 1e4)$	0.57	0.49	10061
$P(Y > 2e4)$	0.37	0.48	10061
$P(Y > 3e4)$	0.18	0.39	10061
Annual wage earnings	21256.1	15024.9	7493
$\mathbf{1}(t \in [0, 2])$	0.25	0.44	10061
$\mathbf{1}(t \in [3, 5])$	0.23	0.42	10061
$\mathbf{1}(t \geq 6)$	0.52	0.50	10061
Panel B			
Woman	0.50	0.50	863
Age at migration	29.32	9.04	863
Employed before migrating	0.71	0.46	863
Education	0.14	0.34	863
Low education	0.40	0.49	863
Medium education	0.32	0.47	863
High education	0.29	0.45	863
Support (family)	0.47	0.50	863
Support (friends)	0.10	0.31	863
Support (both)	0.05	0.22	863
No support	0.37	0.48	863
Panel C			
First job through contacts	0.56	0.50	863
Years until first job	3.27	3.02	863
Daily wage	43.1	34.3	863
Firm size	470.4	2221.8	863
Firm median wage	74.3	39.5	863
Firm age	13.0	10.5	863
Conat. share	0.070	0.19	863
Other mig. share	0.17	0.20	863

Note: Panel A reports time-varying summary statistics for the years since the first job, average earnings are conditional on being employed on June 30. Panel B reports summary statistics on pre-migration characteristics. Panel C reports summary statistics on the characteristics of the first job held after migration and the firm where the job was held. Wages and earnings are deflated and reported in 2010 Euros.

aggregate characteristics, such as cohort effects, if the "quality" of immigrant is changing over time, nationality, or location of destination within Germany effects. There may also be individual characteristics that only indirectly affect the conational share that also directly affect subsequent outcomes. For example, the conational share is a proxy for having found a job through one's network (c.f. Dustmann et al., 2016); if less productive individuals are more likely to search for jobs through their network, this will also lead to endogeneity bias.

I address the possibility of selection on pre-employment characteristics through (i) the inclusion of fixed effects δ_j for aggregate characteristics j: nationality, year of arrival, and federal state (*Bundesland*) of first residence; and (ii) the inclusion of pre-migration characteristics available retrospectively from the SOEP, X_{2i}. Such detailed pre-migration information is not available in administrative data; its availability in the SOEP is the major advantage of using this dataset. The included characteristics are dummies for gender, being proficient in German before migration, for being employed in the year before migration, for whether the immigrant had pre-existing contacts in Germany before migrating, and for the three possible levels of education before migration, and quadratics in self-reported work experience prior to migration and age at migration.

To check how well the pre-migration characteristics and fixed effects capture selection into the first job, I regress other job and firm characteristics on the conational share, the pre-migration characteristics X_{2i}, and fixed effects δ_j, and report the coefficient on the conational share in each specification in Table 2. Of the job characteristics considered, only the dummy for whether the job was found through contacts is significantly associated with the conational share, conditional on included controls. The starting wage in the first job, in particular, is not associated with the conational share, conditional on these controls. The firm characteristics, on the other hand, are all significantly predicted by the conational share. However, this appears largely driven by the association of the conational share with establishment size. When I additionally control for log establishment size, in column two, the association between the conational share and other firm characteristics, with the exception of the other immigrant share, is substantially reduced.

I conclude from the results in Table 2 that the included pre-migration characteristics and fixed effects likely control for the possibly unobserved determinants of the main job characteristics. Nevertheless, I will include the time taken to find a job and a dummy for whether the job was found through contacts as controls in my main specification, since these may pick up the effect of some residual confounding variable not captured by X_{2i} and the fixed effects.[6] I also include the vector of initial firm characteristics, since these are clearly associated with the initial conational share, so any effect of the conational share should be interpreted as holding other observed firm characteristics constant. I call the vector of job-finding and firm characteristics X_{3i}.[7]

While the results presented in Table 2 are informative about the residual association between the conational share and job and firm characteristics, conditional on X_{2i} and δ_j, they do not

[6]The method of finding the first job is also a theoretically relevant control, since previous work has highlighted that the conational share can be a proxy for whether the job was found through a referral, which has a direct effect on wages and turnover (Dustmann et al., 2016)

[7]Note that the vector X_{3i} is determined simultaneously with the initial conational share, it is not an outcome of it; it is therefore not a bad control in the sense of Angrist and Pischke (2009).

Table 2: Association with other firm/job characteristics

	(1) $\beta^{s_i^{own}}$	(2) $\beta^{s_i^{own}}$
Job characteristics		
Job through contacts	0.23*	0.19*
	(0.09)	(0.10)
Years until first job	-0.00	-0.00
	(0.00)	(0.00)
log(Wage)	-0.19	0.03
	(0.16)	(0.17)
Apprentice	-0.04	-0.04
	(0.02)	(0.03)
Part-time	-0.11	-0.06
	(0.08)	(0.09)
Firm characteristics		
log(Firm size)	-3.40**	
	(0.25)	
log(Median wage)	-0.66**	-0.33**
	(0.09)	(0.09)
Firm age	-11.34**	-4.20*
	(1.58)	(1.65)
Firm age^2	-306.18**	-78.20
	(48.61)	(50.83)
Other mig. share	-0.15**	-0.16**
	(0.03)	(0.04)
N	863	863

Note: The table reports the estimated coefficient on the initial conational share for a series of regressions; each row corresponds to a different dependent variable. All regressions include controls for pre-migration characteristics and fixed effects for nationality, year of migration, and location of first residence. Robust standard errors reported. + p<0.1, * p<0.05, ** p<0.01

allow me to conclusively rule out that there is any selection into the treatment on unobservables that also affect the outcome. I will therefore formally test my identifying assumption by testing the claim that selection on unobservables is unlikely to explain the baseline association between s_i^{own} and Y_{it}, applying the method of Oster (2019). I will present the structure of Oster's test in greater detail in Section 5.[8]

Turning from identification to estimation, to make the estimation problem more tractable, I adopt a semi-flexible approach to modelling the functions $f_1(t)$ and $f_2(t)$. Ideally, I would like to model each as a set of indicator variables for all values that t takes on. However, since my sample is relatively small, I group years together and instead model both functions as a set of indicator variables for being within 0-2 years of the first job, 3-5 years of the first job, or more than 6 years of the first job. The final vector of controls, X_{it}, will subsume X_{1it}, X_{2i} and X_{3i} in a single control vector. The effect of the initial share of other immigrants is allowed to vary over time, just as the effect of the conational share does. In addition to being a relevant firm characteristic that is associated with the conational share, the initial share of other immigrants will be of special interest since its effect will help to adjudicate between the different theories presented in Section 2. The full specification is therefore

$$\begin{aligned} Y_{it} = \sum_{g \in \{own, other\}} & s_i^g \times \mathbf{1}(t \in [0,2]) + s_i^g \times \mathbf{1}(t \in [3,5]) + s_i^g \times \mathbf{1}(t \geq 6) \\ & + \mathbf{1}(t \in [0,2]) + \mathbf{1}(t \in [3,5]) + \mathbf{1}(t \geq 6) + \Gamma X_{it} + \sum_j \delta_j + \epsilon_{it}. \end{aligned} \tag{2}$$

Finally, turning from estimation to inference, in all specifications I cluster standard errors by individual. The treatment variable, s_i^{own} is technically assigned at the level of the firm by nationality by starting year. This would be the theoretically justified level at which to cluster standard errors (Abadie et al., 2017). However, since in a small sample the probability of two individuals of the same nationality starting a job in the same firm in the same year is close to zero, clustering observations at this level is essentially identical to clustering by individual.

4.2 OLS results

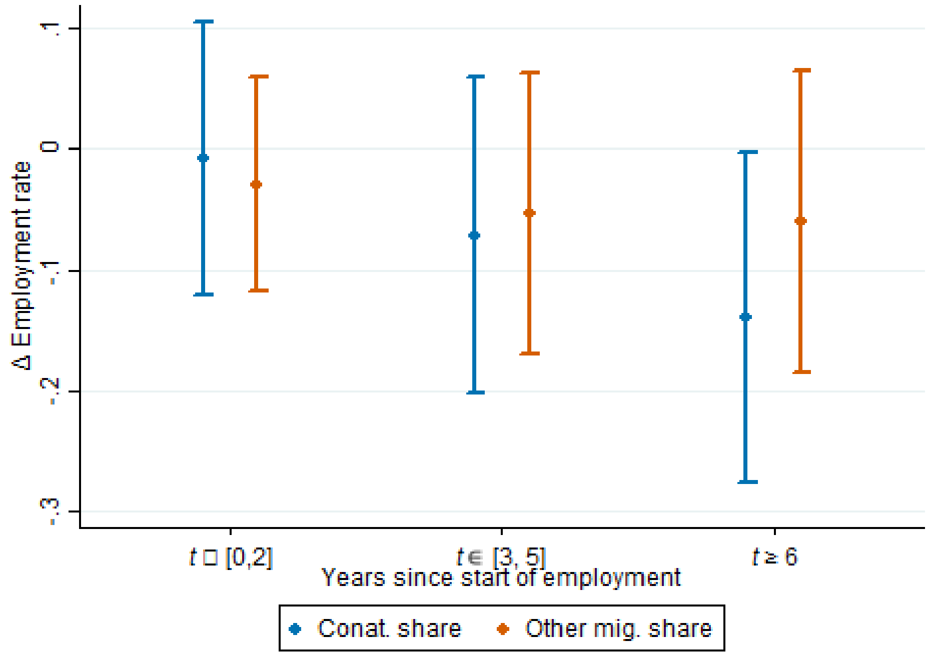
4.2.1 Employment rates

I first report estimates of the association between the starting conational share and individual employment rates. An individual's employment rate is defined as the fraction of days they are employed in a job covered by social security in a year. The results of the full dynamic

[8] In earlier versions of this paper, I have also considered instrumental variables estimates of the effect of the conational share on subsequent outcomes. Asylum seekers and ethnic Germans emigrating from Eastern Europe were subject to a dispersal policy on arrival. This implies that year-on-year variations in the composition of local labour demand, and in particular the expected share of conationals for someone hired in their year of arrival, are exogenous to subsequent labour market outcomes, and can be used as an instrument for the initial conational share. However, asylum seekers and ethnic Germans are a small subset of the sample (around 200 individuals). The instrument is not strong enough to predict the conational share in such a small sample.

specification defined in Equation (2) are plotted in Figure 1. The figure shows the effect on subsequent employment rates of both (i) the starting conational share; and (ii) the starting other immigrant share. While neither coworker share is associated with employment in the short run, given the included controls, the conational share is negatively associated with long-term employment rates; a one percentage point increase in the initial conational share leads to a 0.14 percentage point lower employment rate after six or more years. A one percentage point increase in the other immigrant share, on the other hand, leads to a statistically insignificant 0.06 percentage point decrease in employment rates. The difference between the two associations bears emphasising for at least two reasons. First, it suggests that the significant association between the conational share and subsequent employment rates cannot be explained by first jobs in firms with a higher immigrant share being of worse quality in some way that is not captured by the included controls (in particular firm size, median wage, worker starting wage and part-time status), since the association only exists with the own-group share, and not for other immigrants. Second, observing that only the conational share is associated with subsequent employment suggests that the mechanism underlying this association must be specific to the conational share.

Figure 1: Employment effect of composition of coworkers



Notes: Dynamic estimates of the employment effect of the initial conational share and the other immigrant share, using the specification defined in Equation (2). 95 per cent confidence intervals are calculated using standard errors clustered by individual and shown.

To put the magnitude of the long-term association into context, Glitz (2014) finds that the average employed immigrant in Germany in 2008 had 18 percentage points more conational coworkers than would be expected under a random allocation of workers, or 13 percentage points after partialling out the effects of region of residence, gender, education, and industry. The unemployment rate of the foreign-born in Germany at the time was 12.3 per cent, 5.8 percentage points higher than the unemployment rate of the native-born (OECD, 2020). Scaling the long-term effect of the conational share by average segregation translates to a $0.14 \times 18 = 2.5$ percentage point lower employment rate, or 1.8 percentage points if observable characteristics are partialled out of the measure of segregation. The magnitude of the long-term association between the initial conational share and unemployment is therefore large relative to the difference in employment rates between immigrants and natives in Germany.

While the results in Figure 1 show that immigrants are less likely to be in a job subject to social security in subsequent years if their first job is in a high conational share firm, it is not possible to assert based on this result alone that the individual is more likely to be unemployed. In Table A.2 I explore other measures of an individual's labour force status as outcomes, including my full set of controls X_{it} and fixed effects in all specifications. In columns 1–4 I consider measures drawn from the administrative data: share of days in a year of benefit receipt, share of days as a registered job seeker, share of days in a job training program, and a dummy for being out of the social security system altogether. Only the last of these variables is (positively) associated with the conational share. Individuals out of the social security system might be genuinely out of the labour force, or they might be in self-employment or civil servants. In columns 5–6 I draw on the SOEP survey, which for the years 2013 and 2014 asks if individuals are employed and, if so, in what activity. In particular, I define dummy variables equal to one for individuals who report either self-employment or working as a civil servant. While the sample is much smaller, there is no economically or statistically significant long-run association between these variables and the initial conational share. I therefore conclude that a higher initial conational share is associated with an increased probability of an individual dropping out of the labour force in the longer term.

4.2.2 Wage earnings

In Table 3 I repeat the full specification, again including fixed effects and controls for pre-migration, initial job and initial firm characteristics, for different measures of wages and earnings. The social security data only include daily wages and an indicator for part-time status. In column 1 I therefore estimate the association between the initial conational share and average daily earnings, defined, for individuals who work at least one day during the year, as total earnings subject to social security in a year divided by total number of days worked, deflated to 2010 values. There does not appear to be a significant relationship between the initial conational share and average daily wages, conditional on employment. The estimated magnitude is also not particularly large; a one-percentage-point increase in the initial conational share increases earnings 0.2 log points. To account for any possible effect of the initial conational share on average daily hours worked, I repeat the estimation for respectively full- and part-time workers. Part-time status and the daily wage are here measured on June 30 of a given year, the results are reported in columns 2 and 3 of Table 3. While the initial conational share is positively associated with daily wages of full-time workers in the short-term, the magnitude is again relatively small (a 0.17 log-point increase in wages for a one-percentage-point increase in the conational share) and there is no longer-term association. For part-time workers there is no significant association at any horizon.

While the evidence reported in columns 1–3 of Table 3 suggests there is little significant association between the initial conational share and earnings, these estimates suffer from a type of selection bias. Individuals who are employed, whether full-time or part-time, in spite of having a high conational share in their first job are potentially positively selected on unobserved employability relative to other immigrants, introducing a conditional-on-positive selection bias (Angrist and Pischke, 2009). This kind of selection would likely bias the estimated association between the initial conational share and potential subsequent earnings upward relative to the true association in the full, unobservable, population. As such, it is not possible to conclude whether the true effect of the initial conational share on wages is positive but smaller than the estimated effect, zero, or negative but biased toward zero when conditioning on individuals being employed.

To avoid conditional-on-positive selection bias, in columns 4–7 of Table 3 I use the full sample and regress a dummy for annual earnings being above a series of cutoffs on the initial conational share and the full set of controls and fixed effects. This approach is conceptually similar to a quantile regression, however the interpretation of regression coefficients is more straightforward. The cutoffs I consider are 0, 10,000, 20,000, and 30,000 Euros. An increased conational share does not appear to uniformly shift the distribution of earnings. There is some evidence of a positive short-run association between the initial conational share and earnings; in particular a one-percentage-point increase in the initial conational share increases the probability of earning more than 30,000 Euros by a statistically significant 0.14 percentage points. Given the absence of employment effects at this horizon, and given the firm and pre-migration characteristics controlled for, including whether the job was found through a contact, this positive association suggests that immigrants do earn higher wages when working with more conationals, perhaps because they are more productive.[9] The initial conational share is negatively associated with long-term earnings, though given the magnitude of this effect is broadly in line with the negative long-term employment effects documented above, it is not possible to conclude from this that there is any strong evidence of a long-term wage effect of the initial conational share. Note that there is again no effect of the share of immigrants from other countries on earnings.

The finding of a clear negative effect of the starting conational share on long-term employment and at best only a transient and, if anything, positive wage effect is consistent with the finding that the total earnings gap between immigrants and natives is mostly due to differences in employment, not wages conditional on employment (Sarvimäki, 2011). It is also broadly in line with the theoretical mechanisms reviewed in Section 2, where I argued that different theories made conflicting predictions about the short- and long-term wage effects, but a more clear prediction of a negative longer-term employment effect of the initial conational share. Having

[9]This is in line with evidence that immigrants might be more productive when working in more homogeneous teams (Lazear, 1999; Hjort, 2014) or when working with managers who share their origin Åslund et al. (2014); Glover et al. (2017), however I do not observe the occupations of coworkers, so cannot test this mechanism.

Table 3: Relation between initial coworkers and earnings

	(1) ln(Avg. wage)	(2) ln(Wage) | FT	(3) ln(Wage) | PT	(4) $P(Y>0)$	(5) $P(Y>1e4)$	(6) $P(Y>2e4)$	(7) $> P(Y>3e4)$
$\mathbf{1}(t \in [0,2]) \times$ Conat. share	0.19	0.17*	0.43	0.083+	0.036	0.078	0.13*
	(0.13)	(0.083)	(0.29)	(0.047)	(0.061)	(0.061)	(0.050)
$\mathbf{1}(t \in [3,5]) \times$ Conat. share	0.0015	0.053	0.22	-0.017	-0.047	-0.065	0.092
	(0.13)	(0.12)	(0.26)	(0.067)	(0.078)	(0.074)	(0.060)
$\mathbf{1}(t \geq 6) \times$ Conat. share	-0.22	0.042	-0.17	-0.099	-0.15+	-0.16+	-0.017
	(0.16)	(0.14)	(0.27)	(0.070)	(0.082)	(0.086)	(0.075)
$\mathbf{1}(t \in [0,2]) \times$ Other mig. share	0.089	0.19*	0.21	0.0016	0.050	-0.019	-0.0044
	(0.12)	(0.079)	(0.22)	(0.035)	(0.067)	(0.053)	(0.042)
$\mathbf{1}(t \in [3,5]) \times$ Other mig. share	0.071	0.069	-0.46	-0.057	0.025	-0.041	0.018
	(0.13)	(0.089)	(0.29)	(0.058)	(0.075)	(0.066)	(0.049)
$\mathbf{1}(t \geq 6) \times$ Other mig. share	0.025	0.037	-0.43	-0.051	-0.042	0.012	0.026
	(0.15)	(0.11)	(0.27)	(0.059)	(0.078)	(0.077)	(0.068)
Observations	8560	5027	2413	10061	10061	10061	10061
Individuals	863	704	542	863	863	863	863
R^2	0.34	0.44	0.26	0.12	0.21	0.30	0.32

Note: OLS estimates of relationship between initial conational share and subsequent earnings. Y refers to annual labour earnings covered by social security. The regression for average earnings in column 1 is estimated conditional on an individual being employed in a job covered by social security at least one day during the year, daily wages in columns 2 and 3 are measured on June 30 of the relevant year. All coefficients are estimated using the specification defined in Equation (2), standard errors are clustered by individual. + p<0.1, * p<0.05, ** p<0.01

established that there is a significant negative association between the initial conational share and employment rates, I now turn to assessing possible sources of bias that could explain this finding.

5 Robustness of the identifying assumption to possible sources of bias

In this section I review the main threats to the claim that the negative association between the initial coworker share subsequent employment rates reported in Section 4.2 can be interpreted causally.

5.1 Selection on unobservables into the treatment

5.1.1 Heuristic test

The central identification claim of this paper is that the extensive set of controls before and at migration included in my main specification, made possible by the information gathered in the SOEP, allow me to plausibly control for unobserved pre-employment characteristics that might lead to selection into a first job with a higher or lower conational share. A heuristic way of testing this claim is to compare the effect of the conational share on earnings both including and excluding the different sets of controls. In Table 4 I report estimates of the association between the starting conational share and individual employment rates. Before comparing estimates of the dynamic effects defined in Equation (2), I first estimate the average effect of the initial conational share on subsequent employment, first without controls (column 1), then with controls, including the other immigrant share (column 2). I find that a one-percentage-point increase in the conational share is correlated with a 0.14-percentage-point lower employment rate on average without controls, and a 0.09 percentage-point lower employment rate with controls.

To better understand possible selection into first firms employing a relatively high or low share of conationals, columns 3–6 of Table 4 introduce the different sets of controls one by one. In column 3 I report estimates of the dynamic effect of initial conational share, controlling only for an individual's age and age squared. The conational share is negatively associated with subsequent employment rates at all horizons. In column 4 I include the pre-migration characteristics from the SOEP as controls. The effects are not statistically different from column 3 and even *increase* slightly in magnitude when the pre-migration characteristics are included, suggesting individuals whose pre-migration characteristics are associated with higher employment rates are more likely to work in high-conational share firms. In column 5 I add fixed effects for nationality, cohort (i.e. year of migration) and initial state of residence. The short-term association in particular increases to -0.06 and is no longer significant. The medium-term coefficient also increases somewhat, but remains significant, while the long-term coefficient decreases fractionally.

Finally, I add other characteristics of the initial job search and firm, estimating the full dynamic specification defined in Equation (2), and report the results in column 6. As already

Table 4: Relation between initial coworkers and employment

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
Conat. share	-0.14*	-0.089^{+}				
	(0.055)	(0.053)				
$\mathbf{1}(t \in [0,2]) \times$ Conat. share			-0.12**	-0.12**	-0.059	-0.0080
			(0.044)	(0.045)	(0.052)	(0.058)
$\mathbf{1}(t \in [3,5]) \times$ Conat. share			-0.16*	-0.17*	-0.12^{+}	-0.071
			(0.066)	(0.066)	(0.063)	(0.067)
$\mathbf{1}(t \geq 6) \times$ Conat. share			-0.14^{+}	-0.15^{+}	-0.16*	-0.14*
			(0.079)	(0.078)	(0.066)	(0.070)
Other mig. share		-0.050				
		(0.049)				
$\mathbf{1}(t \in [0,2]) \times$ Other mig. share						-0.030
						(0.045)
$\mathbf{1}(t \in [3,5]) \times$ Other mig. share						-0.053
						(0.059)
$\mathbf{1}(t \geq 6) \times$ Other mig. share						-0.060
						(0.064)
Premigration controls	No	Yes	No	Yes	Yes	Yes
Firm controls	No	Yes	No	No	No	Yes
Job controls	No	Yes	No	No	No	Yes
Observations	10061	10061	10061	10061	10061	10061
Individuals	863	863	863	863	863	863
R^2	0.01	0.13	0.01	0.03	0.12	0.13
FE	No	Yes	No	No	Yes	Yes

Note: OLS estimates of relationship between initial conational share and subsequent individual employment rates. The individual employment rate is the fraction of days in a year an individual is employed. All specifications include a quadratic in age. Standard errors are clustered by individual. + p<0.1, * p<0.05, ** p<0.01

shown in Figure 1, the short-term association is now indistinguishable from zero, suggesting that the short-term association between conational share and employment rates observed in column 3 can be entirely explained by selection on observable characteristics into high-conational share firms and by the correlation of the conational share with other job and firm characteristics. The medium-term association is also nearly halved by the inclusion of the full set of controls and is not significant. The longer-term association, however, is quite robust to the inclusion of all controls and fixed effects; it is -0.14 both when only age is included, in column 3, and when all controls are included, in column 6, and is significant at the 5 per cent level in the full specification.

5.1.2 Overview of formal test

It is possible to evaluate the evidence presented in Table 4 more formally using the test proposed by Oster (2019). Similarly to the heuristic test, the formal test involves comparing (i) a non-causal association between a variable of interest and an outcome that might be at least partially explained by selection on some variable; and (ii) an association between the same variable of interest and outcome, this time controlling for variables that are thought to measure the characteristics on which selection takes place. Because selection is thought to explain a part of the uncontrolled association, one expects the coefficient of interest to change between (i) and (ii). However, if there truly is an underlying causal effect, this change should not be "too big". Just how big is too big is determined by scaling the change in the estimated treatment effect when controls are included by the change in the R^2. Intuitively, only treatment effects that are robust to the inclusion of covariates that actually explain the outcome should be labelled robust. Altonji et al. (2005) used this insight to develop an estimator of the ratio between (a) the (unobserved) covariance between the variable of interest and the unobserved confounders and (b) the covariance between the variable of interest and the observed confounders that would make the true causal effect of the variable of interest zero.

To construct their test, Altonji et al. (2005) assume that the R^2 of the regression would be one if all confounders were included. Oster (2019) observes that this is unduly restrictive if there is an idiosyncratic component to the outcome of interest or if variables are measured with error. She therefore develops a generalised version of the test that allows the maximum R^2 to be less than one. The output of Oster's test is again an estimate of the ratio between (a) the covariance between the unobserved confounders and the treatment variable; and (b) the covariance between the treatment variable and the included confounders that would be consistent with the true treatment effect being zero. I refer to this estimated ratio as Oster's δ.

The maximum possible value of the regression R^2, i.e. when all observed and unobserved confounders are included in the regression, R_{max}, is a key ingredient in estimating Oster's δ. Oster (2019) suggests that $R_{max} = \min\{1.3 \times \tilde{R}, 1\}$, where $\tilde{R}$ is the R^2 from the long regression including all controls, is a reliable benchmark. Reviewing evidence from randomised experiments, where selection on unobservables can be ruled out *a priori* if randomisation succeeded, she finds that using this value of R_{max} would lead the researcher to conclude that 10 per cent of experimental results were due to selection on unobservables. In reviewing a selection of articles from well-known journals, she finds that around 50 per cent of published effects would be explained by selection on unobservables using this standard.

5.1.3 Results of formal test

The calculation of Oster's δ is only defined for a scalar treatment variable. I therefore report the estimated δ both for selection on unobservables in the time-invariant specification, and separately for the each time horizon in the dynamic specification. To estimate the δ, the researcher must also specify the set of controls that are intended to capture selection into the treatment. I am principally concerned about individuals selecting into high- or low-conational share first jobs based on unobservable characteristics that are predetermined relative to their taking up those jobs. I argue that my included pre-migration characteristics, drawn from the SOEP, and characteristics at migration, captured by my fixed effects for year of migration, location of arrival, and nationality, are good controls for unobservable predetermined individual characteristics. However, I argued previously that initial firm and job characteristics may also capture some residual selection. I therefore only include age and age squared in the short regression (and, in the dynamic specification, the interactions of the initial conational share and years since migration that are not being tested for selection on unobservables).[10]

Table 5 reports the estimated values of Oster's δ for my employment specification in column 1. In the static specifications, in the first row, the δ for the employment regression is 1.8. Following Altonji et al. (2005), Oster (2019) suggests that 1 is a reasonable cutoff for declaring results robust to selection on unobservables, since $\delta < 1$, implies that the true treatment effect could be zero even if there is less selection into treatment based on the unobservables than on the observables. The value in the static specification is above the cutoff; for the observed association to be explained by selection on unobservables, these unobservables would have to be almost two times as strongly correlated with the initial conational share than the observables are. The pattern of estimates of δ for the dynamic effects clearly mirrors the pattern that emerged from the heuristic test. The short-term δ is close to zero, the effect is not at all robust to selection on unobservables, while the medium- and long-term effects are increasingly robust, the value of δ in these two cases is 1.91 and 5.96. I can therefore conclude with a high degree of confidence that the negative long-term effect of the initial conational share is robust to selection on unobservables.

In columns 2 and 3 I report estimates of δ for the effect of the conational share on wages conditional on either full- or part-time employment, for comparison In this case, the estimated δ is typically negative. This occurs because the included covariates push the estimated effect *away* from zero; the unobservables would therefore have to push the estimated effect in the other direction for the true effect to be zero. Unlike the test of Altonji et al. (2005), Oster's δ is well-defined in the case where the included covariates increase the estimated effect; the effect is now declared to be robust to selection on unobservables if $\delta < -1$. The wage effects are not particularly robust to selection on unobservables; the static δ for full-time workers is -1.33,

[10] My conclusions about the likelihood of selection on unobservables do not change if I focus only on my pre-migration characteristics and fixed effects and either include firm and job controls in both regressions, or exclude them from the calculation of δ entirely.

Table 5: Estimates of Oster's δ

	$P(E=1)$	$ln(Wage)\|FT$	$ln(Wage)\|PT$
Average effect	1.80	-1.33	-0.22
$t \in [0,2]$	0.18	-1.48	-1.79
$t \in [3,5]$	1.91	-0.54	-1.13
$t \geq 6$	5.96	7.11	0.75
Π	1.3	1.3	1.3
R^2_{max}	0.17	0.57	0.34
N	10061	5027	2414

Note: Estimates of Oster's δ, the ratio of the selection on the observable to the selection on the unobservables implied by model estimates and an assumed value of R^2_{max}. I assume $R^2_{max} = \min\{\Pi\tilde{R}, 1\}$, where $\tilde{R}$ is the R^2 of the long regression, including all controls. For details, see main text.

which is marginally robust, while $\delta = -0.22$ for the static specification for part-time workers. When looking at the values of δ in the dynamic specifications, the short-term effect appears most robust to selection on unobservables, as $\delta < -1$ in both cases. $\delta = 7.11$ in the long-term for full-time workers, however, the point estimate increases from 0.031 without controls to 0.042 with controls, both of which are economically insignificant, suggesting the large magnitude of δ in this case has no practical implications.[11]

The results of these tests for selection on unobservables strengthen the claim that the associational effect of the initial coworker share on subsequent employment, estimated in Section 4.2.1, likely captures the true causal effect. In particular, they provide formal support for the claim that the rich set of pre-migration characteristics, including pre-migration employment, work experience, proficiency in German, having contacts in Germany before migrating, and fixed effects capturing differences across cohorts, nationalities, or location of arrival in Germany, adequately capture selection into high- or low-conational share firms.

5.2 Robustness of functional form

The results in the previous section establish that the set of conditioning variables, which includes the pre-migration characteristics, and features of the job search and of the first firm, among others, is sufficient to capture any selection into firms with a high or low share of conationals. However, even if the conditioning set contains the relevant determinants of selection, there is no guarantee that the functional form assumed in Equation (2) correctly models the selection into

[11]Since the point estimate in this case moves away from zero as controls are included, the estimate of δ should in theory be negative. Bevis et al. (2020) find that the Stata command `psacalc`, which estimates δ is unstable when $\delta < 0$. This presumably explains why the sign is wrong in this case, although the magnitude may still be correct, given the large increase in the R^2, from 0.067 to 0.44, when controls are included.

firms. For example, there may be important interactions between some pairs of variables in the conditioning set that determine both selection into the initial firm and later employment rates that are not properly accounted for when both variables are included separately.

To test for such a possibility, I re-estimate the effect of the initial conational share on employment rates via post-regularisation (Belloni et al., 2013; Chernozhukov et al., 2015). The effect on both the outcome variable—the employment rate—and the treatment variables of interest—the interactions of the initial conational share with the years since migration dummies—of a high-dimensional set of control variables, which includes many interactions between the variables included in the baseline linear specification, is estimated via LASSO. The residuals from the regression of the outcome on the controls is then regressed on the residuals from the regressions of the treatments on the controls.[12] The results of this procedure are reported in Figure A.2. The estimated effect is, if anything, larger than the OLS estimates. In both the medium- and longer-term a one-percentage point increase in the initial conational share is associated with 0.16 percentage point lower employment rates, a result that is significant at the five per cent level. These results can reassure us that including the controls linearly, as in Equation (2), is not leading to any false positives.

Another type of model misspecification bias could also arise if the effect of the conational share on subsequent outcomes is nonlinear or even nonomonotonic. Indeed Ansala et al. (2021) provide descriptive evidence that the number of months worked in the six years after the first job is increasing in the coworker share for low values of the coworker share and decreasing for values of the coworker share above 5–10 per cent. I explore such a possibility in Figure A.3, where I plot the average employment rate for different categories of the initial conational share, conditional on included controls. All averages are expressed as deviations from the employment rate of individuals whose initial conational share is less than 5 per cent in their first two years of employment.[13] There does appear to be some mild evidence of a nonmonotonic effect. Long-term employment are somewhat higher for individuals who start out in a firm where the conational share ranges from 5–10 per cent than for individuals whose initial conational share ranges from 0–5 per cent and are lowest for individuals whose initial conational share ranges from 10–50 per cent. Nevertheless, the assumption of a linear effect does not appear to be too bad of an approximation.

5.3 Selection on the treatment into return migration

I have already noted that my sample is made up of survivors, immigrants who were still in Germany in 2013 and 2014 in order to be interviewed. It is generally accepted that return

[12]Post-regularisation can therefore be described as a non-parametric analogue to the Frish-Waugh-Lovell theorem. The full details of the procedure, the the full set of included variables and the retained variables are described in Appendix B.

[13]These averages are estimated by replacing the interactions of the conational share with years since migration in Equation (2) with a full set of interactions between the years since migration and a set of dummies for the base year immigrant share taking values from $[0, 5)$, $[5, 10)$, $[10, 50)$, $[50, 90)$, and $[90, 100]$; individuals with a conational share in the 0–5 per cent range in their first two years of employment are the omitted category.

migrants had worse labour market outcomes, summarised by earnings, before returning than immigrants who stay (Borjas, 1985; Lubotsky, 2007; Sarvimäki, 2011). This tells us that earnings have a negative effect on return migration, or that return migration and earnings share some common unobservable cause—return migrants might be intrinsically less productive individuals—either of which can bias estimates of the rate of earnings convergence of immigrants to natives over time (Abramitzky et al., 2014). However, when studying the effect of some initial condition, whether the ethnic network at migration or, as in my case, the conational share in the first job, on subsequent labour market outcomes, the sign of the selection bias will depend not only on the effect of earnings on return migration, but also on any effect the initial conational share might have on return migration.

5.3.1 Formal derivation of the sign of the bias

To focus on intuition and to emphasise the fact that the bias induced by selective return migration is independent of the bias induced by selection into treatment on unobservables, I derive the sign of the selection bias under the simplifying assumption that (i) the initial conational share, S is randomly assigned; and (ii) there are no systematic determinants of subsequent employment rates Y besides S. Furthermore, assume that the conational share is either low or high, i.e. $S \in \{0, 1\}$. Assuming the effect of S on Y is linear, the structural equation for Y is simply:

$$Y = a + \beta S + \varepsilon_Y. \tag{3}$$

The structural error term ε_Y is mean-zero[14] and independent of S, since there is no confounding. To model selection, I assume that latent utility C^* is a linear function of S, Y, and a mean-zero structural error term:

$$C^* = \alpha_S S + \alpha_Y Y + \varepsilon_{C^*}, \tag{4}$$

where $\alpha_i \in \mathbb{R}$, $i \in \{Y, S\}$. An individual is assumed to return migrate, $C = 1$, if latent utility is below some fixed threshold:

$$C(S, Y) = \begin{cases} 1 & \text{if } C^* < K, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Equation (5) captures the fact that C is endogenously determined by both S and Y. The sign of α_i, $i \in \{Y, S\}$, encodes hypothetically testable assumptions about the effect of the observable variables Y and S on C. I now show how the selection bias from conditioning the analysis on $C = 0$ depends on the signs of α_S, α_Y, and β. Since the structural equation is linear and S is assumed to be randomly assigned, the true parameter of interest, β, can be defined as

$$\beta = \frac{\mathrm{Cov}(Y, S)}{\mathrm{Var}(S)} \tag{6}$$

[14] Furthermore, we must have $\varepsilon_Y \in [-a, 1 - (a + \beta)]$, since $Y \in [0, 1]$

Since we only observe individuals with $C = 0$, however, the OLS estimand on this restricted sample is

$$\begin{aligned}
\hat{\beta} &= \frac{\text{Cov}(S, Y|C = 0)}{\text{Var}(S|C = 0)} \\
&= \beta + \frac{\text{Cov}(S, \varepsilon_Y|C = 0)}{\text{Var}(S|C = 0)} \\
&= \beta + \frac{\text{Cov}(S, \varepsilon_Y|C^* \geq K)}{\text{Var}(S|C^* \geq K)} \qquad (7)
\end{aligned}$$

The sign of the bias induced by conditioning on the endogenous variable C will therefore depend on the sign of the conditional covariance of S and ε_Y, since the conditional variance of S is positive. Note that $\text{Cov}(S, \varepsilon_Y) = 0$ in the full sample by assumption, but not in the restricted sample of non-return migrants. The sign of the conditional covariance can be calculated as

$$\begin{aligned}
&\text{Cov}(S, \varepsilon_Y|C^* \geq K) \\
&\quad = \text{E}[S\varepsilon_Y|C^* \geq K] - \text{E}[S|C^* \geq K]\text{E}[\varepsilon_Y|C^* \geq K] \\
&\quad = \text{E}[\varepsilon_Y|C^* \geq K, S = 1]\text{Pr}(S = 1|C^* \geq K) \qquad (8) \\
&\qquad - \text{E}[S|C^* \geq K]\text{E}[\varepsilon_Y|C^* \geq K] \\
&\quad = \{\text{E}[\varepsilon_Y|C^* \geq K, S = 1] - \text{E}[\varepsilon_Y|C^* \geq K]\}\text{Pr}(S = 1|C^* \geq K), \qquad (9)
\end{aligned}$$

where the second equality follows from the law of iterated expectations and the third from the fact that S is a Bernoulli random variable, so its expectation is the probability that $S = 1$. The sign of the conditional covariance will depend on the sign of the difference of the two conditional expectations in parentheses in Equation (9), $\text{E}[\varepsilon_Y|\cdot]$. Note, however, that ε_Y is a mean-zero random variable and that its distribution is truncated when calculating the expectations $\text{E}[\varepsilon_Y|\cdot]$. The sign of the conditional expectations will therefore depend on whether the right or the left tail of the distribution is truncated. Furthermore, the difference between the expectations will depend on which distribution is more severely truncated. The truncation condition $C^* \geq K$ can be re-written

$$\alpha_Y \varepsilon_Y \geq K - (\alpha_S + \alpha_Y \beta)S - \alpha_Y a - \varepsilon_{C^*}, \qquad (10)$$

This inequality makes clear how the sign of the bias of $\hat{\beta}$ with respect to β will depend not only on (i) the total effect of employment on return migration, captured by α_Y; but also potentially on (ii) the total effect of the conational share on return migration, that is without netting out the part of the effect that is mediated by employment, i.e. $\alpha_S + \alpha_Y \beta$. Intuitively, the sign of α_Y determines whether the distribution of ε_Y is left- or right-truncated, and the sign of $\alpha_S + \alpha_Y \beta$ determines whether the distribution is more or less severely truncated when $S = 1$. If both α_Y and $\alpha_S + \alpha_Y \beta$ are of the same sign, the bias will be negative, while if α_Y and $\alpha_S + \alpha_Y \beta$ are of opposite signs, the bias will be positive. A formal proof of this claim is in Appendix C.

An interesting special case arises when the true effect of interest $\beta = 0$. Now the gross effect of the conational share on return migration is simply the direct effect, α_S. In this case, if α_Y and

α_S are of the same sign, then $\hat{\beta} < 0$, while if they are of opposite signs, then $\hat{\beta} > 0$. Therefore, if the estimated $\hat{\beta} < 0$ and one has reason to believe that α_Y and α_S are of opposite signs, then the observed association cannot be entirely explained by selection into return migration; it must be that $\beta < 0$.

5.3.2 Evidence of selection on the treatment into return migration

In Section 4.2 I estimated that $\hat{\beta} < 0$. Assuming the selection bias is not so strong as to change the sign of the effect, one could conclude that $\beta < 0$. There is good evidence that lower earnings and employment make an individual more likely to re-emigrate (Lubotsky, 2007; Sarvimäki, 2011; Abramitzky et al., 2014), implying that $\alpha_Y > 0$, i.e. the opportunity cost effect dominates the effect of any target savings behaviour. All that remains to be determined is the sign of α_S, the association between S and C^* after partialling out Y. In a dataset that does not contain any return migrants, at least at the time of observation, it is not possible to show direct evidence on the sign of α_S. Nevertheless, it is possible to provide indirect empirical evidence on the relationship between the initial conational share and selection into outmigration by comparing cohorts that were first employed in Germany more or less recently in the year the individuals were sampled, in my case, 2014. If there are no year-of-first-employment effects, i.e. the starting conational share is the same for all newly employed cohorts, and a higher initial conational share induces greater rates of return migration, then the initial conational share will be higher in more recently employed cohorts, as fewer of the individuals who started out in a high-conational share firm have yet re-emigrated.[15]

In Figure A.4 I show the unconditional relationship between time since first employment in Germany and the initial conational share in 2014, binning observations by year of first employment and plotting a quadratic trend in time since first employment in 2014. There is some evidence of the initial conational share decreasing and then plateauing with time since first employment, suggesting that α_S might be negative. While there is some evidence of the initial conational share increasing again for individuals who have been in Germany more than 15 years, this may be simply the result of observing fewer individuals who have been in Germany that long.

The coefficients of the quadratic trend, reported in column 1 of Table A.3 are not significant. Until now I have abstracted from selection into the initial conational share, S. In reality, the unconditional association between S and C plotted in Figure A.4 is the combination of (i) the direct causal effect of the conational share on return migration; (ii) the indirect effect of the conational share on return migration mediated by employment outcomes; and (iii) a noncausal association between the conational share and return migration due to the possible existence of common causes of the conational share and employment outcomes which will be transmitted via employment to return migration. To be able to infer the sign of α_S, one needs to control for the effect of employment rates on return migration, blocking both the indirect causal path and the

[15] If, on the other hand, the initial conational share has a trend over time, it will not be possible to identify both the cohort effects and the effect of the conational share on return migration using data only on stayers in a given year.

non-causal association. OLS estimates of the time trend controlling for the average individual employment rate between first employment year and 2014 are reported in column 2 of Table A.3. The time trend is unchanged and still statistically insignificant, although the magnitudes are still economically relevant: an individual in her first year of employment in 2014 had a coworker share on average 3 percentage points higher than an individual in her fifth year of employment who had not yet return migrated.

While the evidence presented here relies on the strong assumption that there are no cohort effects in the initial conational share and is estimated on a small sample, it nevertheless suggests points to $\alpha_S \leq 0$, with $\alpha_S = 0$ the most likely value. In this case, given $\alpha_Y > 0$ and $\beta < 0$, I conclude that $\alpha_S + \alpha_Y \beta < 0$. The bias induced by selective return migration will be positive and $\beta < \hat{\beta}$.

6 Discussion of possible mechanisms

Having established that the effect of the conational share on subsequent employment is robust to different possible sources of bias, in this section I explore what evidence there is for the different theories outlined in Section 2. In particular, I suggested there that the main mechanisms through which the initial conational share might worsen the longer-term employment rate of immigrants were by reducing their incentives to acquire host country-relevant human capital, particularly language skills, or by worsening the quality of their social network, thereby reducing the arrival rate of job offers.

I review evidence for different possible explanations for the estimated effect in Table 6 where I regress alternative outcomes on the conational share, conditional on the full set of controls and fixed effects defined in Equation (2). In column 1 I test for an effect of the initial conational share on longer-term Germany-specific human capital accumulation, measured as an indicator for being proficient in German, as self-reported in the SOEP. The sample is restricted to the years 2013 and 2014, when the SOEP was conducted, meaning my estimates are likely to be less precise. A-one-percentage point increase in the initial conational share is associated with a 0.3–0.5-percentage-point lower probability of being proficient in German in the first five years of the immigrant's time in Germany, however not in the longer term.[16] Working with more conationals does, therefore, appear to slow down an individual's learning German. However, the effect is not persistent, suggesting that it is perhaps unlikely to explain the longer term reduction in employment caused by the initial conational share. Corroborating this claim, the association of the initial conational share and of the other immigrant share with subsequent German proficiency are almost identical. This suggests that even if lowered German proficiency in the medium term were a mechanism by which the initial conational share lowered employment in the longer term, it could not be the only, or even the primary mechanism by which this happens. Otherwise, one would observe, contrary to the fact, that the other immigrant share has a similar negative

[16] Note that the set of controls includes a dummy for having been proficient in German *before* migrating, derived from a separate question in the SOEP survey questionnaire.

Table 6: Outcomes related to different theoretical explanations.

	(1) Proficiency	(2) Job separation	(3) Conat. share	(4) Other migrant share	(5) Job coworker	(6) Match exp.
$\mathbf{1}(t \in [0,2]) \times$ Conat. share	-0.50**	0.12*	0.71**	0.012	-0.023	-0.40*
	(0.16)	(0.048)	(0.044)	(0.024)	(0.091)	(0.18)
$\mathbf{1}(t \in [3,5]) \times$ Conat. share	-0.28	0.0048	0.37**	0.12*	-0.040	0.39
	(0.17)	(0.053)	(0.078)	(0.053)	(0.069)	(0.42)
$\mathbf{1}(t \geq 6) \times$ Conat. share	-0.10	0.014	0.16**	0.055	-0.063	-0.068
	(0.087)	(0.038)	(0.056)	(0.045)	(0.077)	(0.13)
$\mathbf{1}(t \in [0,2]) \times$ Other mig. share	-0.45*	0.090*	0.0033	0.75**	0.17	-0.13
	(0.18)	(0.042)	(0.025)	(0.027)	(0.12)	(0.42)
$\mathbf{1}(t \in [3,5]) \times$ Other mig. share	-0.23+	0.052	-0.032	0.41**	0.081	-0.089
	(0.13)	(0.051)	(0.024)	(0.047)	(0.073)	(0.16)
$\mathbf{1}(t \geq 6) \times$ Other mig. share	-0.075	0.065+	0.0076	0.36**	-0.042	-0.087
	(0.079)	(0.038)	(0.034)	(0.036)	(0.060)	(0.14)
Observations	1687	10061	7687	7687	530	626
Individuals	850	863	863	863	530	552
R^2	0.28	0.04	0.30	0.40	0.14	0.30
Source	SOEP	IAB	IAB	IAB	SOEP	SOEP

Note: OLS estimates of the relationship between the initial coworker share and various variables dran from the IAB and the SOEP. All specifications include a quadratic in age, controls for pre-migration characteristics, firm characteristics and method of job-finding, as well as the full set of fixed effects defined in the text. The last row reports the source of the dependant variable. Standard errors clustered by individual. + p<0.1, * p<0.05, ** p<0.01

association with subsequent employment.

In column 2, I show that the conational share is associated with the probability of experiencing a job separation (unconditional on being employed) in the first two years, but not in the longer-term. This increase in turnover in the short term, however, is also present for individuals working with a higher share of immigrants from other countries. The increased turnover therefore also cannot alone explain the negative effect of the conational share on employment in the longer-term.[17]

In the remaining columns of Table 6, I explore how the composition of the initial workplace is associated with various job and workplace characteristics in subsequent years, conditional on employment. In columns 3 and 4 I assess the persistence of the conational and other immigrant shares. Both shares are persistent across time since the first job, in spite of the fact that higher conational and other immigrant shares were found to be associated with higher turnover in the short run. This persistence suggests that, as in other settings, social networks play a role in job finding and that the coworkers in the initial job may be an important source of information about the labour market when searching for subsequent jobs.

In columns 5 and 6 I draw on information about the current job from the SOEP to further asses the role of networks in job-finding. The initial conational share is not associated with the probability that the current job was found through one's coworkers (only available in 2014), which may suggest the composition of the initial set of coworkers does not affect the intensity of the use of former coworkers in subsequent job-search.[18] The initial conational share is also not associated in the longer term with the probability of being in a job that matches one's experience. However, since these two variables are only available for individuals actually working in the years the question is asked in the SOEP, the samples used in these regressions are considerably smaller and potentially selected, making it difficult to draw firm conclusions.

If the negative employment effect of working with relatively more conationals is due to the lower quality of social networks this induces, one would ideally test this explanation directly using information on the quality of the conational coworkers.[19] However, such a test is not possible in the IAB-SOEP Migrant Sample, which only includes individual information on the respondents themselves. To indirectly test this explanation, I therefore repeat my main regression specification, but interact the conational share with local labour market by nationality-specific employment rates.[20] I report the estimated association between the conational share and sub-

[17]Note that this effect is estimated conditional on having found the first job through a contact so should not be contrasted with the positive effect of finding a job through a referral (see Dustmann et al., 2016, who argue that high-conational-share jobs are more likely to have been found through a referral). Instead it may be that, netting out the positive effect of a referral, individuals who start out in a firm where they had a contact, or that more readily hire immigrants in general, are on average worse-matched to the job, a hypothesis supported by the evidence in column 6, and are therefore more likely to leave.

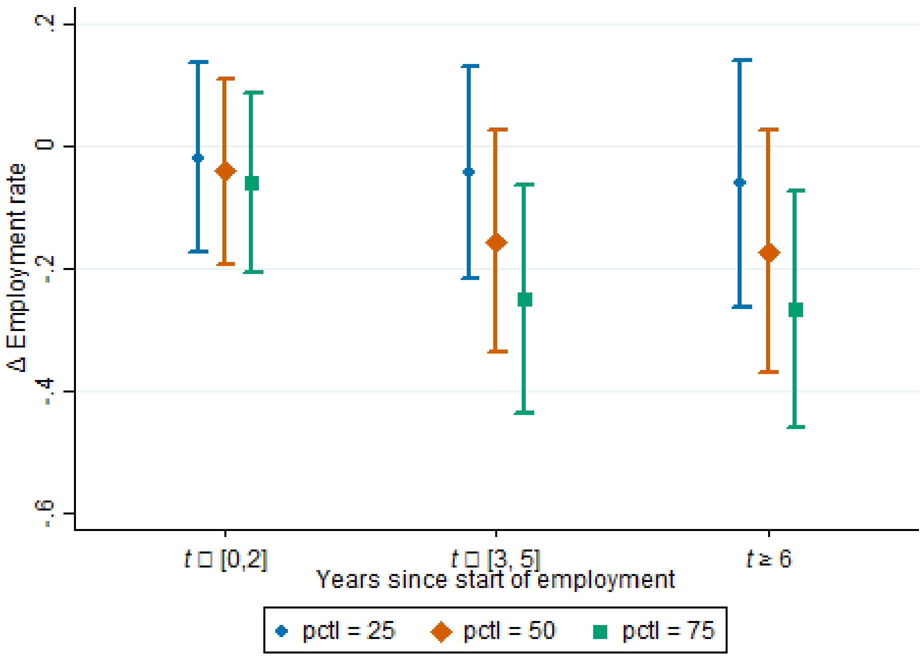
[18]Although here, as elsewhere, the outcome measures the use of referrals in the accepted offer, one does not observe the source of the full set of job offers (see Montgomery, 1991), nor how the unemployed are searching for work.

[19]For similar tests see for example Edin et al. (2003) in the case of ethnic residential networks, or Cingano and Rosolia (2012) or Glitz (2017) for coworker networks in general.

[20]For data confidentiality reasons, the area m employment rate for individuals of nationality c in year t is measured as a nationality by local labour market by year fixed effect in a regression of individual

sequent employment rates for individuals whose local conational employment rate is at the 25th, 50th, or 75th percentile in the sample in Figure 2. Clearly, the estimated effect is more negative when local employment rates are higher. One possible explanation for this potentially puzzling pattern is that those conationals who make it into employment are more positively selected when the local nationality-specific employment rate is lower. Working with relatively many conationals would in that case be less prejudicial, since these individuals are of on average higher quality. More generally, the pattern of heterogeneity present in Figure 2 appears more consistent with the underlying mechanism being differences in the quality of social networks, since it is not clear why the effect of the conational share on an immigrant's acquisition of host country human capital should vary with the local employment rate of the individual's conationals.

Figure 2: Interaction of conational share with local conational employment rates



Notes: Heterogeneous effect of initial conational share by local employment rate. Estimated from the full specification defined in Equation (2), interacting the conational share with a measure of local conational employment rates. 95 per cent confidence intervals are calculated using standard errors clustered by individual.

Finally, there are also interesting patterns of heterogeneity in the effects of the conational share by gender and education that may provide a measure of support for the claim that lower-quality social networks explain my findings. In columns 1 and 2 of Table A.4 I re-estimate my main specification separately for men and women. The effect is clearly strongest for women, for whom a one-percentage-point increase in the conational share lowers the long-term employment probability by 0.33 percentage points. Immigrant women are typically less attached to the labour

employment rates on a quadratic in age and the set of fixed effects, using a two per cent sample of all individuals making social security contributions. Only estimated fixed effects for cells containing at least twenty individuals are retained.

force than immigrant men, and have lower employment rates (see e.g. Sarvimäki, 2011). It seems reasonable that they would therefore be more likely to drop out of the labour force entirely if their job offer rate declines, or the distribution of offered wages deteriorates. This, and the fact that the conational share should affect the costs and benefits of learning German equally for men and women, provides some support for network-based explanations of the negative employment effect.

There is also an interesting pattern of heterogeneity by pre-migration education level, reported in columns 3–5. In particular, medium- and highly-educated immigrants, those with at least an apprenticeship qualification, are more susceptible to the negative effects of starting out with a low conational share. However, for highly educated individuals, the effect is also present for the other immigrant share. These patterns are consistent with previous evidence showing that highly educated immigrants benefit more from improvements in the quality of their ethnic network (Edin et al., 2003).

7 Conclusion

In this paper I have shown that starting one's career in an establishment with a high share of conationals has negative long-term effects on an immigrant's labour market outcomes and particularly their employment rate. This is in contrast to the literature on initial residential conditions for newly arrived immigrants, where a high share of conationals in an immigrant's location of residence, by expanding the size of an individual's network, is generally thought to have positive effects on an immigrant's labour market outcomes. The effect is also specific to an immigrant's conationals; there is no statistically significant penalty for working with immigrants from other countries of origin. Descriptive evidence suggests that the negative effect is not due to a reduced acquisition of Germany-specific human capital. Instead, I show indirect evidenct that working with more conationals worsens the quality of an immigrant's social network, making it harder to find jobs in future. However, observations on the labour-market outcomes of the individuals who make up the immigrant worker's social network would be necessary to directly test this hypothesis.

One common feature of the existing results on the effects of initial residential conditions is that they potentially suffer from selection bias due to differential selection into return migration based on the treatment of interest. This paper provides the first formal treatment, to my knowledge, of the sign of the bias this is likely to create for estimates of the effect of initial conditions. The results contained in this paper could be productively used in future research to empirically assess the sign of the different components of the bias in these settings. Such an exercise would require a dataset that can identify future return migrants and non-return migrants, something that is not possible with the present dataset.

References

Abadie, A., S. Athey, G. Imbens, and J. Wooldridge (2017): "When Should You Adjust Standard Errors for Clustering?" .

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Abramitzky, R., L. P. Boustan, and K. Eriksson (2014): "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 122, 467–506.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of Political Economy*, 113, 151–184.

Andersson, F., M. García-Pérez, J. Haltiwanger, K. McCue, and S. Sanders (2014): "Workplace Concentration of Immigrants," *Demography*, 51, 2281–2306.

Angrist, J. D. and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, N.J.: Princeton University Press.

Ansala, L., O. Åslund, and M. Sarvimäki (2021): "Immigration history, entry jobs and the labor market integration of immigrants," *Journal of Economic Geography*.

Arellano-Bover, J. (2020): "Career consequences of firm heterogeneity for young workers: First job and firm size," .

Åslund, O., L. Hensvik, and O. N. Skans (2014): "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market," *Journal of Labor Economics*, 32, 405–441.

Åslund, O. and O. N. Skans (2010): "Will I see you at work: Ethnic Workplace Segregation in Sweden, 1985-2002,," *ILR Review*, 63, 471–493.

Aydemir, A. and M. Skuterud (2008): "The immigrant wage differential within and across establishments," *Industrial and Labor Relations Review*, 61, 334–352.

Azlor, L., A. P. Damm, and M. L. Schultz-Nielsen (2020): "Local labour demand and immigrant employment," *Labour Economics*, 63.

Barth, E., B. Bratsberg, and O. Raaum (2012): "Immigrant wage profiles within and between establishments," *Labour Economics*, 19, 541–556.

Battisti, M., G. Peri, and A. Romiti (2018): "Dynamic effects of co-ethnic networks on immigrants' economic success," .

Beaman, L. A. (2012): "Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the U.S." *Review of Economic Studies*, 79, 128–161.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012): "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 80, 2369–2429.

Belloni, A., V. Chernozhukov, and C. Hansen (2013): "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 81, 608–650.

——— (2014): "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 28, 29–50.

Bevis, L., K. Kim, and D. Guerena (2020): "Soils and South Asian stunting: Soil zinc deficiency drives child stunting in Nepal," .

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020): "How much should we trust estimates of firm effects and worker sorting?" .

Borjas, G. J. (1985): "Assimilation, changes in cohort quality, and the earnings of immigrants," *Journal of Labor Economics*, 3, 463–489.

Boucher, V. and M. Goussé (2019): "Wage dynamics and peer referrals," *Review of Economic Dynamics*, 31, 1–23.

Brücker, H., M. Kroh, S. Bartsch, J. Goebel, S. Kühne, E. Liebau, P. Trübswetter, I. Tucci, and J. Schupp (2013): "The new IAB-SOEP Migration Sample: an introduction into methodolofy and the contents," .

Burdett, K. and D. T. Mortensen (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 39, 257–273.

Calvó-Armengol, A. and M. O. Jackson (2004): "The effects of social networks on employment and inequality," *American Economic Review*, 94, 426–454.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., J. Heining, and P. Kline (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Chernozhukov, V., C. Hansen, and M. Spindler (2015): "Post selection and post regularization inference in linear models with many controls and instruments," *American Economic Review: Papers & Proceedings*, 105, 486–490.

Chiswick, B. R. (1978): "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, 86, 897–921.

Cingano, F. and A. Rosolia (2012): "People I Know: Job Search and Social Networks," *Journal of Labor Economics*, 30, 291–332.

Damm, A. (2009): "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence," *Journal of Labor Economics*, 27, 281–314.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016): "Referral-based job search networks," *Review of Economic Studies*, 83, 514–546.

Eckstein, Z. and Y. Weiss (2010): "On the Wage Growth of Immigrants: Israel 1990-2000," *Journal of the European Economic Association*, 2, 665–695.

Edin, P.-A., P. Fredriksson, and O. Aslund (2003): "Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment," *Quarterly Journal of Economics*, 118, 329–357.

Eliason, M., L. Hensvik, F. Kramarz, and O. Nordstrom Skans (2019): "Social Connections and the Sorting of Workers to Firms," .

Galenianos, M. (2013): "Learning about match quality and the use of referrals," *Review of Economic Dynamics*, 16, 668–690.

Gee, L. K., J. Jones, and M. Burke (2017a): "Social networks and labor markets: How strong ties relate to job finding on facebook's social network," *Journal of Labor Economics*, 35, 485–518.

Gee, L. K., J. J. Jones, C. J. Fariss, M. Burke, and J. H. Fowler (2017b): "The paradox of weak ties in 55 countries," *Journal of Economic Behavior and Organization*, 133, 362–372.

Glitz, A. (2014): "Ethnic segregation in Germany," *Labour Economics*, 29, 28–40.

——— (2017): "Coworker networks in the labour market," *Labour Economics*, 44, 218–230.

Glitz, A. and R. Vejlin (2020): "Learning through coworker referrals," *Review of Economic Dynamics*.

Glover, D., A. Pallais, and W. Pariente (2017): "Discrimination as a self-fulfilling prophecy: evidence from french grocery stores," *Quarterly Journal of Economics*, 132, 1219–1260.

Granovetter, M. (1995): *Getting a job: a study of contacts and careers*, University of Chicago Press, 2 ed.

Hellerstein, J. K. and D. Neumark (2008): "Workplace Segregation in the United States: Race, ethnicity, and skill," *The Review of Economics and Statistics*, 90, 459–477.

HJORT, J. (2014): "Ethnic divisions and production in firms," *Quarterly Journal of Economics*, 129, 1899–1946.

KAHANE, L., N. LONGLEY, AND R. SIMMONS (2013): "The effects of coworker heterogeneity on firm-level output: Assessing the impacts of cultural and language diversity in the national hockey league," *Review of Economics and Statistics*, 95, 302–314.

LAZEAR, E. P. (1999): "Globalisation and the Market for Team-Mates Author(s): Edward P. Lazear Source: The Economic Journal, Vol. 109, No. 454, Conference Papers (Mar., 1999), pp. C15-C40 Published by:," *The Economic Journal*, 109, 15–40.

LUBOTSKY, D. H. (2007): "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 115, 820–867.

MANNING, A. (2011): *Imperfect competition in the labor market*, vol. 4, Elsevier B.V.

MCPHERSON, M., L. SMITH-LOVIN, AND J. M. COOK (2001): "Birds of a feather : Homophily in social networks," *Annual Review of Sociology*, 27, 415–444.

MONTGOMERY, J. D. (1991): "Social networks and labor-market outcomes: Toward an economic analysis," *American Economic Review*, 81, 1408–1418.

——— (1992): "Job search and network composition: Implications of the strength-of-weak-ties hypothesis," *American Sociological Review*, 57, 586–596.

MUNSHI, K. (2003): "Networks in the modern economy: Mexican migrants in the U. S. labor market," *Quarterly Journal of Economics*, 118, 549–599.

OECD (2020): "International Migration Statistics," .

OSTER, E. (2019): "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business and Economic Statistics*, 37, 187–204.

OTTAVIANO, G. I. AND G. PERI (2012): "Rethinking the effect of immigration on wages," *Journal of the European Economic Association*, 10, 152–197.

PERI, G. AND C. SPARBER (2009): "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, 1, 135–169.

ROSEN, S. (1986): "The theory of equalizing differences," *Handbook of Labor Economics*, 1, 641–692.

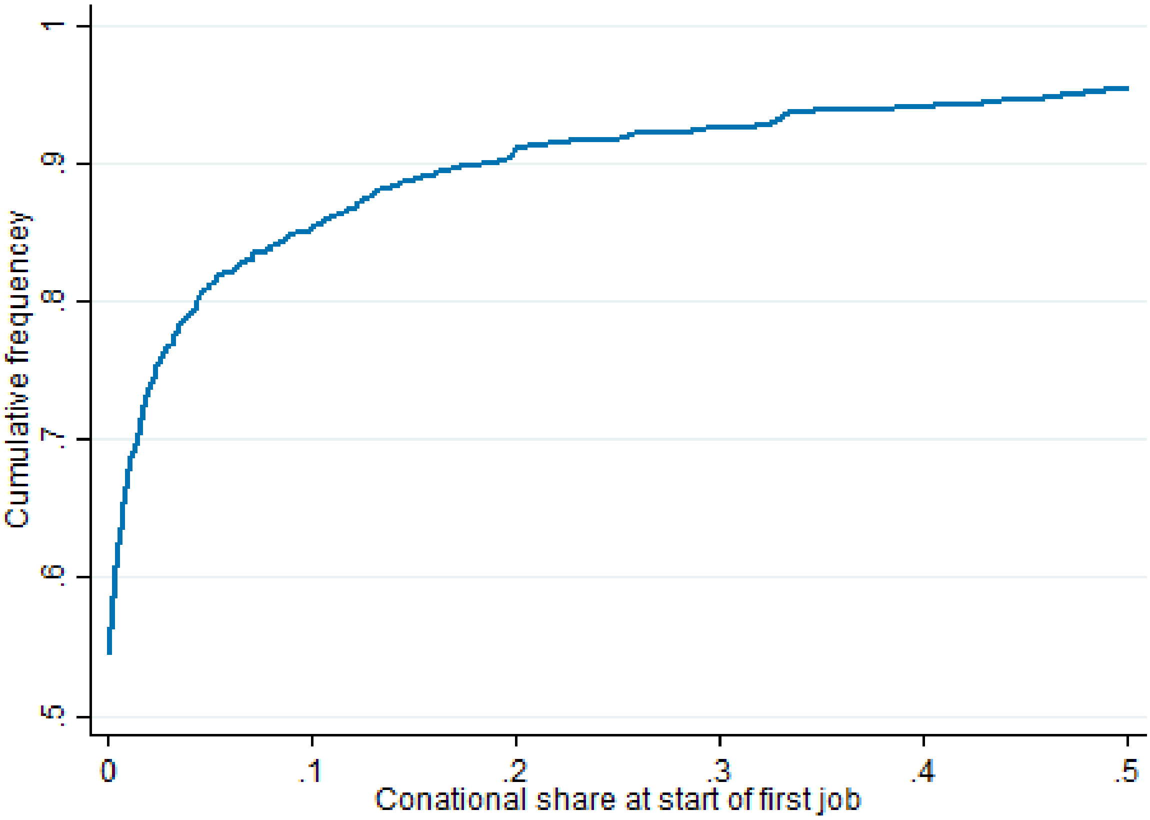
SARVIMÄKI, M. (2011): "Assimilation to a welfare state: Labor market performance and use of social benefits by immigrants to Finland," *Scandinavian Journal of Economics*, 113, 665–688.

SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2019): "Firming up inequality," *Quarterly Journal of Economics*, 134, 1–50.

Sorkin, I. (2018): “Ranking firms using revealed preference,” *Quarterly Journal of Economics*, 133, 1331–1393.

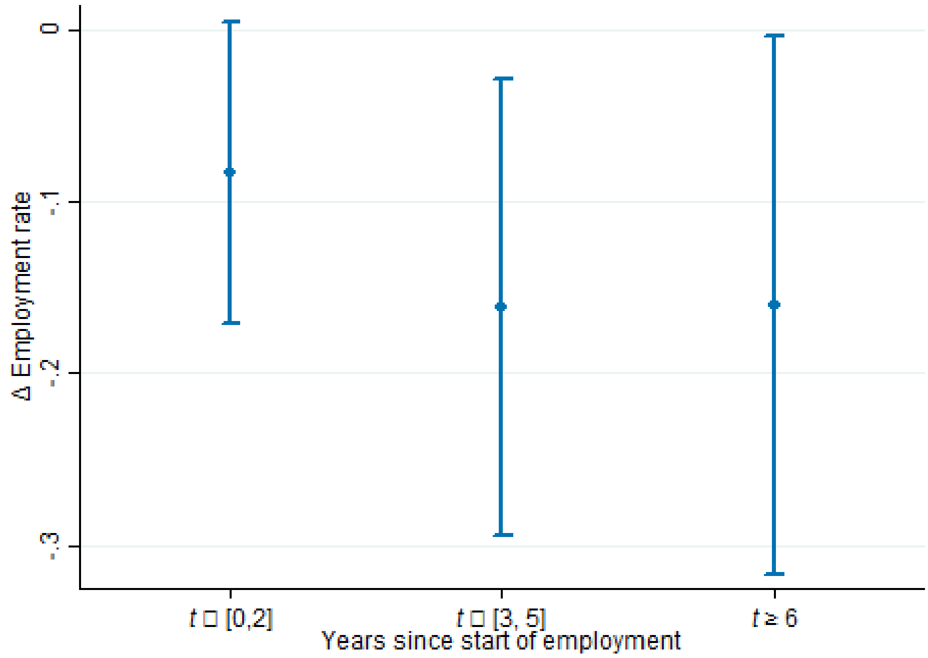
A Supplementary figures and tables

Figure A.1: CDF of conational share in first job



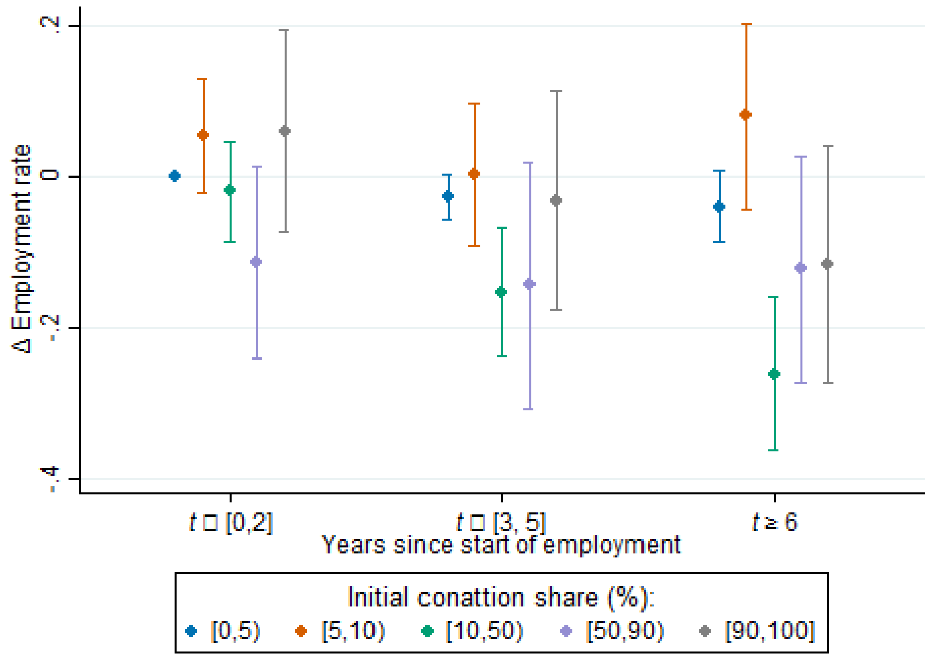
Notes: Empirical CDF of the initial conational share in the first job held by an immigrant in my sample. The distribution is truncated at 50, for ease of representation.

Figure A.2: Post-regularised estimate of employment effect



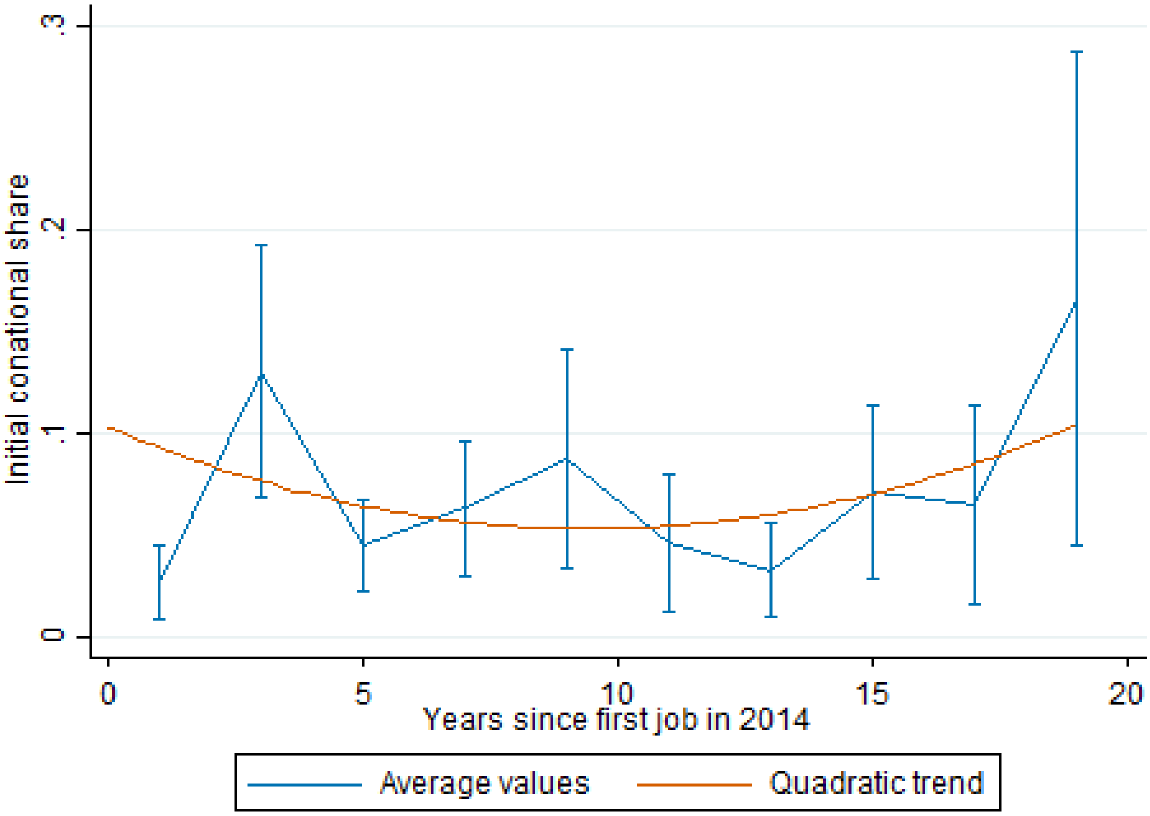
Notes: Post-regularisation estimates of the effect of the initial conational share on subsequent employment rates. See Appendix B for details on the procedure. 95 per cent confidence intervals are calculated using standard errors clustered by individual.

Figure A.3: Non-linear employment effect of composition of coworkers



Notes: Indicators for each category, coworker share in $[0, 5)$ in the first two years of employment is the omitted category. The full set of controls and fixed effects is included, 95 per cent confidence intervals are calculated using standard errors clustered by individual.

Figure A.4: Starting conational share over time



Notes: Differences in initial conational share for immigrants who first worked in Germany t years ago in 2014. Average values are grouped in two-year bins for data protection reasons.

Table A.1: Country groups

	N	Share
Russia	323	37.43
Romania	114	13.21
Poland	93	10.78
ex-Yugoslavia	71	8.23
Turkey	65	7.53
Asia	52	6.03
Italy	41	4.75
Other Europe	38	4.40
Africa	29	3.36
Greece	2*	2.55
Others	//	////
Total	863	100.00

Note: Refers to country of birth for individuals born without German nationality. The table has been censored in accordance with IAB data protection requirements.

Table A.2: Relation between initial coworker share and other labour market outcomes

	(1) Benefit receipt	(2) Jobseeker	(3) In training	(4) Not in IEB	(5) Self-employed	(6) Civil servant
$\mathbf{1}(t \in [0,2]) \times$ Conat. share	-0.021	-0.030	-0.043**	-0.050	0.15	-0.0030
	(0.054)	(0.058)	(0.012)	(0.033)	(0.16)	(0.013)
$\mathbf{1}(t \in [3,5]) \times$ Conat. share	-0.026	-0.026	-0.043**	0.028	0.38*	-0.030
	(0.059)	(0.060)	(0.015)	(0.046)	(0.18)	(0.019)
$\mathbf{1}(t \geq 6) \times$ Conat. share	0.047	0.055	-0.025+	0.11	0.076	0.0070
	(0.068)	(0.072)	(0.013)	(0.066)	(0.061)	(0.0077)
$\mathbf{1}(t \in [0,2]) \times$ Other mig. share	-0.036	-0.10*	-0.016	0.0036	0.021	0.014
	(0.048)	(0.049)	(0.021)	(0.024)	(0.035)	(0.014)
$\mathbf{1}(t \in [3,5]) \times$ Other mig. share	-0.037	-0.027	-0.017	0.013	-0.055	-0.012
	(0.054)	(0.055)	(0.021)	(0.032)	(0.041)	(0.013)
$\mathbf{1}(t \geq 6) \times$ Other mig. share	0.12+	0.13+	0.0022	-0.040	-0.0053	-0.0074
	(0.063)	(0.065)	(0.016)	(0.030)	(0.032)	(0.0060)
Observations	10061	10061	10061	10061	1506	1506
Individuals	863	863	863	863	849	849
R^2	0.15	0.15	0.06	0.08	0.14	0.08
Source	IAB	IAB	IAB	IAB	SOEP	SOEP

Note: Benefit receipt is a dummy for receiving earnings replacement benefits or unemployment benefits, defined respectively under Social Code Book (SGB) III and SGB II, in the course of the year. Jobseeker is a dummy for being registered as a job seeker with an employment agency. In training is a dummy for participating in a federal or state active labour market policy measure. Not in IEB is a dummy for no social security data being available in a given year. Self-employed and Civil servant are self-reported dummy variables from the SOEP survey, available in 2013-14. All specifications follow Equation (2), standard errors clustered by individual. + p<0.1, * p<0.05, ** p<0.01.

Table A.3: Relationship between initial conational share and return migration

	(1)	(2)
	Conat. share	Conat. share
t	-0.010	-0.0082
	(0.0064)	(0.0068)
$t \times t$	0.00056	0.00047
	(0.00035)	(0.00036)
Average Employment rate		-0.053
		(0.033)
Constant	0.10**	0.13**
	(0.026)	(0.028)
Observations	791	791
R^2	0.004	0.008

Note: Evidence of a relationship between the conational share and time since first job, via the relationship between time since first job, t, and the conational share. Robust standard errors reported. + p<0.1, * p<0.05, ** p<0.01

Table A.4: Heterogeneity of employment effect

	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}(t \in [0,2]) \times$ Conat. share	-0.051	0.069	-0.046	0.00063	0.076
	(0.097)	(0.070)	(0.085)	(0.12)	(0.14)
$\mathbf{1}(t \in [3,5]) \times$ Conat. share	-0.16	-0.011	-0.014	-0.14	-0.037
	(0.10)	(0.082)	(0.085)	(0.13)	(0.17)
$\mathbf{1}(t \geq 6) \times$ Conat. share	-0.33**	-0.048	-0.0013	-0.24+	-0.30+
	(0.12)	(0.075)	(0.077)	(0.14)	(0.17)
$\mathbf{1}(t \in [0,2]) \times$ Other mig. share	-0.12+	0.066	-0.079	-0.047	-0.072
	(0.069)	(0.064)	(0.083)	(0.081)	(0.082)
$\mathbf{1}(t \in [3,5]) \times$ Other mig. share	-0.13	0.019	-0.050	0.024	-0.25*
	(0.088)	(0.080)	(0.098)	(0.095)	(0.11)
$\mathbf{1}(t \geq 6) \times$ Other mig. share	-0.15	0.0099	-0.057	0.0078	-0.32*
	(0.11)	(0.080)	(0.10)	(0.090)	(0.13)
Observations	4667	5394	4404	3194	2463
Individuals	432	431	344	273	246
R^2	0.16	0.19	0.18	0.22	0.26
Sample	Women	Men	Low	Med	High

Note: OLS estimates of the relationship between initial conational share and subsequent individual employment rates. The individual employment rate is the fraction of days in a year an individual is employed. Columns 1 and 2 report results conditional on gender, columns 3–5 conditional on the pre-migration educational attainment being either lower than apprenticeship, an apprenticeship, or higher than apprenticeship. Standard errors clustered by individual. + p<0.1, * p<0.05, ** p<0.01

B Post-regularisation implementation details

In Equation (2), I impose restrictive assumptions on the form of the regression function to make it tractable. To check that my estimates are robust to more flexible functional forms without overfitting my relatively small sample, I would like to allow for a wide set of interactions between my control variables and only retain ones that are truly relevant. Traditional dimension-reduction methods of penalised estimation, such as the Least Absolute Shrinkage and Selection Operator (LASSO) treat all regressors as equal, and may not retain my regressor of interest, S in the set of included predictors of the outcome Y. Furthermore, the LASSO and related methods are not intended to estimate the marginal effect of any one variable on the outcome Y, so even if S is retained as a regressor by the LASSO, it is incorrect to interpret the estimated coefficient on S as an estimate of the true marginal effect of S on Y.

For this reason, methods for applying the LASSO to causal and structural models and conducting inference on a set of linear parameters of interest break the set of predictors of Y into two groups: one low-dimensional group of regressors of interest (in this case S and its interactions with years since first employment, though here I focus on S for expositional ease) and one high-dimensional set of nuisance regressors, whose inclusion is necessary to guarantee that the structural model is correctly specified, X. Elements of X are then chosen by regressing Y and S one-by-one on the set X using the LASSO. The marginal effect of S on Y can then be estimated by calculating the residual of the LASSO regression of Y on X and regressing this on the residual of the LASSO regression of S on X, an approach known as post-regularisation (Belloni et al., 2013; Chernozhukov et al., 2015).[21]

I consider the following set of control variables X: (i) orthogonalised fifth-degree polynomials in age, pre-migration experience, age at migration, log wages in first job, and log firm size, firm log median wages, and firm age, all in the first job; (ii) dummy variables for each nationality group, year of migration, federal state in which first located, and education group, as well as dummy variables for being employed and for being proficient in German pre-migration, for having a first job that was part-time or an apprenticeship, for gender, for having support from contacts in Germany when moving, and for finding the first job through contacts; (iii) all one-way interactions for the complete set of dummy variables; (iv) all one-way interactions between the dummy variables and the terms of the fifth-degree polynomials; and (v) dummy variables for years since migration and their interactions with the initial other immigrant share.[22] In total, this makes for 1220 control variables in my high-dimensional nuisance regressor set.

By design, if two regressors are highly correlated, the LASSO will usually only retain one of them, which cannot be interpreted to mean that only the retained variable matters for the

[21]Note the conceptual similarity of this approach to the Frisch-Waugh-Lovell theorem, where one regresses the residual of a regression of Y on a low-dimensional X on the residual from regressing T on X.

[22]This implies that I retain the assumption that the effect of the included covariates is constant over time, with the potential exception of the other immigrant share. In results available on request, I check that my results are robust to including interactions of all dummies and polynomial terms with the years since migration dummies. The estimated effect of the initial conational share remains negative and significant, however interpreting the larger set of retained covariates is less straightforward.

outcome. Nevertheless, it can be instructive to consider the set of retained variables as a check of the researcher's priors. For both the employment rate and earnings conditional on being employed, the LASSO retains the set of time since first job dummies, the interaction of the share of other immigrants with $\mathbf{1}(t \in [3, 5])$, a dummy for Romanian nationality, the interaction of the Romanian dummy with dummies for part-time, first job through contacts, gender, and the full set of education dummies, and linear terms in age of first establishment and log first establishment size. The employment specification also includes in particular a dummy for having contacts in Germany at migration and pre-migration German proficiency, while the earnings specifications include, in particular, linear terms for the log starting wage and log median wage in the first firm and the interaction of quadratic terms for the same variables with a dummy for being high-educated.

It is instructive and perhaps reassuring to consider that measures of employability, such as pre-migration German proficiency and having a pre-existing network of contacts in Germany matter (positively) for subsequent employment, but not wages conditional on employment, while measures of the quality of the first job, in particular starting wage and median firm wage, are important predictors of subsequent wages, but not of subsequent employment. Conversely, the differential effect of several factors for Romanians, the second-largest group in my sample, was not necessarily expected *a priori*.

C Proof of the sign of the bias induced by selective return migration

I claim that if $\alpha_Y > 0$, the bias will be of the opposite sign to $\alpha_S + \alpha_Y \beta$, while if $\alpha_Y < 0$, the bias will be of the same sign as $\alpha_S = \alpha_Y \beta$. To see this, note that if $\alpha_Y > 0$, the condition $C^* \geq K$ truncates the left tail of the distribution of ε_Y; the expectations in Equation (9) will be positive. Furthermore, if $\alpha_S + \alpha_Y \beta > 0$, then the supplementary condition $S = 1$ truncates the distribution less severely than when the condition is not imposed, since $S \in \{0, 1\}$. As a result, we will have

$$\mathrm{E}[\varepsilon_Y | C^* \geq K, S = 1] < \mathrm{E}[\varepsilon_Y | C^* \geq K] \tag{11}$$

and the bias will be negative. If, on the other hand, $\alpha_Y < 0$, the right tail of the distribution is truncated and the expectations in Equation (9) are negative. If $\alpha_S + \alpha_Y \beta > 0$, the supplementary condition $S = 1$ again means the distribution is less severely truncated, implying now that

$$\mathrm{E}[\varepsilon_Y | C^* \geq K, S = 1] > \mathrm{E}[\varepsilon_Y | C^* \geq K] \tag{12}$$

and the bias will be positive.