

ESSAYS ON COWORKER NETWORKS
AND
EMPLOYMENT DYNAMICS

Omar Hussein

A thesis submitted for the degree of
Doctor of Philosophy in Economics

Department of Economics
University of Essex

March 2026

Declaration

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

This research was funded by the South and East Network for Social Sciences (SeNSS), an ESRC-funded Doctoral Training Partnership (project N. 2762505).

Omar Hussein

Acknowledgments

So many wonderful people have supported me throughout this journey, and I am very grateful to them. I would like to thank my supervisors Carlos Carrillo-Tudela and Xiaoyu Xia for their advice and guidance. I thank Tarek Selim and Emma Duchini for their thoughtful advice. I am grateful to Paul Fisher for his mentorship and for our collaborations — from him, I have learned a lot.

I have made so many amazing friends along the way. I'm especially grateful to Tommaso, Hettie, and Aitor for the many great conversations we've shared. I also thank Pablo and Edith, my office mates, for their friendship and humor.

I would like to thank Sarah, Cherif F., and Cherif C., for being my friends for more years than I can count, and for our unforgettable sing-along car rides.

I would like to thank Annalivia, my partner, for her encouragement and love. She stood by me through the ups and downs of this journey, and I will always be grateful to her.

Finally, this wouldn't have been possible without my family. My grandmothers, Alia and Hanaa, have inspired me in countless ways. My brother, Hazem, is one of the smartest and kindest people I know. My nephew, Younis — well, just being Younis is more than enough. I am deeply thankful to my mom, Hala, and my dad, Mohamed, for everything they have done for me — from helping me study and driving me to training, to supporting me when I needed it most, and, above all, for encouraging me to pursue my dreams.

Summary

This thesis consists of three distinct but related papers on the implications of information transmission between coworkers on labour market outcomes and employment dynamics. I use administrative data from Germany (SIEED) as the main source of data in all three papers.

The first paper empirically investigates the relationship between coworker occupational homophily — defined as the share of direct coworkers with the same occupation — and the probability of changing occupation among workers who switch jobs. The results suggest a strong negative effect of occupational homophily on the probability of changing occupation. This effect is more pronounced for high-skilled service and production jobs, and when coworkers share other characteristics such as gender and age group.

The second paper examines how the size of coworker networks influences job search efficiency and its implications for worker mobility and wage dispersion. I develop an equilibrium on-the-job search model with two-sided heterogeneity, wage bargaining, and endogenous coworker networks, where workers form connections with coworkers when they join a firm and retain them during periods of unemployment. I estimate the model using indirect inference, and I find a large effect of networks on job-to-job mobility, but a relatively modest effect on wage dispersion.

The third paper investigates the role of the occupational composition of coworker networks in occupational choice and its implications on mobility, wages, and occupational allocation. I propose a model that incorporates search through coworker networks into an on-the-job search model with heterogeneous workers, firms, and oc-

cupations. The data suggest that the probability of choosing an occupation following a transition increases with the share of coworkers switching to that occupation. I estimate the model using indirect inference. A counterfactual analysis indicates that coworker networks account for a moderate part in wage dispersion, and a significant part in mobility and the allocation of workers into high skilled occupations.

Contents

Introduction	i
1 Occupational Homophily in Coworker Networks and Occupational Change	1
1.1 Introduction	1
1.2 Conceptual framework	6
1.2.1 Occupational homophily	6
1.2.2 Theoretical background	7
1.3 Empirical framework	10
1.3.1 Regression framework	10
1.3.2 Identification	13
1.4 Data and descriptives	16
1.4.1 Data	16
1.4.2 Descriptive statistics	18
1.5 Empirical results	22
1.5.1 The effect of occupational homophily on occupational change . .	22
1.5.2 The role of worker and coworker characteristics	24
1.6 Conclusion	31
2 Equilibrium Search with Endogenous Coworker Networks	34
2.1 Introduction	34
2.2 Model	37
2.2.1 Environment	37
2.2.2 Worker problem	41
2.2.3 Steady state	45

CONTENTS

2.3	Quantitative analysis	50
2.3.1	Data	50
2.3.2	Parameterisation and identification	51
2.3.3	Estimation	55
2.3.4	Results	55
2.3.5	Implications on mobility and wage dispersion	59
2.4	Discussion and extensions	60
2.4.1	Social capital accumulation	60
2.4.2	Social ability and sorting	63
2.5	Conclusion	65
3	Coworker Networks and Occupational Allocation	67
3.1	Introduction	67
3.2	Model	71
3.2.1	Environment	71
3.2.2	Job search and coworker networks	74
3.2.3	Steady state	80
3.3	Quantitative analysis	82
3.3.1	Data	83
3.3.2	Parameterisation and identification	84
3.3.3	Estimation	91
3.3.4	Results	92
3.3.5	Implications on wages, mobility, and occupational allocation	94
3.4	Conclusion	96
	Bibliography	103
	Appendix A: Chapter 1	104
A.1	Additional descriptive statistics	104
A.2	Occupational definitions	110
A.3	Variable construction	112

CONTENTS

A.3.1	Mapping coworker networks	112
A.3.2	Wage imputation	113
A.4	Additional results	114
Appendix B: Chapter 2		118
B.1	Mathematical appendix	118
B.1.1	Monotonicity of the value of employment in productivity	118
B.1.2	Derivation of the value of employment	119
B.2	Estimation procedure	120
B.3	Additional tables and figures	122
Appendix C: Chapter 3		123
C.1	Sequential auctions framework	123
C.2	Additional tables	125

List of Figures

1.1	Occupational homophily and occupational change in Germany	2
1.2	Framework; occupational homophily and occupational change	10
1.3	An example of the formation of a coworker network	14
1.4	Degree and homophily distributions	20
2.1	Share of high ability workers in employment by firm rank and firm size	64
3.1	The productivity CDF of a job v and the sampling CDF of v conditional on working in y	78
3.2	Worker productivity by firm productivity rank and occupation	86
A.1	Share of job transitions resulting in occupational change, by year	106
A.2	Duration of non-employment and occupational change	107
A.3	Distribution of change in log wages	108
A.4	Degree and homophily distributions	109
B.1	Log wage and log output distribution	122

List of Tables

1.1	Distribution of Blossfeld occupations	18
1.2	Summary statistics; worker sample	19
1.3	Occupational homophily and occupational change; main results	23
1.4	Occupational homophily and Blossfeld occupational change; hetero- geneity by field and occupation of origin	25
1.5	Occupational homophily and Blossfeld occupational change; hetero- geneity by position in the wage distribution within Blossfeld occupations	27
1.6	Occupational homophily and occupational change; heterogeneity along other dimensions of homophily	29
1.7	Occupational homophily and Blossfeld occupational change; hetero- geneity by gender and age	30
2.1	Summary statistics by education group; SIEED 2012-2015	51
2.2	Parameters and moments	54
2.3	Model fit	56
2.4	Parameter values	57
2.5	Variance decomposition	58
2.6	Counterfactual exercise; no search through coworkers when employed .	59
3.1	Summary statistics	83
3.2	Estimates of ability and task productivity	85
3.3	Coworker networks and occupational choice	90
3.4	Estimation results	92
3.5	Counterfactual exercise; dedicate search effort to general channels	94
A.1	Summary statistics; worker, coworker, IV samples	104

LIST OF TABLES

A.2	Summary statistics; year, location and industry	105
A.3	KldB–Blossfeld correspondence; administration occupations	110
A.4	KldB–Blossfeld correspondence; service occupations	110
A.5	KldB–Blossfeld correspondence; production occupations	111
A.6	Occupational homophily and occupational change; heterogeneity by year	114
A.7	Occupational homophily and occupational change; heterogeneity by network size, duration of non-employment, and establishment size change	115
A.8	Occupational homophily and Blossfeld occupational change; hetero- geneity by skill and education	116
A.9	Occupational homophily and occupational change; alternative specifica- tions to occupational homophily	117
C.1	Wage regression results	125
C.2	Counterfactual exercise; dedicate search effort to general channels, by level of education	126

Introduction

Reliance on social networks is common in labour markets. Workers often search among their connections for job opportunities, and firms ask employees to refer potential candidates to fill vacancies (Topa, 2019). Coworker networks are an important part of these networks. These networks are composed of connections formed during the workers' employment spells. The connections with coworkers are weak ties¹ (Granovetter, 1973) – characterised by infrequent interactions. However, they can be useful because they are numerous and because coworkers are in a good position to observe their colleagues' skills and career paths. The literature has shown that coworker interactions can improve productivity through learning (Jarosch et al., 2021; Herkenhoff et al., 2024) and peer effects (Cornelissen et al., 2017; Hong and Lattanzio, 2022), and that coworkers can help their past colleagues find jobs once they go into unemployment (Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021; Gyetvai and Zhu, 2025).

The literature on coworker effects in job finding has primarily focused on unemployed workers displaced by mass layoffs (Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021; Gyetvai and Zhu, 2025).² However, the literature looking at the effects of coworkers on job finding for the employed (and subsequently job-to-job transitions) is rare.³ In addition, the literature generally does not look at the type of job – for example occupation – the worker finds through coworkers and how it differs from jobs found through other channels.

¹ In contrast to strong ties, which are family and friends; characterised by frequent interactions.

² This strand of research typically relies on administrative data, defining coworkers as individuals employed in the same establishment or firm at the same time.

³ A notable exception is Caldwell and Harmon (2019) who study the effect of coworkers on job finding, therefore their ability to renegotiate job contracts while employed.

This thesis aims to build on this literature by examining how coworker networks shape job and occupational mobility for the employed and the unemployed, and their implications on employment dynamics and the wider labour market. Specifically, I address the following questions; (1) how do the characteristics of coworker networks — such as their size and composition — relate to mobility? and (2) what are the implications for wage dispersion, mobility, and occupational allocation?

This thesis consists of three distinct but related papers addressing these questions. The common theme of the thesis is the focus on the nature of coworker networks as a social capital that can be leveraged to search for jobs during unemployment and on the job, which is accumulated and dis-accumulated, and as an endogenous object generated by the movement of workers between firms. A feature that is taken into account in the empirical exercises in Chapter 1 and incorporated in the structural models of Chapters 2 and 3.

The main data source for the three chapters is the SIEED, an establishment-level linked employer–employee dataset from Germany. It contains complete employment histories for all workers in a large sample of establishments, which makes it possible to identify coworker networks.

The papers are briefly summarised below.

The first paper empirically investigates the relationship between coworker occupational homophily — defined as the share of direct coworkers with the same occupation — and the probability of changing occupation among workers who switch jobs. I use an instrumental variable strategy to address the potential endogeneity arising from the simultaneity of a worker’s decision to form networks and to change occupation. Specifically, I exploit the variation in the network composition resulting from the firm hiring workers *after* the worker of interest to identify the causal effect of homophily.

The results suggest a strong negative effect of occupational homophily on the probability of changing occupation. This effect is more pronounced for high-skilled service and production jobs, when coworkers rank higher in the occupation specific wage distribution, and when coworkers share other characteristics such as gender and age

group.

The second paper examines how the size of coworker networks influences job search efficiency and the consequent implications for worker mobility and wage dispersion. I develop an equilibrium on-the-job search model with endogenous coworker networks, two-sided heterogeneity, and wage bargaining. Workers form connections with coworkers when they join a firm and retain them during periods of unemployment. Search efficiency is a function of the measure (i.e., size) of these connections. I assume that workers lose their connections upon moving to a new employer, which allows the measure of connections to be summarised by the size of their current employer.

In equilibrium, more productive firms are larger, and workers in larger firms have a higher job finding rate. The presence of network effects alters the employment structure of the economy, and thus mobility and wage dispersion. The underlying mechanism is that firms with mid-level productivity can attract workers, but as they grow, it becomes easier for their employees to move up the job ladder, which in turn reduces their size. Networks channel workers into firms at the high end of the productivity distribution. I estimate the model using indirect inference with German administrative data, and I find a large effect of networks on job-to-job mobility, but a relatively modest effect on wage dispersion.

The third paper investigates the role of the occupational composition of coworker networks in occupational choice and its implications on mobility and wages. I propose a model that incorporates search through coworker networks into an on-the-job search model with heterogeneous workers, firms, and occupations. Workers differ in ability and occupations in the level of skill requirements. High ability workers have a comparative advantage in high skilled tasks compared to low-ability workers. I keep the process of link formation and loss from the previous paper, but I follow a different approach to the definition of a coworker link – *Labour flow networks*. The network is composed of links between firms that are generated by the movement of workers between them. The worker assumes these links when they start a job can use them to

search for new jobs. The sampling probability of a job depends on the ratio between the measure of links to that job and the worker's total measure of links.

In equilibrium, it is easier for workers higher up the job ladder to find employment in highly productive occupations, as the share of links to such occupations increases higher up the job ladder. This is more pronounced for high ability workers. Similarly to the second paper, the model suggests that networks channel workers into highly productive jobs (occupation-firm pairs). One advantage of the model is that it allows the characterisation of the relationship between the occupational composition of the network and occupational choice. The data suggest that the probability of choosing an occupation following a transition increases with the share of coworkers switching to that occupation. I estimate the model using indirect inference. A counterfactual analysis indicates that coworker networks account for a moderate part in wage dispersion, and a significant part in mobility and the allocation of workers into high skilled occupations.

Chapter 1

Occupational Homophily in Coworker Networks and Occupational Change

1.1 Introduction

Coworker networks are an important part of a worker's social capital. Individuals form these connections over their employment spells, and rely on them to find new jobs, exit future unemployment, or climb the job ladder (Hellerstein and Neumark, 2020).¹ Much of the literature on the subject focuses on the effects of the characteristics of coworker networks on wages, tenure, and unemployment (Cingano and Rosolia, 2012; Cornelissen et al., 2017; Glitz, 2017; Glitz and Vejlin, 2021; Jarosch et al., 2021; Saygin et al., 2021). However, little is known on how they affect occupational change. Occupational change is an important feature in the labour market with multiple welfare implications. It can improve the match between the worker skill and the occupation (Groes et al., 2015), or lead to the loss of occupation specific human capital (Gathmann and Schönberg, 2010), and results in large earnings variations (Fitzenberger et al., 2015; Bachmann et al., 2020; Carrillo-Tudela et al., 2022).

The key hypothesis of this paper is that workers who have formed many connections with coworkers in the same occupation will have better access to that occupation, and are thus less likely to change occupation. A preliminary exploration of the rela-

¹ Broadly, reliance on social networks is ubiquitous in the labour market; Topa (2011) suggested that half of the jobs are filled through informal contacts, and Topa (2019) estimated that between 25% and 80% of vacancies are filled through contacts.

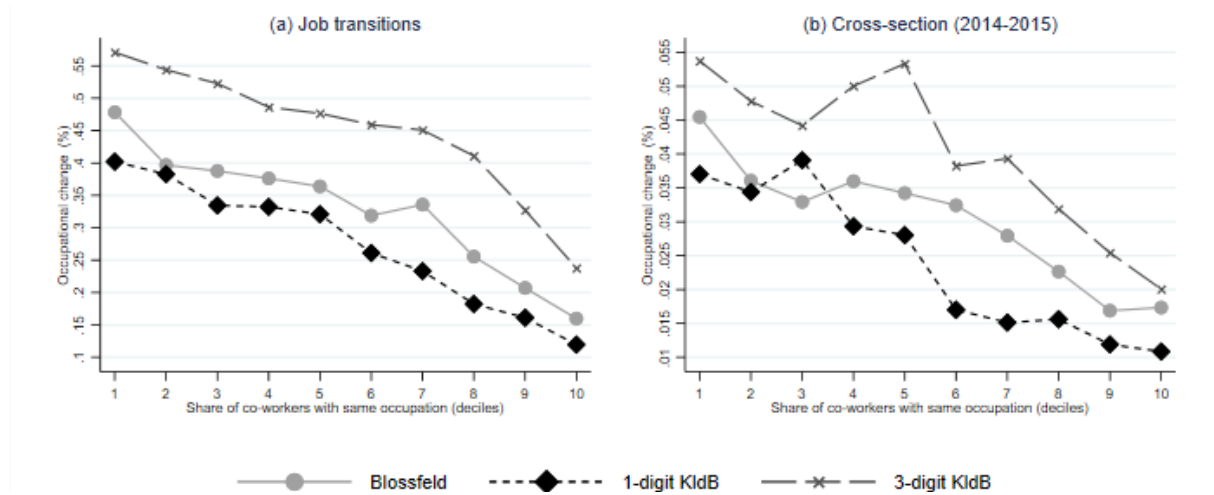


Figure 1.1: Occupational homophily and occupational change in Germany

Notes: Author's calculations using SIEED. Panel (a) establishment changes 2012-2017. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. Panel (b): occupation observed on January 1st 2014 compared to January 1st 2015. Sample of workers age 24-55 who worked only in panel establishments in the last three years. In both panels, coworkers are defined as those who worked in the same establishment at the same time for at least 30 days in the last 3 years.

relationship between the share of direct coworkers (former and current) with the same occupation – henceforth *occupational homophily* – and the probability of changing occupation in Germany reveals a strikingly steep and negative trend (Figure 1.1). The first decile of occupational homophily changed their occupation 55% of the time, while the tenth changed 25%. This, however, does not necessarily reflect interactions between coworkers, as other factors related to the nature of the occupation or the employer, as well as individual preferences towards the occupation and network, could generate such a trend.

In this paper, I investigate empirically the effect of occupational homophily on the probability of changing occupation in Germany. I aim to provide an estimate of this relationship that can be interpreted as a result of network interactions.

The mechanism underlying the negative relationship between homophily and occupational change I aim to identify is the information flow generated by the interactions between coworkers. This argument rests on three premises. First, workers tend to have greater awareness of job opportunities within their own field or occupation,

rather than in the broader labour market (Marin, 2012). Second, coworker networks are a channel of information transmission; workers often obtain information about job opportunities from their connections, and seek referrals to potential employers (Topa, 2019). Third, the connections between coworkers are weak ties (in the terminology of Granovetter, 1973); workers share information with or refer their coworkers only if they are confident about the quality of their fellow coworkers (Antoninis, 2006; Marin, 2012).² Thus, if workers are more likely to hold information about job opportunities in their own fields, and relay them to coworkers who share the same occupation, then workers with connections in their own fields are more likely to encounter job opportunities in these fields, and less likely to change occupation.

I conduct my analysis on a linked employer-employee administrative dataset from Germany; Sample of Integrated Employer-Employee Data (SIEED).³ It covers the entire employment histories of a large sample of workers in Germany between 1975 and 2018, and allows me to track the complete coworker network of a large subset of this sample.⁴ I focus on establishment changes (job transitions) between 2012 and 2017, looking at the effect of occupational homophily at the end of the spell in the establishment of origin on the probability of holding a different occupation in the destination establishment.

After controlling for a set of observable confounders, the simultaneity in the worker's decision to form connections and change occupation might create biases in the estimates. I thus employ an instrumental variable approach to tackle the possible endogeneity in the model. I exploit the conditionally exogenous variation in the

² This is in contrast to the links with family and friends, who might provide information unconditionally.

³ I choose Germany for this analysis for two reasons. First, it has one of the highest rates of job and occupational mobility in Europe, with about 8% of workers changing jobs each year and half of these changes resulting in a switch of occupation (Bachmann et al., 2020). The SIEED dataset provide similar estimates. Second, previous studies have examined various coworker network metrics and their effects on outcomes in Germany, which serves as a useful benchmark for this study. For example, Glitz (2017) found a positive effect of coworker network employment rate on the probability of finding a job for workers displaced in mass-layoffs, while Cornelissen et al. (2017) identified a small positive effect of the productivity of coworkers on that of their peers. Jarosch et al. (2021) found a positive effect of the average wage of coworkers on future wage growth, especially for coworkers who share the same occupation.

⁴ Two workers are considered coworkers – connected – if they worked in the same establishment for at least 30 days in the 3 years prior their transition.

occupational homophily of a worker's network due to the firm hiring coworkers *after* them. I use the occupational composition of this subset of coworkers as an instrument for homophily of the worker's whole network. The identifying assumption is that – conditional on observables – the process of hiring new coworkers only depends on the decision of the firm and the new hires, but not that of the workers themselves or their unobserved characteristics. I then examine if the effect varies by occupation, wage rank, and other dimensions of homophily such as age and gender.

The results suggest a negative and large effect of occupational homophily on occupational change; a 10 percentage point increase in the share of coworkers with the same occupation is associated with a 1.6-2 percentage point decrease in the probability of changing the occupation, depending on the definition of occupation. There exists some heterogeneity in the effect by both worker and coworker characteristics. It is more pronounced for high skilled service occupation and production occupations compared to administrative occupations. Additionally, the effect is stronger when coworkers with the same occupation earns more than the worker in question. Furthermore, the results indicate that other dimensions of homophily, such as gender or age, also play a role. The impact of occupational homophily is greater when the coworker shares the same gender or age group as the worker.

A large body of literature has looked at the role of social networks in the labour market, given their widespread use. From a theoretical point of view, they facilitate the transmission of information between workers, and workers and firms, in labour markets characterised by search and information frictions (e.g., Mortensen and Vishwanath, 1994; Kugler, 2003; Bentolila et al., 2010; Galenianos, 2013, 2014; Dustmann et al., 2016; Heath, 2018). The theoretical literature generally predicts that individuals with access to contacts experience better labour market outcomes, i.e., higher wages and lower unemployment duration.⁵ While the empirical literature largely agrees on the negative (reducing) effects of networks on unemployment duration (e.g., Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021), the results on wages are mixed. Some

⁵ See Topa (2011, 2019) for a review on the theoretical and empirical literature on social networks in the labour market.

studies find a positive effect (such as Burks et al., 2015; Dustmann et al., 2016), and others find a negative effect (such as Bentolila et al., 2010; Ponzio and Scoppa, 2010).

One explanation to these mixed results is that social networks may cause a mismatch between the worker's skill and their occupation. That is, if a worker's job finding rate is higher for occupations where they have more connections, we will observe a negative effect of social contacts on wages if the occupation does not fit the worker's skill (Bentolila et al., 2010; Zaharieva, 2018; Alaverdyan and Zaharieva, 2022). Occupational change could thus be an intermediate causal path from networks to outcomes. By studying the role of networks in occupational change, this paper can reconcile some of the disagreement between the theoretical literature and some of the empirical results.

This paper is methodologically related to the growing body of literature using large establishment level administrative data to map coworker networks to study the characteristics of contacts on labour market outcomes such as wages, productivity, tenure, and unemployment duration (Cingano and Rosolia, 2012; Cornelissen et al., 2017; Saygin et al., 2021; Glitz and Vejlin, 2021; Gyetvai and Zhu, 2025; Jarosch et al., 2021; Eliason et al., 2022). Here, I explore new outcomes; how the occupational characteristics of coworkers affect where an individual finds a job, not only how quickly they exit unemployment or their wages. The studies on networks and occupational mobility are either theoretical (Zaharieva, 2018; Alaverdyan and Zaharieva, 2022) or test the effects indirectly (Bentolila et al., 2010). The value of using establishment level data is that allows me to tackle the relationship directly.

Finally, this paper contributes to the literature on the determinants of occupational change (e.g., Longhi and Brynin, 2010; Longhi and Taylor, 2013; Carrillo-Tudela et al., 2016; Bachmann et al., 2020; Carrillo-Tudela and Visschers, 2023). It introduces social networks as a new mechanism; previous studies do not generally consider networks as a factor in occupational change. As occupational movements are an important way workers cope with economic shocks (see, for example, Carrillo-Tudela et al., 2023), the paper sheds a light on social networks as a channel facilitating the reallocation of

workers between sectors, and a possible insurance device against these shocks.

The rest of this paper is structured as follows. Section 1.2 discusses the conceptual framework. Section 1.3 describes the empirical framework of this paper, and Section 1.4 the data. Section 1.5 shows the empirical results. Section 1.6 concludes.

1.2 Conceptual framework

In this section, I present the concept of occupational homophily I am using for this paper, as well as the theoretical rationale behind the relationship between occupational homophily in coworker networks and occupational change.

1.2.1 Occupational homophily

In this paper, I define occupational homophily as the share of coworkers with the same occupation. Homophily is the tendency of individuals with similar characteristics to associate (Lazarsfeld, 1954; McPherson et al., 2001). Homophily has different origins; *induced homophily* and *choice homophily* (Kossinets and Watts, 2009). Induced homophily is created by the social environment or the structure in which the individuals operate, and choice homophily is attributed to individual preferences.

I focus on induced homophily in my definition of occupational homophily, and analyse the role of the structure of the network in occupational change. Here, the degree of occupational homophily is the direct consequence of the occupational composition of the coworkers of a given workers. Coworker networks are formed in the workplace, so the worker's network is homophilous if they work in a firm with many others with the same occupation due to the nature of the firm, its productive requirements, and the patterns of mobility of its workforce. For example, an accountant working in an accounting firm will have a higher share of accountants among their coworkers than an accountant working in a fishing company. In addition, since there are frictions to occupational mobility (due, for example, to imperfect skill transferability), many of the former coworkers of the accountant in the accounting firm will also more likely be

accountants at any given point.

There are different choices when it comes to how to construct a variable reflecting occupational homophily in coworker networks using the data. For example, other than the share of coworkers with the same occupation, it could be the number of workers with the same occupation. I choose the share for my main analysis for two reasons. First, this follows the empirical literature on the effects of the composition of coworker networks on labour market outcomes, which generally uses shares (e.g., Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021).⁶ The second reason is a more practical one; using the raw numbers will result in a few observations with large values having a disproportionate effect on the coefficient estimates.⁷ As a robustness check, I ran the analysis with the number of coworkers with the same occupation and its log as independent variables (see Table A.9 Appendix A.4), and the results remain qualitatively the same as when using shares.⁸

1.2.2 Theoretical background

The main hypothesis of this paper is that workers with a higher share of coworkers with the same occupation are more likely to stay in that occupation. This rests on three premises. First, workers are more likely to have a greater awareness of job opportunities within their own fields than the broader labour market. Second, coworkers are a useful source of information about job opportunities. Third, workers are more likely to share that information with coworkers with the same occupation.

The first premise concerns the type of information about job opportunities the

⁶ Shares also fit better the models of information transmission in social networks (for example, Calvo-Armengol and Jackson, 2004; Calvo-Armengol and Jackson, 2007), where a worker samples randomly from their connections and the type of information they receive depends on the individual they sample. Here, the distribution of information received depends on the distribution of types rather than their numbers.

⁷ This is because the number of coworkers has an extremely skewed distribution, with most coworkers have relatively few connections and a relatively small number of coworkers with a large number of connections (see Figure 1.4). Dealing with that issue by using, for example, a log transformation will force us to drop the individuals without any connections, although having no connections with the same occupation has an economic significance.

⁸ The probability of occupational change decreases with the number of coworkers with the same occupation and, in all specifications, either increases with the number of coworkers with a different occupation or is not significantly affected by it.

worker possesses, and possibly relays to their coworkers. Workers are more likely to come across opportunities in their own occupations through their own experiences, current or past job search. For example, Marin (2012) suggests that workers are more likely to be aware of job opportunities that are related to their own field or occupation. Carrillo-Tudela et al. (2023) found that, during the COVID-19 pandemic in the UK, about 56% of employed workers searched for a job in the same occupation as the one they currently hold, or previously held, and 69% of the non-employed searched for a job in an occupation they previously held.⁹ Generally, while the flow of occupational mobility is high and depends on the context (e.g., sector, country, demographics, etc.), most workers stay in the same occupation following a transition (Bachmann et al., 2020). This indicates that workers tend to search for jobs in their own occupations, thus influencing the type of job opportunities they possess.

The second premise relates to the broad role of social networks as a channel for information transmission about job opportunities in labour markets characterised by search frictions. Coworker networks are a part of these networks. Here, workers may share this information with the connections form during their employment spells, and refer or recommend these connections to potential employers. A multitude of studies have found evidence in support of this mechanism. For example, Cingano and Rosolia (2012), Glitz (2017), and Saygin et al. (2021) found that a higher share of employed coworkers in a given individual's network reduces their unemployment duration or increases their re-employment probability. Caldwell and Harmon (2019) find a positive relationship between labour demand in the former coworkers' current firm and job-to-job mobility. These studies indicate that, broadly, workers with better networks – larger or with a higher share of workers with useful information – have higher job arrival rates.¹⁰

⁹ These results are from Table 3 in Carrillo-Tudela et al. (2023).

¹⁰ There are two types of theoretical literature on information transmission about job opportunities in the job market. The first approach considers informal social networks as an alternative source of information to formal methods (e.g., Mortensen and Vishwanath, 1994; Bentolila et al., 2010; Galenianos, 2013; Zaharieva, 2013). Here, the job arrival rate has a network component and a general (other) channels component. The other strand of literature models explicitly the information transmission process (e.g., Calvo-Armengol and Jackson, 2004; Calvo-Armengol and Zenou, 2005; Calvo-Armengol and Jackson, 2007; Horváth, 2014). In its simplest form (in Calvo-Armengol and Jackson, 2004), workers are

The third premise is that workers are more likely to share the information with those with the same occupation. This is for two main reasons. First, homophily; people generally tend to associate with others who share the same characteristics (McPherson et al., 2001). Workers will have a higher contact rate with similar coworkers thus increasing the probability of information exchange.¹¹ Gyetvai and Zhu (2025) found that the employment rate among coworkers with the same occupation has a stronger effect in reducing unemployment duration, compared to those holding different occupations. Second, workers tend to share information or recommend coworkers only if they are confident in their ability (Antoninis, 2006; Marin, 2012).¹² Information sharing between coworkers is considered an investment in reputational capital; a coworker might incur a reputational penalty if the individual they refer is not productive (Topa, 2019). Coworkers with the same occupation are better positioned to observe each others' quality, and thus more likely to share that information with each other.

In summary, if workers hold information about job opportunities in their own occupations, and that this information is subsequently relayed to coworkers with the same occupation, then we can expect that workers with a higher proportion of coworkers with the same occupation are less likely to change occupations. In other words, occupational homophily might increase the job arrival rate in the original occupation. Thus, two identical workers will exhibit different job mobility patterns given the occupational composition of their respective networks.

Figure 1.2 illustrates the hypothesis. In an ideal setting, assume we have two identical workers; i and h , and situated in two mostly identical networks. The difference is that i 's network features a low level of homophily – i 's occupation is 1, and most

embedded in social networks. They randomly receive pieces of information on job opportunities, and choose whether to pass them along. The employed will pass the information along to their contacts, and the unemployed will keep it for themselves. This framework predicts a positive correlation between the probability of receiving information about job opportunities and the network employment rate.

¹¹ McPherson et al. (2001) suggest that the most significant dimension of homophily in the workplace are gender and education, which in turn determine the occupation held by the worker.

¹² A strand of literature studies the role of referrals in reducing uncertainty about potential job candidates (Montgomery, 1991; Casella and Hanaki, 2008). A social connection is familiar with ability of the candidate the refer, and relays that information to the employer. Glitz and Vejin (2021) provides evidence that workers have a propensity to transition into establishments where their former coworkers are employed, which is interpreted as a result of referrals.

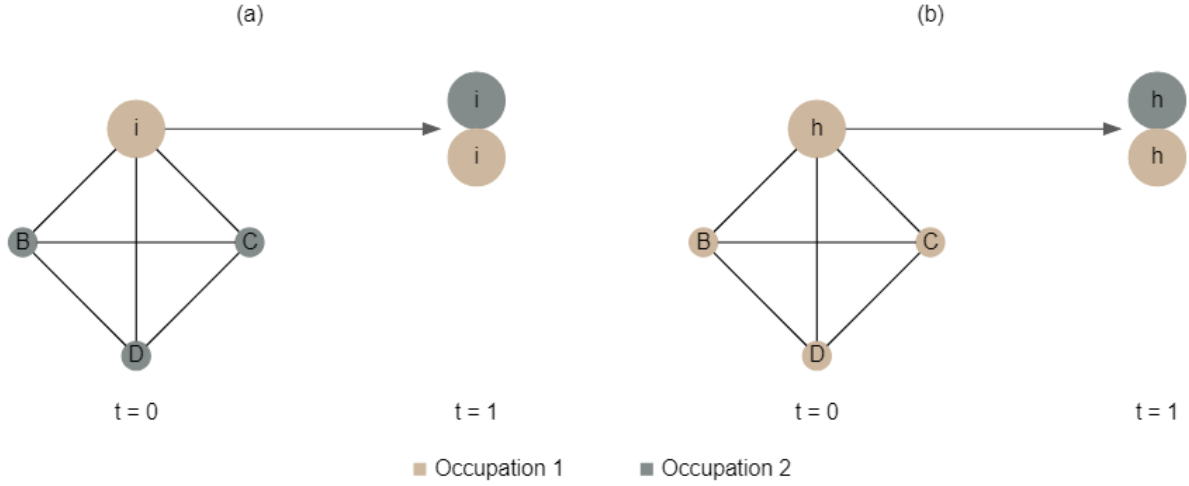


Figure 1.2: Framework; occupational homophily and occupational change

Notes: Nodes represent workers. The headless arrows represent links, and the horizontal arrows are the possible occupational transitions from $t = 0$ to $t = 1$. The node colour represents occupation.

of their coworkers are in 2. On the other hand, h 's network features a high level of homophily; most of their coworkers are in 1. I test if the probability of i changing their occupation is different than that of h . By accounting for possible confounders and other sources of endogeneity, I aim to provide an interpretation to the estimates that could be attributed the mechanisms above. The empirical framework of this paper is shown in the next section.

1.3 Empirical framework

1.3.1 Regression framework

My objective is to estimate the effect of occupational homophily on the probability of occupational change following a job change. The baseline is a linear-in-means local interaction model similar to that of Cingano and Rosolia (2012) and Glitz (2017), where the worker's outcome y_i is assumed to be linearly dependent on the average fixed characteristics of their direct peers. The estimating equation is

$$y_i = \beta \bar{m}_{-i} + \eta \ln(D_i) + \delta x_i + \epsilon_i \quad (1.1)$$

where y_i is a binary variable taking the value of 1 if i 's occupation in the destination establishment is different to that of the origin establishment. $\bar{m}_{-i}$ is occupational homophily, the explanatory variable of interest, and β is the parameter of interest. $\bar{m}_{-i}$ is the share of direct coworkers with the same occupation, measured on the last day of the spell in the origin establishment. It can be expressed as

$$\bar{m}_{-i} = \frac{\sum_{j \neq i} g_{ij} \mathbb{1}[m_i = m_j]}{D_i} \quad (1.2)$$

where m_i is the worker's occupation, g_{ij} is an element of a $N \times N$ unweighted and undirected adjacency matrix G , taking the value of 1 if workers i and j are connected and zero otherwise. D_i is the total number of coworkers of worker i (also known as degree), which is the sums of the rows of the matrix G .¹³ x_i is a set of controls, and ϵ_i is the error term.

The set of variables x_i captures the observed possible confounders. These are factors that are potentially correlated with both occupational homophily and the probability of changing occupation. I group these variables into four categories; worker characteristics, employment characteristics in the origin establishment, establishment characteristics, and common aggregate shocks.

Worker characteristics include the standard set of demographics such as gender, age, level of education, and the overall experience in the occupation of origin, as well as employment characteristics in the network building phase – three years prior their last in day in the origin establishment. The employment characteristics capture the workers' patterns of mobility. Occupational homophily is generated by these patterns, and simultaneously, there might be some factors (e.g., job search behaviour or risk taking behaviour) leading some workers to be more or less mobile, affecting their probability of changing occupation. I thus include dummies of the number of employers and the number of different occupations held, and the average establishment size in the network building phase.

Employment characteristics in the origin establishment include occupation, last

¹³ The procedure to map coworker networks is described in detail in Appendix A.3.1.

daily wage, part-time, tenure in the establishment, and the average wage of coworkers. Occupation of origin is possibly the strongest determinant of both occupation homophily and the probability of changing occupation. If the occupation exhibits high turnover, as a result of the nature of the occupation or it being a declining occupation, we would expect its workers to have mechanically a high level of homophily. Simultaneously these workers may change their occupations more often. I control for the occupation of origin and occupation-year fixed effects to capture these mechanisms.

Origin establishment characteristics are controlled for with establishment effects and the time varying establishment size. Establishment fixed effects capture several elements. First, they capture industry and location, by definition. They also capture some features that cannot be controlled for with industry and location dummies. They determine what the proportions of different occupations are necessary to produce, which in turn determines the occupational composition of coworker networks. They also capture for the average performance of the establishment, as well as other establishment specific factors - such as hiring policies, average turnover, performance, training, and common specific human capital accumulated within.

With the establishment fixed effects, I am comparing the outcomes of individuals who transitioned from the same establishment. Of course, if two workers with the same occupation join the establishment at the same time and leave at the same time, their occupational homophily will be identical. The sources of variation stem from the timing of joining or leaving the establishment, and occupation. If two workers have the same occupation, the difference in occupational homophily comes from the timing of joining the establishment and that of the transition out of the establishment. If they have different occupations, the difference is the result of occupational homophily being constructed on the basis of occupation of origin.

Finally, I control for time varying aggregate common shocks that could influence the probability of changing occupation for workers with some level of occupation homophily. These include year, industry-year, and state-year fixed effects.

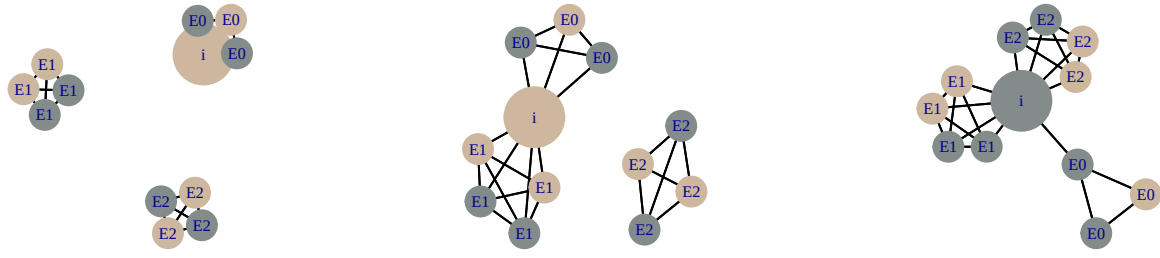
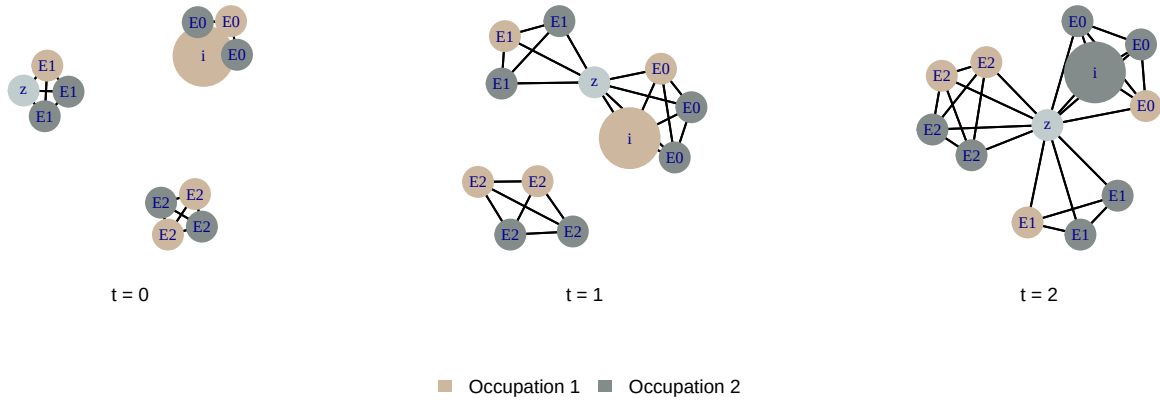
The benchmark is a linear probability model (estimated with OLS).¹⁴ In the next Section (1.3.2), I discuss the identification of the parameter of interest and the instrumental variable strategy to deal with potential bias.

1.3.2 Identification

Identifying the causal effect of occupational homophily on occupational change as a result of information transmission between coworkers requires the independence of the variable of interest from the error term. It is violated if there are unobserved factors omitted from the model that are correlated with the process of network formation and the worker's occupational choice. Not accounting for these factors will bias the estimates of the coefficient of interest, as it will capture factors other than network transmission mechanisms.

Even after controlling for the observed worker characteristics, experience in the labour market in the network building phase, and aggregates, there might remain some unobserved individual behaviours influencing both the explanatory and dependent variables. These are related to the simultaneity of worker's decision to form networks and to change occupation. The first example is occupational homophily by choice – workers who choose firms as to maximise the number of coworkers with the same occupation. For instance, consider an engineer who will always remain an engineer, and prefers to work with and be surrounded by other engineers, so they will choose to work in firms with a large share of engineers. Second, the worker's strategic behaviour. A worker might join a firm with a certain occupational composition with the intention to use their new connections to move to another occupation. In both cases, the mechanisms lead workers to simultaneously choose to work in a firm with a large share of coworkers with the same occupation and choose to stay in the same occupation after a job transition. Thus, OLS estimates will not uniquely capture the information channel, and will overstate the effect of occupational homophily on occupational change.

¹⁴ The marginal effects of a similar logit model give similar results. I chose a linear probability model for its flexibility and interpretability.

Panel A: Worker i as an entrant**Panel B:** Worker i as an incumbent**Figure 1.3:** An example of the formation of a coworker network

Notes: Nodes represent workers, and the links are formed if they worked in the same establishment in the recent past. The node colour represents occupation.

One common strategy to deal with endogeneity in network models is to exploit random assignment to groups and exogenous shocks to networks. In related papers studying the effects of coworkers networks on labour market outcomes, this shock is mass layoffs (e.g., Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021; Eliason et al., 2022). Another solution is to randomise the assignment using instrumental variables affecting the outcome of interest only through the composition of the network (e.g., Rege et al., 2012; Glitz, 2017). For example, Glitz (2017) instruments the network employment rate with the share coworkers displaced in mass layoffs after separation.

I use another source of exogenous shocks; *new hires*. I use the characteristics of coworkers z hired in worker i 's establishment *after* them as an instrument for occu-

pational homophily. This is based on the fact that coworker networks are formed by two processes; first, the movement of workers between establishments, and second, the movement of coworkers around them as a result of turnover within these establishments. Occupational homophily will reflect the occupational structure of the establishments they join, and the turnover within their occupation.

Figure 1.3 illustrates the resulting pattern and shows how the mobility of workers alters the composition of the network. In the first case (Panel A), call worker i an *entrant*, and in the second an *incumbent* (Panel B).¹⁵ The instrument of choice z_i is thus the number of i 's entrant coworkers with the same occupation when i is an incumbent (that is the number of z nodes as shown in Figure 1.3).¹⁶

The identifying assumption is that, conditional on observables, the instrument has to alter the worker's network independently from the worker's decision to form connections and to change occupation. While the first process is generated by i 's labour market decisions, I argue that z 's choice to join the firm is independent from these decisions. From i 's point of view, the influx of new coworkers is a random shock to their network. This is because these additions are induced by the movement of their coworkers around them, which is determined by the decisions of other workers to join the firm, and the firm to hire them. The firm might hire new workers due to exogenous demand increases, or a replacements for other coworker who moved, which are not necessarily correlated with the worker's decision to change occupation.

This instrument is valid only if a worker's decision to move is uncorrelated with the hiring of a new coworker. This condition would be violated if the new hire is a replacement for the worker in question. In that case, the hiring decision would be a direct consequence of the worker's decision to move. To address this concern, the instrument includes only coworkers who joined the establishment at least six months before the worker changes establishments. A second concern arises in smaller firms, where individual workers may influence hiring decisions. In such cases, a worker

¹⁵ Following the terminology of Hensvik and Skans (2016).

¹⁶ The instrument z_i is a raw number instead of a ratio over network size, i.e., $\frac{z_i}{D_i}$, because network size is potentially endogenous. However, the results are almost the same using the ratio.

could affect the firm's occupational composition directly. To mitigate this issue, I exclude coworkers hired in smaller establishments (fewer than 25 employees) from the set of new hires.¹⁷

1.4 Data and descriptives

1.4.1 Data

I use the dataset Sample of Integrated Employer-Employee Data (SIEED), published by the Institute for Employment Research (IAB) of the Federal Employment Agency (BA). It is a linked employer-employee administrative dataset containing information about wages, employment characteristics, skill levels, as well as establishment characteristics for a large sample of German workers between 1975 and 2018. These workers are sampled as follows; 1.5% of all establishments of Germany are chosen randomly, then, the whole employment histories of all the workers who have ever worked in these establishments are added (Schmidtlein et al., 2020). Establishments are units of the same company, situated within the same municipality and performing the same economic activity. This means that, with the exception of very large establishments, an establishment is likely to represent a single workplace.

I focus on prime age workers (between 24-55) who changed establishment¹⁸ between 2012-2017.¹⁹ I include both job-to-job and non-employment to employment to non-employment transitions. I restrict the analysis to workers who change establishment to compare the outcomes of workers who searched for or are willing to take up a new job. In addition, I select workers who spent more than one year in the origin establishment, and joined the destination establishment within one year of leaving. I

¹⁷ For the establishments smaller than 25 employees, the value of the instrument is zero. The results remain the same when this restriction is relaxed.

¹⁸ Because the aim is to study job changes, I exclude the changes that occurred in the last year the establishment observed, where more than 80% of its workers moved to the same establishment. These might be establishment identification changes, changes in ownership or location. See Hethey-Maier and Schmieder (2013) for a discussion of the definitions of establishment changes.

¹⁹ I exclude observations before 2012 because new occupational codes were introduced in 2011, leading to inaccuracies.

thus focus on transitions from stable jobs for workers who are attached to the labour market. I also exclude workers who held a second job at the end of the spell in the establishment.

I select the workers for whom we can observe the complete set of present and past coworkers. Those are who worked exclusively in the sampled establishments (panel establishments) in 3 years preceding their transition – the network building phase.²⁰ I also exclude those who worked in very large establishments (more than 300 workers), because any two given workers are more likely to have made contact in smaller workplaces. Finally, I exclude the firms that contain less than 5 of these workers to provide enough within establishment variation.

Coworkers are those who worked with the selected workers in the same establishment at the same time for at least 30 days in the past 3 years. I exclude those who were not employed at the time of transition, as these may be unable or unwilling to provide information on job opportunities (similar to the logic in Glitz, 2017). The sample used to construct the instrumental variable is a subset of the coworker network that only includes those who joined the establishment after the worker in question. Occupational homophily, as well other network characteristics, are measured on the last day of the worker's spell in the origin establishment.

I use two occupational classifications; the 3-digit KldB classification and the Blossfeld (1987) occupational classification (henceforth Blossfeld). The first is the standard German occupation definition and it's detailed; containing around 340 occupations. As the cell sizes in these occupation can be small, a more aggregate measure is useful for sub-sample analysis. It is also valuable to assess if the effect of occupational homophily varies by level of aggregation. For this, I use the definition of occupation in Blossfeld (1987).²¹ It maps the less aggregated definition into 12 broad occupational definitions.

²⁰ There are two types of establishments in this dataset; panel establishments and *non-panel* establishments. Panel establishments are the 1.5% that were originally sampled. We can observe their entire workforce at any time. The second type is non-panel establishments. As we observe the whole employment history of the selected workers, some of their spells are in non-sampled establishments. We do not observe for the entire workforce of these establishments.

²¹ The conversion code between KldB and Blossfeld for the IAB data is provided by Dauth and Epelheimer (2020). See Tables A.3-A.5 for the mapping between KldB and Blossfeld.

Table 1.1: Distribution of Blossfeld occupations

	N	% in total	% changing occupation
Total	41300	100	40.2
Production			
Agriculture	438	1.1	39.7
Unskilled manual	3463	8.4	40.9
Skilled manual	5194	12.6	39.0
Technicians	1585	3.8	44.4
Engineers	1242	3.0	36.6
Services			
Unskilled services	7755	18.8	31.4
Skilled services	1633	4.0	33.8
Semiprofessions	4997	12.1	17.4
Professions	724	1.8	38.8
Administration			
Unskilled administrative and commercial	4258	10.3	40.5
Skilled administrative and commercial	8253	20.0	25.8
Managers	1758	4.3	43.5

Notes: The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The conversion code is provided by Dauth and Eppelsheimer (2020).

They are divided into three fields (production, services and administration), and each field is divided by skill level (see Table 1.1). The nice feature of this classification is that it divides occupations horizontally and vertically, reflecting the type of task performed and level of skill.

1.4.2 Descriptive statistics

Table 1.2 shows the main descriptive statistics of the sample. The worker sample contains 41,300 individuals transitioning from 3,303 origin establishments split roughly equally over the years included in the analysis.²² Most workers in the sample stay in the same occupational following a transition, but the share of workers who changed occupation is still high, at 45% and 31% in the 3-digit KldB and Blossfeld classifications, respectively.²³

The average age is 39 (median of 39), and 45% of the sample is female. As for the level of education, 13% are without qualifications (low education), 69% received some vocational training (medium education), and 18% have a university degree (high edu-

²² See Table A.2 for the composition of the sample by year, state, and industry.

²³ Figure A.1 in Appendix A.1 show that this share is stable during the period of analysis (2012-2017).

Table 1.2: Summary statistics; worker sample

	Mean	Standard Deviation	10th Percentile	50th Percentile	90th Percentile
Occupational change					
3-digit KldB	0.31				
Blossfeld	0.45				
Individual characteristics					
Age	39.16	9.32	26	39	52
Female	0.45				
German	0.89				
Low education	0.13				
Medium education	0.69				
High education	0.18				
Employment characteristics					
Number of employers (last 3 years)	1.01	0.07	1	1	1
Average establishment size (last 3 years)	84.66	64.41	16.50	67	181.75
Number of occupations held (last 3 years)					
3-digit KldB	1.26	0.51	1	1	2
Blossfeld	1.19	0.43	1	1	2
Experience in occupation (days)					
3-digit KldB	2,606.18	2,420.32	535	1,725	6,292
Blossfeld	3,311.04	2,707.68	672	2,390	7,440
Origin establishment					
Last daily wage**	89.02	74.07	20.23	74.05	161.02
Part-time	0.28				
Tenure in establishment (days)	2,516.83	1,982.72	730	1,857	5,277
Establishment size	82.08	65.90	14	63	179
Destination establishment					
Days until start of new job	42.40	75.75	1	1	154
Starting daily wage**	89.91	70.58	22.08	76.14	161.08
Establishment size	332.79	1,834.13	6	58	530
Network characteristics					
Number of connections					
All coworkers (degree)	146.41	126.40	26.50	112	313
Current coworkers	82.03	65.90	14	63	181
Former coworkers	64.38	85.90	8	37	149
Occupational homophily					
3-digit KldB	0.31	0.26	0.02	0.24	0.70
Blossfeld	0.45	0.27	0.09	0.45	0.82
Average wage of coworkers**	85.06	43.15	44.22	73.22	145.07
Instrumental variable					
Number of entrants with the same occupation					
3-digit KldB	20.07	34.22	0	7	55
Blossfeld	32.97	50.44	0	16	87
Observations	41,300				
Number of origin establishments	3,303				

Notes: The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. Coworkers are defined as those who worked in the same establishment at the same time for at least 30 days in the last 3 years. The instrumental variable is the number of coworkers with same occupation who were hired after the worker. **Wages above the social security contribution ceiling are top coded. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009); Card et al. (2013). See appendix A.3.2 for the imputation details.

cation). Most workers in the sample worked in one establishment in the past 3 years (the maximum is 3), and held one occupation. This indicates that the majority of workers formed their networks in a single establishment and did not change occupations during the network-building phase. The average tenure in the origin establishment is

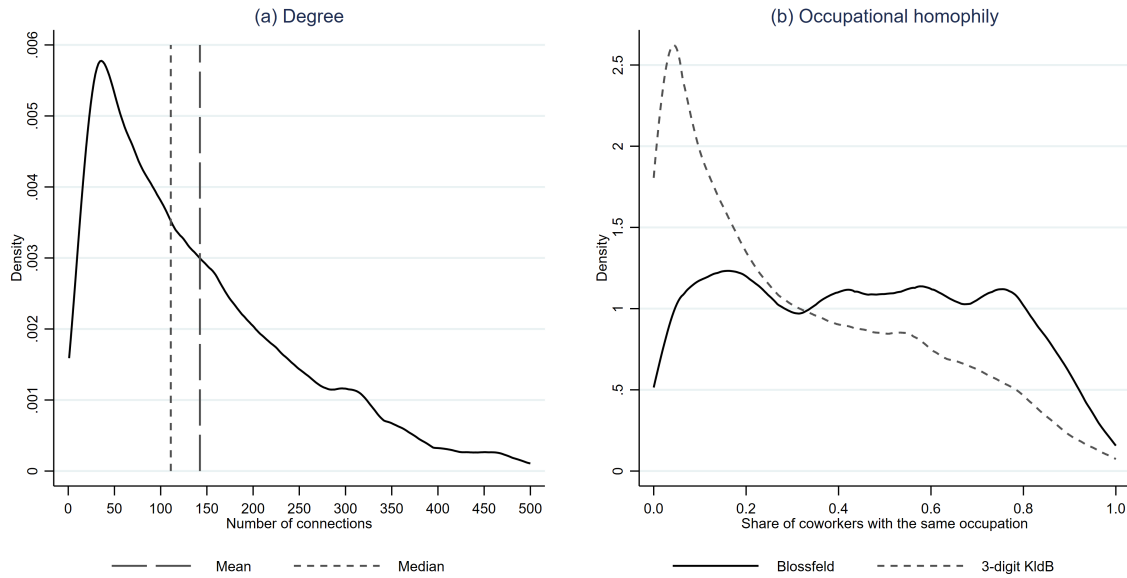


Figure 1.4: Degree and homophily distributions

Notes: Author's calculations using SIEED. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

7 years, and the average overall experience in the 3-digit KldB (Blossfeld) occupation is 7 years (9 years).

Now, we compare the worker employment characteristics in the origin establishment to the destination establishment. On average, workers joined the destination establishment 43.4 days after they left the origin establishment, and about half the sample joining the destination establishment without a break (in one day). The average origin establishment size is 82 workers. It should be noted that the average size of the destination establishment is higher than the origin (332 workers) because I do not impose restrictions on the size of the destination establishment. The last average daily wage in the origin establishment is similar to that of the new wage in the destination establishment (89 euros to 89.9 euros). However, the wage in the destination establishment is less dispersed, with standard deviations of 70.58 and 74.07, respectively.

The sample of workers used to construct the network variables consists of 592,085 workers (see Table A.1). Comparing the fixed characteristics of the coworker network sample with the main sample; we see that, first, the share of females is similar (47% and 45%, respectively). They are on average from a younger cohort than the main sample

(average age in 2015 is 38.8 and 39.6 respectively). The average daily wage calculated over the analysis period is lower in the coworker sample (69.2 euros to 89.7). This difference is due to the selection of workers who had a stable job in this period – a restriction not present in the network sample. The subset of coworkers to construct the instrumental variables consists of 437,855 workers, and is similar to the network sample with respect to these fixed characteristics.

The rest of Table 1.2 and Figure 1.4 review the main features of the network variables. The average size of the network (degree) is 146 coworkers, with an average of 82 current coworkers who are still employed in the worker’s origin establishment, and 64 former coworkers who have separated before the cutoff day. The degree distribution exhibits a power-law like distribution with a heavy tail (Panel (a) in Figure 1.4); many workers with a few connections, and a few workers with many connections. Since most workers stayed in one establishment in the network building phase, this pattern reflects the size of the origin establishment, and the tenure within it. Tenure within establishments explains why the average network size is higher than the average establishment size.

The shape of the occupational homophily variables depends on the definition of occupation. On average, the share of coworkers with the same detailed 3-digit occupation is 31%, while it is higher for the Blossfeld occupation at 45%. This is due to the fact that it is less likely for two workers to perform the exact same task in the same small establishment at the same time. Looking at the distribution of the two definitions in panel (b) in Figure 1.4, we see a mass of workers between 0 and 0.2 in the detailed 3-digit definition, which is not present in the Blossfeld definition. In the Blossfeld definition, the distribution remains relatively constant across all possible values, with lower concentrations at the extremes.

1.5 Empirical results

1.5.1 The effect of occupational homophily on occupational change

Table 1.3 reports the main results of this paper. I compare the results by specification and occupational definition. Column (1) shows the the estimates without controls. Columns (2) and (3) include the benchmark set of controls; demographics, employment, tenure and experience variables, occupation of origin dummies, and some time varying firm characteristics. Columns (1) and (2) are linear probability models and (3) reports the marginal effects of a logit a model. Columns (4) and (5) introduce industry-state interactions, and (6) and (7) include establishment fixed effects, with the even numbered columns estimated with OLS and the odd numbered columns with 2SLS.

The results show that the effect of occupational homophily on occupational change is negative, significant, and large in magnitude for all specifications. Focusing on the Blossfeld definition in Panel A, we see that benchmark controls in Columns (2) and (3) account for a large part of the unobserved heterogeneity, and the coefficient falls from -0.324 in the unconditional estimates in Column (1), to -0.240 in the OLS regression with the benchmark controls in Column (2). We also see that the estimates with the benchmark controls in Columns (2) and (3) are similar to those with industry-state interactions with a coefficient of -0.24 in Column (4). The effect goes down slightly when we introduce establishment fixed effects to reach -0.2 in Column (6).

Columns (5) and (7) show the 2SLS results. The instrument is strong with a first stage F-statistic of more than 130 for both cases. The OLS results are very similar to those of 2SLS. This indicates that the benchmark controls account for most of the factors driving both occupational homophily and occupational change. The estimates in Column (7) suggest that a 10 percentage point increase in the share of coworkers with the same Blossfeld occupation (37% of the standard deviation)²⁴ is associated with a 2 percentage point decrease in the probability of changing the Blossfeld occupation.

Now, we turn to the more detailed KldB definition. We see that the results are

²⁴ See Table 1.2.

Table 1.3: Occupational homophily and occupational change; main results

	Benchmark			Industry×location interactions		Establishment Fixed effects	
	OLS (1)	OLS (2)	Logit (3)	OLS (4)	2SLS (5)	OLS (6)	2SLS (7)
Panel A: Blossfeld							
Occupational homophily	-0.324*** (0.008)	-0.240*** (0.014)	-0.251*** (0.017)	-0.240*** (0.015)	-0.232*** (0.042)	-0.205*** (0.018)	-0.194*** (0.053)
ln(degree)		-0.005 (0.010)	-0.006 (0.011)	-0.008 (0.009)	-0.007 (0.010)	-0.031** (0.014)	-0.030** (0.015)
<i>First stage</i>							
Instrumental variable					0.002*** (0.000)		0.002*** (0.000)
First stage F-stat					175.42		132.46
Panel B: 3-digit KldB							
Occupational homophily	-0.353*** (0.009)	-0.197*** (0.020)	-0.228*** (0.024)	-0.189*** (0.018)	-0.171*** (0.042)	-0.125*** (0.021)	-0.165** (0.066)
ln(degree)		-0.008 (0.012)	-0.005 (0.014)	-0.014 (0.010)	-0.012 (0.011)	-0.044*** (0.014)	-0.048*** (0.015)
<i>First stage</i>							
Instrumental variable					0.003*** (0.000)		0.002*** (0.000)
First stage F-stat					289.26		129.43
Observations	41235	41300	41299	41239	41300	41300	41300
Number of establishments						3303	3303
Controls		X	X	X	X	X	X
<i>Fixed effects</i>							
Occupation×year				X	X	X	X
Industry				X	X		
State				X	X		
Industry×state				X	X		
Industry×year				X	X	X	X
State×year				X	X	X	X
Industry×state×year				X	X		
Establishment					X	X	

Notes: Dependent variable is the probability of changing occupation. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. Column (1) shows the estimates with no controls. (2)-(3) are the baseline estimates with controls including occupation in the origin establishment, education, age (squared), female, German, log last wage in the origin establishment, part-time, tenure in firm (squared), experience in occupation (squared), number of employers, number of occupations held, and average establishment size (squared) in the last 3 years, establishment size (squared), year fixed effects, and average wage of coworkers. The logit coefficients in column (3) are marginal effects at the means. Columns (4)-(5) add industry, state, industry×year, industry×state, industry×state×year, Blossfeld occupation×year fixed effects. Columns (6)-(7) introduce establishment fixed effects. The instrumental variable is the number of coworkers with the same occupation who were hired after the worker in question (entrants). Standard errors are robust and clustered at the establishment levels (in parenthesis). * p<0.1, ** p<0.05, *** p<0.01.

similar to those of the Blossfeld definition. The effect is negative, large, and similar across specifications. The coefficient in Column (7) show that a 10 percentage point increase in the share of coworkers with the same 3-digit KldB occupation (37% of the standard deviation) is associated with a 1.6 percentage point decrease in the probability

of changing the Blossfeld occupation.

The results are in support of the hypothesis that workers who have a lot of connections with the same occupation are on average more likely to stay in that occupation following a job change. This relationship cannot be explained by the workers' demographic characteristics, their occupation, the characteristics of the workplace, or their strategic behaviour to form connections in order to find a job in a certain occupation.²⁵ In the next Section 1.5.2, I conduct heterogeneity analysis along several dimensions (occupation, wages, demographics), investigating if the effects are different by the type of the worker and their coworkers.

1.5.2 The role of worker and coworker characteristics

Occupation of origin

In this section, I investigate how the effect of occupational homophily varies across occupations in the establishment of origin. Information may play a different role according to the occupation; three potential mechanisms could explain the variation in the effect of homophily across occupations. First, some occupations may pose a challenge for firms to assess the productivity of potential hires ex-ante. In such cases, referrals from coworkers may carry more weight in the decision to hire, making the effect of homophily stronger. Second, the capacity of coworkers to observe the skill and productivity of their colleagues varies across occupations. Since coworkers are generally more likely to refer or transmit information about job opportunities to their fellow coworkers if they are sure about their qualifications, if it is easier for coworkers to assess the performance of their peers, networks may play a more significant role. Third, some occupations – by nature – require workers to work more closely together, leading to more frequent interactions.

I split the sample by Blossfeld field (production, services, and administration), as well as occupation within each field to examine their significance and relative magnitudes. I use the Blossfeld definition and I estimate the model using 2SLS with speci-

²⁵ Additional results and robustness checks are in appendices A.4.

Table 1.4: Occupational homophily and Blossfeld occupational change; heterogeneity by field and occupation of origin

	Field	Blossfeld Occupation				
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Production Occupations						
	All Production	Agriculture	Unskilled manual	Skilled manual	Technicians	Engineers
Occupational homophily	-0.245*** (0.071)	-1.763 (2.524)	-0.564*** (0.173)	0.141 (0.154)	-0.710*** (0.188)	-0.444** (0.181)
ln(degree)	-0.022 (0.014)	-0.088 (0.133)	-0.070** (0.029)	-0.018 (0.024)	0.125*** (0.044)	0.015 (0.056)
First stage F-stat	1006.312	1.599	198.704	309.412	186.008	147.822
Observations	11922	438	3463	5194	1585	1242
Panel B: Service Occupations						
	All Services	Unskilled services	Skilled services	Semi professions	Professions	
Occupational homophily	-0.399*** (0.055)	-0.401 (0.261)	-0.602*** (0.141)	-0.376*** (0.077)	-0.956*** (0.316)	
ln(degree)	0.008 (0.011)	-0.007 (0.021)	0.106*** (0.041)	-0.025 (0.021)	-0.132** (0.062)	
First stage F-stat	877.889	100.145	321.466	439.037	41.694	
Observations	15109	7755	1633	4997	724	
Panel C: Administration Occupations						
	All Admin.	Unskilled admin.	Skilled admin.	Managers		
Occupational homophily	-0.092* (0.049)	-0.035 (0.228)	0.193** (0.082)	-0.198 (0.270)		
ln(degree)	-0.014 (0.012)	-0.052** (0.026)	0.026 (0.017)	0.042 (0.036)		
First stage F-stat	779.878	143.110	271.502	131.769		
Observations	14269	4258	8253	1758		

Notes: Dependent variable is the probability of changing the Blossfeld occupation. Estimated with 2SLS following the specification in column (5) Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The instrumental variables is the number of coworkers with the same occupation who were hired after the worker in question (entrants). Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

fication of Column (5) in Table 1.3 (industry $\times$ state interactions). Table 1.4 shows the results of the exercises; each panel represents a field, Column (1) shows the overall effect within the field, and the remaining columns divide the sample by individual occupation.

The effect of occupational homophily seems to vary greatly by field, but is generally increasing in the skill level. It is most pronounced in service occupation (coefficient of -0.4) followed by production (coefficient of -0.25), but low in magnitude and weakly

significant for administration occupations (coefficient of -0.09). In the production field, the effect is negative across the different level of skill except for agriculture and skilled manual occupation. However, the effect in the service occupation is negative and significant for the upper level of skill; skilled services, semiprofessions and professions, but insignificant for unskilled service. For the administration occupations, the coefficients for managers and unskilled administrative and commercial occupations are insignificant. The coefficient switches signs and becomes positive and significant for skilled administrative occupations. For this group, the higher the share of workers with the same occupations, the higher the probability of changing occupations.

In summary, the effect of occupational homophily on occupational change is larger for service occupations compared to production and administrative occupations. The effect is also generally higher the higher level of skill.

Position in the wage distribution within occupation

Here, I examine how the position of the worker in the wage distribution *within* occupation relative to that of their coworkers affects the influence of occupational homophily on the probability of changing occupation. The underlying motivation is that, in many theoretical models on information transmission between coworkers (e.g., Calvo-Armengol and Jackson, 2004; Calvó-Armengol and Zenou, 2005; Calvó-Armengol and Jackson, 2007), workers share job opportunities with coworkers only if they cannot use them themselves, and coworkers use this information only if the wage associated with the vacancy is higher than their current wage. The implication is that the negative effect of occupational homophily on the probability of changing occupation would be more pronounced for workers ranked lower in the within occupation wage distribution with coworkers ranked higher in the distribution.

To create a comprehensive view of wage distribution within the occupation, I utilise all workers observed in the dataset between 2012 and 2017 (3,812,471 workers). I create yearly panels with the cut-off on the 30th of June, and divide the sample into 3 wage terciles within each Blossfeld occupation each year. Workers and their coworkers

Table 1.5: Occupational homophily and Blossfeld occupational change; heterogeneity by position in the wage distribution within Blossfeld occupations

	1st tercile (1)	2nd tercile (2)	3rd tercile (3)
Occupational homophily			
1st tercile	-0.639*** (0.113)	-0.357*** (0.096)	-0.205 (0.127)
2nd tercile	-0.271** (0.129)	-0.356*** (0.091)	0.131 (0.118)
3rd tercile	-0.532*** (0.171)	-0.624*** (0.115)	-0.268*** (0.071)
ln(degree)	-0.029** (0.014)	-0.039*** (0.012)	-0.009 (0.014)
First stage F-stat	145.164	215.579	199.715
P-value of coefficient equality			
1st = 2nd	0.018	0.999	0.043
1st = 3rd	0.588	0.078	0.677
2nd = 3d	0.195	0.036	0.004
Observations	11786	15904	13610

Notes: Dependent variable is the probability of changing the Blossfeld occupation. Estimated with 2SLS following the specification in column (5) of Panel A, Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. For each column, the sample is split according to wage terciles within occupation. Occupational Occupational homophily by tercile is number of coworkers with the same occupation who belong to a given tercile, divided by degree. The instrumental variables are constructed accordingly; the number of entrant coworkers in the wage tercile with the same occupation. Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

are assigned to their respective terciles. The occupational homophily is split accordingly. Let $r_{j,m,t} \in \{1, 2, 3\}$ be the tercile of worker j in occupation m and year t , then the share of coworkers with the same occupation in the tercile r will be

$$\bar{m}_{-i}^r = \frac{\sum_{j \neq i} g_{ij} \mathbb{1}[m_i = m_j] \mathbb{1}[r_{j,m,t} = r]}{D_i} \quad \text{for } r = \{1, 2, 3\} \quad (1.3)$$

and the model becomes

$$y_i = \sum_{r=1}^3 \beta^r \bar{m}_{-i}^r + \eta \ln(D_i) + \delta x_i + \epsilon_i \quad (1.4)$$

I follow the same strategy as the previous section. I use 2SLS, where the instruments are the number of entrants with the same occupation in a given tercile.

Table 1.5 reports the results of this exercise. Each column represents where the worker in question is situated the within occupation wage distribution. We look at

the effect of the share of coworkers with the same occupation in the 3 terciles in each column.

The results indicate that the effect of occupational homophily is generally strongest if coworkers rank higher than the worker in the wage distribution (with the exception of the bottom tercile). For the lowest ranked workers – the bottom tercile in Column (1)– the effect is negative and significant for all terciles, but strongest if coworkers are in the same tercile and those in the top tercile (coefficients of -0.64 and -0.53, respectively), compared to those in the middle. For the middle tercile in Column (2), the effect is still negative across all terciles, but significantly stronger when coworkers are in the top tercile. The effect in the top tercile is only significant for coworkers in the same tercile, suggesting that the probability of changing occupation of these workers is not affected by their coworkers with the same occupation who earn less than them.

Overall, these findings are in line with the mechanisms in the literature on information transmission outlined above; the effect of occupational homophily is larger when coworkers earn more than the worker in question.

Heterogeneity along other dimensions of homophily

We now turn to investigate whether the effect of occupational homophily on occupational change is different when workers have coworkers who share, not only the same occupation, but also the same demographic characteristics, such as gender, education, and age. As discussed in Section 1.2, the literature suggests that workers tend to associate with others with the same demographic characteristics (McPherson et al., 2001). A part of the effect of occupational homophily on occupational change could be attributed to, for example, occupational segregations. This would imply that workers interact with others with the same characteristics, who happen to hold the same occupation.

I follow a strategy similar to the previous section. I split the occupational homophily variable into two parts. The first is the number of coworkers with the same occupation who share the same characteristics as the worker divided by the number

Table 1.6: Occupational homophily and occupational change; heterogeneity along other dimensions of homophily

	Gender		Education		Age cohort	
	Blossfeld (1)	KldB (2)	Blossfeld (3)	KldB (4)	Blossfeld (5)	KldB (6)
Occupational homophily						
Same characteristic	-0.212*** (0.056)	-0.193*** (0.069)	-0.203*** (0.057)	-0.156** (0.068)	-0.334*** (0.071)	-0.294*** (0.086)
Different characteristic	-0.147** (0.059)	-0.084 (0.074)	-0.174*** (0.061)	-0.190** (0.078)	-0.138** (0.059)	-0.109 (0.074)
ln(degree)	-0.030** (0.015)	-0.047*** (0.015)	-0.030** (0.015)	-0.048*** (0.015)	-0.031** (0.015)	-0.048*** (0.015)
First stage F-stat	66.749	62.715	63.721	68.816	67.631	68.767
P-value of coefficient equality	0.111	0.044	0.565	0.553	0.009	0.043
Observations	41300	41300	41300	41300	41300	41300
Number of establishments	3303	3303	3303	3303	3303	3303

Notes: Dependent variable is the probability of changing occupation. Estimated with 2SLS following the specification in column (7) Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. For each column the sample is split according to the characteristic. Same characteristic (different characteristic) occupational homophily is number of coworkers in the same characteristic (different characteristic) divided by degree. The instrumental variables are constructed accordingly; the number of entrant coworkers in the same characteristic (different characteristic) with the same occupation. Age cohort is defined as coworker who are less than 5 years older or younger than the worker. Standard errors are robust and clustered at the establishment level (in parenthesis). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

of connections (degree), and the second is occupational homophily minus the first. I use 2SLS with two instruments; the number of entrants with the same occupation and characteristic, and the number of entrants with the same occupation but different characteristics.

Table 1.6 reports the results of this analysis. It is divided by gender, education, and age cohort, and estimated for both the Blossfeld and 3-digit KldB classifications.

The effect is generally stronger coworkers with the same occupation who are of the same age cohort (less than 5 years older or younger than the worker), and to a lesser extent, the same gender. As for the level of education, there is no statistically significant difference between coworker with the same level of education and those with a different level of education; the p-value of coefficient equality between coworkers with the same education and those with different education is large.

Additionally, the difference in the coefficients is generally stronger in the less aggregate KldB classification. The difference in the coefficients between same gender and

Table 1.7: Occupational homophily and Blossfeld occupational change; heterogeneity by gender and age

	Gender				Age group					
	Female		Male		24 - 35		36 - 45		46 - 55	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Occupational homophily										
All	-0.321***		-0.151***		-0.297***		-0.250***		-0.123**	
	(0.039)		(0.043)		(0.046)		(0.053)		(0.051)	
<i>By gender</i>										
Same gender		-0.368***		-0.105*						
		(0.043)		(0.054)						
Different gender		-0.191**		-0.232***						
		(0.084)		(0.054)						
<i>By age group</i>										
< 35					-0.388***		-0.302***		-0.161	
					(0.083)		(0.115)		(0.109)	
36-45					-0.377*		-0.021		0.162	
					(0.215)		(0.232)		(0.231)	
> 45					-0.414***		-0.747***		-0.448***	
					(0.144)		(0.152)		(0.124)	
ln(degree)	-0.018*	-0.021**	-0.007	-0.006	-0.001	-0.010	-0.017	-0.034***	-0.008	-0.023*
	(0.010)	(0.010)	(0.009)	(0.009)	(0.012)	(0.014)	(0.012)	(0.013)	(0.012)	(0.014)
First stage F-stat		523.5		450.0	932.8	276.744	731.2	217.6	661.6	273.4
Observations	18468	18468	22832	22832	16521	16521	11795	11795	12984	12984

Notes: Dependent variable is the probability of changing the Blossfeld occupation. The specification corresponds to that of column (5) of Panel A, Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The occupational homophily variable is split according to the characteristic. For each column, the sample is split according to the characteristics. Occupational homophily by characteristic is number of coworkers with the same occupation who belong to a given characteristic, divided by degree. The instrumental variables are constructed accordingly; the number of entrant coworkers with the characteristic in question. Robust standard errors are robust in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

different gender, and between same age cohort and different age cohort, is stronger for the KldB classification relative to the Blossfeld definition. In fact, the effect is not significant for coworkers with the same KldB occupation who are of a different age cohort or gender. On the other hand, while the effect of coworkers with different characteristics is weaker, it is still significant for the KldB classification. This means that occupational similarity is still a factor, even if workers do not share the same demographic characteristics.

The effect of occupational homophily on occupational change is higher for coworkers who share the same characteristics. This suggests that the other dimension of homophily (gender and age) play a role in driving that effect.

The effect by gender and age

I further desegregate the effect by the gender and age group of workers. The estimates are shown in Table 1.7. The odd numbered columns show the estimate of

the original occupational homophily, and the even numbered show occupational homophily split by the characteristic of the coworkers. The model is estimated for the Blossfeld definition. In terms of gender, the effect is significantly stronger for females compared to males (coefficients of -0.32 and -0.15 in Columns (1) and (3), respectively). In terms of coworker characteristics, the effect is mainly driven by female coworkers; the coefficient is higher for female coworkers with the same occupation compared to male coworkers for both males and females. This result aligns with Marin (2012), who found that women are more willing to share information about job opportunities than men.

In terms of the age dimension, the effect is decreasing with age group, as shown in Table 1.7. Specifically, the strongest effect is observed for workers between the ages of 24 and 35, with a coefficient of -0.29 in Column 5. The effect is slightly weaker for the 35-45 age group, with a coefficient of -0.25 (Column 7), and generally weak for the oldest age group between 46-55, with a coefficient of -0.12 (Column 9). The pattern of the effect varies depending on the interaction between the age group of the worker and their coworkers. For the youngest group, the effect appears to be homogeneous for all age groups, as shown in Column 6. However, for the middle group, the effect is strong and significant for coworkers in the oldest age group, substantially weaker for the youngest age group, and insignificant for coworkers in the same age group (Column 8). For the oldest age group, the effect is primarily driven by coworkers who are of the same age group (Column 9).

1.6 Conclusion

This study examined the relationship between occupational homophily, defined as the share of direct coworkers with the same occupation, and the probability of changing occupation among workers in Germany. Initial descriptive evidence from a large linked employer-employee administrative dataset showed that workers with higher occupational homophily are less likely to change occupation following an establishment change. This relationship could be explained by information transmission about

job opportunities or referrals between coworkers, as well as other factors related to the nature of the occupation, firm, or individual preferences. This paper had the objective of accounting for these factors, and providing an estimate of the effect that could be attributed to network interactions. After controlling for a large set of possible confounders and employing an instrumental variable strategy, a large and negative effect remains. A 10 percentage point increase in occupational homophily is associated with a 1.6 to 2 percentage point decrease in the probability of changing occupation. The effect is more pronounced for high skilled services and production jobs, if coworkers rank higher in the wage distribution within occupation, and when coworkers share other characteristics such as gender and age.

These results indicate that the type of occupation a coworker holds also matters for individual labour market outcomes. They show that workers who have many connections in their occupation will have a better access to future job opportunities in that occupation. These findings complement previous research on how coworkers affect labour market outcomes in Germany (Cornelissen et al., 2017; Glitz, 2017; Jarosch et al., 2021), and possibly in other countries (Cingano and Rosolia, 2012; Glitz and Vejlin, 2021; Saygin et al., 2021; Lengyel and Eriksson, 2017).

The paper has some limitations that should be taken into account when interpreting its results. One of the limitations is that the job search intensity might be endogenous, which means that those who have better networks search more intensely for a job and thus change occupation. This is a case of selection bias, which could affect the estimation of the causal effect of occupational homophily on occupational change. One possible solution is to use mass layoffs as an exogenous shock to job change (as in Cingano and Rosolia, 2012; Glitz and Vejlin, 2021; Saygin et al., 2021; Eliason et al., 2022). However, the question is whether we can generalise the estimates to, for example, job-to-job transitions.

Another limitation is that the paper only examines occupational change and not the more general occupational choice, i.e., it does not look at the effect of the share of

coworkers of a certain occupation, on the probability of switching to that occupation.²⁶ Occupational choice is a more complex problem as it involves modelling worker outcomes as a function of the average outcome of coworkers (endogenous peer effects), which is challenging to identify (see Manski, 1993; Bramoullé et al., 2009). In contrast, in the case of occupational change, the outcome of the worker in question is a function of an object that is not occupational change of their coworkers. Occupational homophily is considered an exogenous peer effect, which is simpler to identify.

Future research could provide a theoretical model that incorporates coworker network transmission mechanisms into a model of occupational choice. This can give a clearer picture of how workers obtain information through their networks, and how the type and quality of coworkers affects the kind of opportunities the worker will encounter in the future, and the implications of these interactions on the labour market as a whole. Finally, such model can help identify endogenous peer effects, and predict the occupational path of individuals based on the occupational composition of the coworkers they encounter along their employment spells.

²⁶ Although preliminary descriptive evidence from SIEED shows that, in most occupations, there is a positive effect the share of coworkers of a certain occupation, on the probability of switching to that occupation.

Chapter 2

Equilibrium Search with Endogenous Coworker Networks

2.1 Introduction

Empirical research has documented the importance of coworker networks - past and current - in job finding (e.g., Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021; Caldwell and Harmon, 2019). Workers who possess better and larger networks have a better chance of finding a job due to information transmission between coworker. There is a mechanical component to the formation of these networks. Coworker links are formed during the worker's employment spells and evolve by the movement of workers between firms.¹ The worker's network is larger in larger firms and in firms with a high churn. Many on-the-job search model (for example Burdett and Mortensen, 1998; Postel-Vinay and Robin, 2002; Cahuc et al., 2006) predict firm size dispersion due to firm productivity differentials or their wage setting strategies. The dispersion in firm size can lead to heterogeneity in coworker network size and thus may affect other job market outcomes.

This paper investigates the link between information transmission between coworkers, firm (network) size, wage, mobility, and match quality. I study an equilibrium search model with two sided heterogeneity and on-the-job search, where workers in larger firms search more efficiently for jobs when they are employed and when unemployed. I then estimate the model using German linked employer-employee

¹ The number of links formed in a given firm is number of coworkers plus the number of separations that occurred in the firm during the worker's spell.

data.

The model has three key components. First, firm productivity heterogeneity with on the-job-search (e.g., Burdett and Mortensen, 1998; Postel-Vinay and Robin, 2002; Cahuc et al., 2006). As in many such models, the model predicts that more productive firms are larger in equilibrium because they are better capable of attracting workers. Second, the model incorporates coworker networks as a social capital that is accumulated and dis-accumulated. Workers form connection when they join a firm and maintain them during periods of unemployment. These connections are then lost upon moving to a new firm while employed, and if an individual spends an extended period in unemployment. The model generates an ex-post heterogeneity in job finding rates in equilibrium as workers in larger firm search more efficiently. Third, the model features wage bargaining following Cahuc et al. (2006). The firm takes into account the wage growth opportunities networks provide the worker when the wage contracts are negotiated. The effect networks will have on wages depends on the worker bargaining power and the spread of the firm productivity distribution.

I estimate the model using administrative data from Germany using indirect inference (Gourieroux et al., 1993). I find that coworkers have a large effect on job-to-job mobility, but a modest effect on unemployment to employment mobility, wage dispersion, and match quality.

The paper contributes to several strands of literature. The first is the role of coworkers in the labour market. It follows the idea of labour flow networks (Guerrero and Axtell, 2013), where the links between coworkers and firms is generated by mobility. I link the formation of these networks to the behavioural job mobility decisions found in on-the-job search models (e.g., Postel-Vinay and Robin, 2002; Cahuc et al., 2006) to evaluate the role of these networks in broader job market outcomes. The empirical literature focuses on the effect of coworkers through information transmission (for example Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021) or other mechanisms such as peer effects (Cornelissen et al., 2017), learning in teams (Jarosch et al., 2021; Herkenhoff et al., 2024). I follow the former, where the size of the firm as a proxy for

coworker network size affect the job search efficiency of its employees.

In terms of theory, the closest study to this paper is Caldwell and Harmon (2019), who study how coworkers affect mobility and wage growth in an on the job-search-model. Past coworkers who come across a vacancy transmit the information to their links who can use it to renegotiate the wage or move to that post. While they explicitly modelled the information transmission process, they did not show how the network is formed and shaped by that process. In this paper, I derive closed form solutions to the shape of the network as the distribution of firm size. The shape of the network gives clues on how it was formed and how information transmission between affects job finding and mobility. Another related theoretical paper is Arbex et al. (2019). It is a wage posting on the job-search-model (following Burdett and Mortensen, 1998) where the position of the worker in a social network determines the job they draw from an offer distribution. While the social network structure is fixed in Arbex et al. (2019), in this paper, it is an equilibrium outcome generated by worker mobility.

Another strand is Arbex et al. (2019) and Mortensen and Vishwanath (1994), where workers can draw from different distributions depending on their social connections. These models consider referrals; the workers “pull” their connections into their firms. In this paper, I consider a “push” approach; worker joins a firm and relies on their connection in this firm to find better job, not necessarily in that firm. This is more consistent with models such as Calvo-Armengol and Jackson (2004), Calvó-Armengol and Zenou (2005), and Calvó-Armengol and Jackson (2007), where the worker transmits information about job opportunities they encountered in their job search, not necessarily in their own firm.

Finally, a feature of the model is that workers recall their coworkers from previous periods. In case of unemployment, their job finding rate is a function of the firm size of their previous employer. This relates to the notion of search capital (Carrillo-Tudela et al., 2011; Carrillo-Tudela and Smith, 2017), where workers can recall the past job offers they encountered, using these firm connections to negotiate wages in the future.

The rest of the paper is structured as follows. Section 2.2 presents the theoretical

model. Section 2.3 describes the estimation procedure of the model, the estimates, and quantifies the effect of coworker networks on wage dispersion and mobility. Section 2.4 discusses the model and presents some possible extension. Section 2.5 concludes.

2.2 Model

This section describes an equilibrium on-the-job search model, where workers form connections in the workplace that can improve their job search efficiency.

2.2.1 Environment

Time is continuous and goes forever. The economy is at the steady state. There is a unit mass of risk neutral workers and a unit mass of firms. Agents live forever and discount at a rate r . There is free entry. The economy is characterised by complete information, and the only frictions in the economy are search frictions.

Workers differ in ability ϵ , where ability is distributed according to the continuous cumulative distribution $H(\cdot)$. Workers can be employed or unemployed. The utility flow for unemployment is ϵz . Job search is sequential and random, and on-the-job search is allowed, and workers can search for jobs through their coworker networks or over other (general) channels.

The firm

Firms are heterogeneous with respect to productivity and the potential match between a worker and a firm produces ϵp units of goods, where p is the firm productivity. p is distributed according to a continuous cumulative distribution $F(\cdot)$. Matching is random so the sampling distribution of the firm is identical to the productivity distribution.

As the main focus of this paper is the role of coworker networks, the firm in this model is a unit characterised by productivity p , and a place where its employees have the possibility to interact. A firm is not necessarily a company, but could be a work-

place, a department, or a team. Empirically, the researcher can decide how to define a firm in terms of who is considered a coworker that influences the worker's outcomes. For instance, a firm can be an establishment, so coworkers are all the connections formed in that establishment. Alternatively, one can assume homophily, so firms are cells of employees working in the same establishment and sharing similar characteristics, such as occupation or level of education.²

This approach has limitations and advantages. The disadvantage of this approach is that it abstracts from production complementarities and the patterns of interactions between different types of workers in the same workplace. If the definition of the firm is too broad, the composition of the firm will have little impact on labour market outcomes. Conversely, if the definition is too narrow, the model assumes that only similar workers affect a given worker's outcome. The advantage, as we will see in the next sections, is that this approach can capture key aspects of coworkers networks (formed during employment spells, accumulated and lost), yet maps naturally to tractable on-the-job search models, such as Postel-Vinay and Robin (2002) and Cahuc et al. (2006).

Coworker networks

The individual firm is the centre of the notion of coworker networks. It is the location where the workers interact and share information. When a worker joins a firm, they form connections that can potentially increase their job search efficiency. The higher number of connections, the more efficient they are. When a worker loses their job, they keep their previous connections for a period of time.

Coworker connections are weak links; i.e., brief and in the short term (Granovetter, 1973). This has two implications. First, the value of coworkers lies in their large numbers. Even if the probability of one coworker having information about job opportunity is low, the probability of at least one coworker having information could be high when the network is large enough. Second, coworker links are not forever – un-

² The literature follows different approaches when it comes to how the notion of firm in the model corresponds to that in the data. For example, in Caldwell and Harmon (2019) and Herkenhoff et al. (2024), it corresponds to the establishment, whereas in Jarosch et al. (2021), it corresponds to both establishments and establishment-occupation cells.

like personal connections. These connections are lost if there are no circumstances for the workers to interact regularly.

This motivates a simplifying assumption; workers can keep track of coworkers from only one employer. An employed worker forgets their connections from their previous employer when they move to a new employer. Similarly, an unemployed worker loses their previous connections when they find a new job. This follows from the fact that when worker once moves firms, they will no more be in close proximity to the other workers.³ I further will discuss the implication when we relax this assumption in Section 2.4.2. The main advantage of this assumption is that the firm size summarises the worker's network size.⁴

When it comes to the network shape, this assumption means that network is disconnected. Each firm is an island where the workers inside the island are connected to each other but not to the outside world. The workers within each island will face the same structure of information about job opportunities in the outside world. This simplifies the model as the worker when formulating the strategy will only need to consider the firm productivity, size, and the aggregate state of economy. In turn, the workers' strategies and job search behaviour will influence the overall shape of the network.

Job search

Workers can search over regular channels or through their coworker networks. When a worker joins a firm p with the average size $l(p)$, they form connections they can use to search for a new job when they are unemployed and when they are employed. The efficiency of the search depends on the firm size. The arrival rate for the

³ Another approach is to assume that once a worker leaves a firm, the personal connections they made with coworkers become friendships, and assume that workers have different capacities to maintain these friendships (that could be correlated or uncorrelated with productive ability). Here, the workers will be heterogeneous with respect to the arrival rate through other channels. This will have interesting implications – e.g., sorting –, but in this paper I will assume that workers will have the same arrival rate through other channels.

⁴ This result is similar to limited observability (the worker cannot observe their peers state) and limited memory (the worker cannot recall the peers' past job offers) in Arbex et al. (2019) where the worker will only need to look at aggregates when they make these decisions.

employed in a firm p is $\xi(p) = \gamma + \gamma_s(p)$, where γ the job arrival rate through the regular channels and $\gamma_s(p)$ through social connections.

The Poisson arrival rate for the unemployed is $\chi(p)$ which depends on the size of their previous employer. When the job is destroyed – with an exogenous rate δ – the unemployed keeps the connections for a while. With an exogenous rate ϕ , the unemployed worker loses their connections and moves to a state where they can only search through regular channels the arrival rate is fixed at λ .

Wage setting

When a worker meets a firm, they bargain over the wage following Cahuc et al. (2006). The worker bargaining power depends on their employment status; let $\beta^u \in [0, 1]$ be the bargaining power of the unemployed workers and $\beta \in [0, 1]$ that of the employed. When β or $\beta^u = 0$ the firm holds all the bargaining power and extracts the entirety of the surplus of the match,⁵ and when β or $\beta^u = 1$ the worker holds all the bargaining power and the firm pays them their marginal productivity.

The wage setting rule for the employed is as follows. Let $V(\epsilon, w, p)$ be the value of employment at wage w in a firm with a productivity p . When a worker employed at firm p with wage w meets new a firm $\tilde{p}$ above a threshold marginal productivity $q(\epsilon, w, p)$, they bring the job offer to their employer and renegotiate their wage. The threshold productivity satisfies $V(\epsilon, w, p) = V(\epsilon, q(\epsilon, w, p), q(\epsilon, w, p))$.

Now, unlike Cahuc et al. (2006) it is not immediately clear if meeting $\tilde{p} > p$ induces a move to $\tilde{p}$. For example if the arrival rate $\tilde{p}$ is sufficiently lower than p , i.e., firm p is not as productive but provides better future employment opportunities through networks. Here, the maximum value of employment in the incumbent firm p given by $V(\epsilon, \epsilon p, p)$ is greater than $\tilde{p}$ given by $V(\epsilon, \epsilon \tilde{p}, \tilde{p})$. In this paper, I will focus only on the case where $V(\epsilon, \epsilon p, p)$ is increasing in p , thus, only a poaching firm with a higher productivity will induce mobility.⁶

⁵ In this case, the model collapses to the sequential auctions model (Postel-Vinay and Robin, 2002).

⁶ In Appendix B.1.1, I show that whenever $V(\epsilon, \epsilon p, p)$ is differentiable with respect to p , it must be increasing in p .

The new wage $R(\epsilon, \tilde{p}, p)$ is set according to the following rules

1. if $q(\epsilon, w, p) < \tilde{p} < p$ $V(\epsilon, R(\epsilon, \tilde{p}, p), p) = \beta V(\epsilon, \epsilon p, p) + (1 - \beta)V(\epsilon, \epsilon \tilde{p}, \tilde{p})$
2. if $\tilde{p} > p$ $V(\epsilon, R(\epsilon, p, \tilde{p}), \tilde{p}) = \beta V(\epsilon, \epsilon \tilde{p}, \tilde{p}) + (1 - \beta)V(\epsilon, \epsilon p, p)$

As for the unemployed, let $U(\epsilon, 0)$ be the value of unemployment of an unemployed workers with connections, and $U(\epsilon, p)$ be the value of unemployment of those who lost their jobs at firm p . Here, I will make a simplifying assumption; I assume that unemployed workers have no bargaining power and the firm captures all the surplus of the match, i.e., $\beta^u = 0$. So when an unemployed worker meets a firm $\tilde{p}$, it will set the wage as to equalise the value of employment at wage $R^u(\epsilon, 0, x)$ or $R^u(\epsilon, p, x)$ (without and with connections, respectively) and the value of unemployment. The wage setting rule for the unemployed is

1. $U(\epsilon, 0) = V(\epsilon, R^u(\epsilon, 0, \tilde{p}), \tilde{p})$
2. $U(\epsilon, \tilde{p}) = V(\epsilon, R^u(\epsilon, p, \tilde{p}), \tilde{p})$

I assume the unemployed workers do not have any bargaining power because if $\beta^u > 0$, and if the arrival rate for the unemployed is heterogeneous (as is the case in this paper), the value of unemployment will be heterogeneous. So, some workers will be selective and will not accept all the job offers they receive. Since the arrival rates at the steady state depend on the worker mobility decisions, and those depend on the value of unemployment, the complexity of the model will significantly increase. Setting $\beta^u = 0$ will insure that all unemployed workers will accept any job offer they encounter.

2.2.2 Worker problem

Equation 2.1 below describes the Bellman equation of the unemployed worker without connections.

$$rU(\epsilon, 0) = \epsilon z + \lambda \int_{\underline{p}}^{\bar{p}} V(\epsilon, R^u(\epsilon, 0, x), x) - U(\epsilon, 0) dF(x) \quad (2.1)$$

The value of unemployment without connections is the utility flow from unemployment ϵz plus the job arrival rate from search on the regular channels weighted by the sampling distribution of p times the surplus from the match; the difference between the value of employment at wage w in a firm with a productivity p and the value of unemployment. The reservation wage for a firm p is $R^u(\epsilon, 0, p)$ and satisfies $U(\epsilon, 0) = V(\epsilon, R^u(\epsilon, 0, p), p)$. Any firm they meet will offer $R^u(\epsilon, 0, p)$, which yields $U(\epsilon, 0) = \frac{\epsilon z}{r}$.

The value of unemployment of the unemployed who were employed in a firm p satisfies the Bellman Equation 2.2 bellow.

$$rU(\epsilon, p) = \epsilon z + \chi(p) \int_{\underline{p}}^{\bar{p}} V(\epsilon, R^u(\epsilon, p, x), x) - U(\epsilon, p) dF(x) + \phi(U(\epsilon, 0) - U(\epsilon, p)) \quad (2.2)$$

A worker with connections from firm p finds a job with a rate $\chi(p)$. If they do not, they stay unemployed and move to a state without connections with a probability ϕ . The reservation wage for a type p unemployed worker to type x firm is given by $R^u(\epsilon, p, x)$ and satisfies $U(\epsilon, p) = V(\epsilon, R^u(\epsilon, p, x), x)$. We get $U(\epsilon, p) = \frac{\epsilon z}{r} \forall p$. This means that the reservation wage for any firm for the unemployed is the same regardless the number of connection, conditional on ability. Then, we can set the value of unemployment for all workers $\bar{U}(\epsilon) = U(\epsilon, 0) = U(\epsilon, p)$ and the reservation wage from unemployment to firm x for all workers $R^u(\epsilon, p) = R^u(\epsilon, p, x) = R^u(\epsilon, 0, x) \forall x, p$.

Now we turn to the value of employment. The Bellman equation for the employed is

$$\begin{aligned} rV(\epsilon, w, p) = & w + \xi(p) \int_{q(\epsilon, w, p)}^p V(\epsilon, R(\epsilon, x, p), p) - V(\epsilon, w, p) dF(x) \\ & + \xi(p) \int_p^{\bar{p}} V(\epsilon, R(\epsilon, p, x), x) - V(\epsilon, w, p) dF(x) \\ & + \delta(\bar{U}(\epsilon) - V(\epsilon, w, p)) \end{aligned} \quad (2.3)$$

The value of employment at wage w depends and the wage, the value of employ-

ment in the same firm if they meet a less productive firm but higher than the threshold $q(\epsilon, w, p)$ (second term), and the value of employment if they meet a more productive firm (third).

We use the wage setting rules above, and integrating by parts the second and third terms⁷ on the right hand side, yields

$$\begin{aligned} (r + \delta)V(\epsilon, w, p) = & w + (1 - \beta)\xi(p) \int_{q(\epsilon, w, p)}^p V'(\epsilon, \epsilon x, x) \bar{F}(x) dx \\ & + \beta \xi(p) \int_p^{\bar{p}} V'(\epsilon, \epsilon x, x) \bar{F}(x) dx \\ & + \delta \bar{U}(\epsilon) \end{aligned} \quad (2.4)$$

Where $\bar{F}(x) = 1 - F(x)$. Now to derive $V(\epsilon, \epsilon p, p)$, set w to ϵp , yielding

$$(r + \delta)V(\epsilon, \epsilon p, p) = \epsilon p + \xi(p)\beta \int_p^{\bar{p}} V'(\epsilon, \epsilon x, x) \bar{F}(x) dx + \delta \bar{U}(\epsilon) \quad (2.5)$$

There exists a closed form expression for $V(\epsilon, w, p)$ and $V(\epsilon, \epsilon p, p)$ (derived in detail in appendix B.1.2) where

$$\int_p^{\bar{p}} V'(\epsilon, \epsilon x, x) \bar{F}(x) dx = \epsilon e^{-T(p)} \int_p^{\bar{p}} \frac{\bar{F}(x)}{r + \delta + \beta \xi(x) \bar{F}(x)} e^{T(x)} dx \quad (2.6)$$

with

$$T(p) = - \int_p^{\bar{p}} \frac{\beta \xi'(y) \bar{F}(y)}{r + \delta + \beta \xi(y) \bar{F}(y)} dy \quad (2.7)$$

Note that this implies that the value of employment at wage ϵp is proportional to ability; $V(\epsilon, \epsilon p, p) = \epsilon V(1, p, p)$. Using the wage setting rule, and Equations 2.4 and 2.5, we can derive the wage of a worker moving from p to $\tilde{p}$ in Equation 2.8 below

$$\begin{aligned} R(\epsilon, p, \tilde{p}) = & \epsilon \left(\beta \tilde{p} + (1 - \beta)p \right. \\ & - (1 - \beta)^2 \xi(\tilde{p}) \int_p^{\tilde{p}} V'(1, x, x) \bar{F}(x) dx \\ & \left. - \beta(1 - \beta)(\xi(\tilde{p}) - \xi(p)) \int_p^{\bar{p}} V'(1, x, x) \bar{F}(x) dx \right) \end{aligned} \quad (2.8)$$

⁷ I focus on the case where $\xi(p)$ is differentiable at $[p, \bar{p}]$, making $V(\epsilon, \epsilon p, p)$ also differentiable.

Here the wage $R(\epsilon, p, \tilde{p})$ has three components. First, the average weighted by the worker bargaining power of the productivity of the poaching and incumbent firms (the first two terms). The second line implies the worker is willing to accept a lower wage in exchange for future wage growth opportunities. They will take a larger cut the larger the network size when the arrival rate is increasing in firm size. The third line is the direct effect effect of coworkers on wages. It indicates the worker the amount of wage cuts the worker is willing to take to account for the differences in the wage growth opportunities between the incumbent and the poaching from.⁸

Finally, the lower support of the productivity distribution satisfies $V(\epsilon, \epsilon \underline{p}, \underline{p}) = \bar{U}(\epsilon)$ and is given by

$$\underline{p} = z - \beta \xi(\underline{p}) \int_{\underline{p}}^{\bar{p}} V'(1, x, x) \bar{F}(x) dx \quad (2.9)$$

and the wage from unemployment into a firm p satisfies $V(\epsilon, R^u(\epsilon, p), p) = \bar{U}(\epsilon)$ yielding

$$\begin{aligned} R^u(\epsilon, p) = \epsilon \left(z \right. \\ \left. - (1 - \beta) \xi(p) \int_{\underline{p}}^p V'(1, x, x) \bar{F}(x) dx - \right. \\ \left. \beta \xi(p) \int_p^{\bar{p}} V'(1, x, x) \bar{F}(x) dx \right) \end{aligned} \quad (2.10)$$

Unlike in Postel-Vinay and Robin (2002) and Cahuc et al. (2006), because the bargaining parameter depends on the employment status, the wage following unemployment will be different from that of that for those poached from the least productive firms (p).

In the next section, I derive the key quantities in the steady state, including the steady state arrival rates for the unemployed and employed.

⁸ For example, if the poaching firm is larger than the incumbent firm, thus $\xi(\tilde{p}) > \xi(p)$, then the worker might be willing to take larger wage cuts because the new firm offers them a higher wage growth rate.

2.2.3 Steady state

Let E be the employment rate at the steady state, u_0 the measure of unemployed workers with no connections, and u_s the measure of unemployed workers with connections. Let $e(\epsilon, p)$ be the share of employment with skill ϵ in firm p , and let $u(\epsilon, p)$ be the share of the unemployed with skill ϵ who lost their jobs in firm p . The next three equations characterise the steady state for $e(\epsilon, p)$, $u(\epsilon, p)$, and u_0 .

Equating the measure of workers exiting firm p to those entering firm p , we get

$$Ee(\epsilon, p) \left[\delta + \zeta(p)(1 - F(p)) \right] = f(p) \left[\lambda u_0 h(\epsilon) + u_s \int_{\underline{p}}^{\bar{p}} \chi(x) u(\epsilon, x) dx + E \int_{\underline{p}}^p \zeta(x) e(\epsilon, x) dx \right] \quad (2.11)$$

The left hand side is the outflow; the measure of workers of skill ϵ in firm p (i.e., $Ee(\epsilon, p)$) who lose their jobs with the rate δ or find a new job with a rate $\zeta(p)(1 - F(p))$. The inflow consists of the employed in the less productive firm and the unemployed who find a job in p with a probability $f(p)$.

The flows in and out of unemployment with connections p is

$$\delta Ee(\epsilon, p) = u(\epsilon, p) \left[\phi(1 - \chi(p)) + \chi(p) \right] u_s \quad (2.12)$$

The inflow is the workers employed who lose their jobs from firm p and the outflow is the unemployed who find a job, and those do not find a job and lose their connections.

The steady state for the unemployed without connections is given by

$$\lambda u_0 h(\epsilon) = \phi u_s \int_{\underline{p}}^{\bar{p}} (1 - \chi(x)) u(\epsilon, x) dx \quad (2.13)$$

Where the inflow is the stock of the unemployed without connections who find a job and the inflow is the unemployed with connections who do not find a job and lose their connections.

Integrating Equation 2.12 from $\underline{p}$ to $\bar{p}$, together with Equation 2.13, simplifies Equation 2.11 into Equation 2.14 below.

$$e(\epsilon, p) \left[\delta + \xi(p)(1 - F(p)) \right] = f(p) \left[\delta h(\epsilon) + \int_{\underline{p}}^p \xi(x) e(\epsilon, x) dx \right] \quad (2.14)$$

Note that a the solution to this equation must imply the absence of sorting, i.e., $e(\epsilon, p) = h(\epsilon)e(p)$. This because conditional on ϵ , workers face the same distribution of firms and thus the same entry and separation rates. Integrating over ϵ and solving the above integral equation gives Equation 2.15.

$$\delta F(p) = \delta \int_{\underline{p}}^p e(x) dx + (1 - F(p)) \int_{\underline{p}}^p \xi(x) e(x) dx \quad (2.15)$$

The solution to $u_s, u, E, e(x)$, and $u(x)$ depends on the assumption on the functional form of $\xi(p)$ and $\chi(p)$. In the next section, I will provide an example for $\xi(p)$ that is increasing and concave in the firm size $l(p)$, and gives a closed form solution to $e(p)$.

A closed form solution to the steady state employment

There exists a closed form solution to $e(p)$ if the arrival rate for the employed $\xi(p)$ is such that $\int_{\underline{p}}^p \xi(x) e(x) dx = a g(p) + b \int_{\underline{p}}^p e(x) dx$, where a and b are parameters, and $g(p)$ is a known function of p . Let us assume that that the arrival rates are weakly increasing and concave in the firm size, given by $E l(p) = E \frac{e(p)}{f(p)}$. Which is the measure of workers employed in p divided by the measure of firms with productivity p . We also need to guarantee that $\xi(p) > 0$. An example of this is given by Equation 2.16.

$$\xi(p) = \gamma + \alpha \left(1 - \frac{\eta}{l(p)} \right) = \gamma + \alpha \left(1 - \frac{\eta f(p)}{e(p)} \right) \quad (2.16)$$

α represents the intensity of the effect of network size on the arrival rates. Specifically, it is the change in the mass of workers in firm p finding a job through coworkers due to a change in the firm size.⁹ η is a parameter small enough to guarantee that

⁹ Let $\Gamma(l(p))$ be the mass of workers finding a job through coworkers in a firm with a size $l(p)$, i.e.,

$\frac{\eta}{l(p)} \leq 1 \forall p$ so that the arrival rate is always positive. If we set η to be equal to the average size of the least productive firm, then $\eta = \frac{\delta}{\delta+\gamma}$; the arrival rate for the least productive firm will be γ and positive for all other firms.¹⁰

Let $E(p)$ be the share of employed workers with productivity less than p , and given by $E(p) = \int_p^{\bar{p}} e(x) dx$. We use the expression for $\xi(p)$ to solve for $E(p)$ from Equation 2.15. Equation 2.15 below shows the expression for $E(p)$.

$$E(p) = \frac{F(p)(\delta + \alpha\eta) - \alpha\eta F(p)^2}{\delta + \bar{F}(p)(\gamma + \alpha)} \quad (2.17)$$

Where $\bar{F}(p) = 1 - F(p)$. We differentiate Equation 2.17 with respect to p and we get the $e(p)$ in Equation 2.18.

$$e(p) = f(p) \left[\frac{\delta(\delta + \gamma + \alpha + \alpha\eta(1 - 2F(p))) + \alpha\eta(\gamma + \alpha)\bar{F}(p)^2}{(\delta + \bar{F}(p)(\gamma + \alpha))^2} \right] \quad (2.18)$$

Given that $l(p) = \frac{e(p)}{f(p)}$, the values of $l(p)$ do not depend on the shape of the productivity distribution or the density; they are determined by the ranking of the firm in the productivity distribution. We also see that more productive firms are larger as $l'(p) > 0 \forall p$.¹¹

Now it is useful to express the firm size in terms of the firms position on the productivity distribution $\tilde{l}(x)$ – with a trivial change of variable in the equation above from $F(p)$ to x .

$$\tilde{l}(x) = \frac{\delta(\delta + \gamma + \alpha + \alpha\eta(1 - 2x)) + \alpha\eta(\gamma + \alpha)x^2}{(\delta + (1 - x)(\gamma + \alpha))^2} \quad (2.19)$$

$\Gamma(l(p)) = \gamma_s(l(p))l(p) = \alpha l(p) - \alpha\eta$. Thus the derivative with respect to firm size $\frac{\partial \Gamma(l(p))}{\partial l(p)} = \alpha$.

¹⁰ Note that this is related to a similar micro founded formulation, where the arrival rate is related to the probability of at least on coworker receiving a job offer through regular channels and transmitting it to a given worker. Let a be the probability of a coworker transmitting a job opportunity to a given worker conditional on receiving one, then the arrival rate of job offers through coworkers is $1 - e^{l(p)\log(1-\gamma(1-a))}$. Which is also increasing and concave in firm size. The downside is that there is no close form solution to the steady state measures. However, numerically, the model behaves in a similar manner as that with the chosen functional form above.

¹¹ $l'(p) = 2f(p)\delta(\alpha + \delta + \gamma) \frac{(\gamma + \alpha(1-\eta))}{(\delta + \bar{F}(p)(\gamma + \alpha))^3}$. If $\alpha > 0$ and given $\eta < 1$, $l'(p) > 0$.

As firms in equilibrium differ in firm size, the model will generate an ex-post heterogeneity in the job arrival rates $\xi(p)$ and $\chi(p)$; the effect of coworkers on search efficiency will be endogenously heterogeneous. More productive firms are larger, workers in larger firms have larger networks, and therefore, will search more efficiently when employed and when unemployed.

Wage distribution

As pointed in section 2.2.2, the starting wage from unemployment follows a different rule to that associated with renegotiation or job-to-job mobility. We therefore need to derive the wage distribution for each group separately. Let $g(\epsilon, p)$ be the share of workers with ability ϵ and work in a firm p who earn the initial wage from unemployment. Let $G(w | \epsilon, p)$ be the share of workers with skill ϵ in firm p who renegotiated or moved from a job to another and do not more than the wage w .

To compute $g(\epsilon, p)$, we equate the outflow and the inflow into that state

$$g(\epsilon, p)e(\epsilon, p)E(\delta + \xi(p)) = f(p)(\lambda u_0 h(\epsilon) + u_s \int_{\underline{p}}^{\bar{p}} \chi(x)u(\epsilon, x) dx) \quad (2.20)$$

Where the outflow is the measure of workers who earn the initial unemployment wage and lose their job or meet a firm, and the inflow is the unemployed who find a job at firm p .

By Equations 2.12 and 2.13, and given the absence of sorting implies that $g(\epsilon, p)$ is given by

$$g(\epsilon, p) = \frac{\delta}{l(p)(\delta + \xi(p))} \quad (2.21)$$

To compute $G(w | \epsilon, p)$, we equate the measure of workers entering and exiting the

state gives

$$G(w|\epsilon, p)(1 - g(\epsilon, p))Ee(\epsilon, p) \left[\delta + \xi(p)(1 - F(q(\epsilon, w, p))) \right] =$$

$$f(p) \left[Eg(\epsilon, p)e(\epsilon, p)\xi(p)F(q(\epsilon, w, p)) \right. \quad (2.22)$$

$$\left. + E \int_{\underline{p}}^{q(\epsilon, w, p)} \xi(x)e(\epsilon, x) dx \right]$$

Rearranging gives the wage distribution for workers who negotiated or move to p from employment;

$$G(w|\epsilon, p) = \frac{(\gamma + \alpha)(\delta + \xi(p))E(q(\epsilon, w, p)) + (\delta\xi(p) - \alpha\eta(\delta + \xi(p)))F(q(\epsilon, w, p))}{(1 - g(\epsilon, p))l(p)(\delta + \xi(p)\bar{F}(q(\epsilon, w, p)))(\delta + \xi(p))} \quad (2.23)$$

Note that the overall wage distribution in a firm is continuous except for a mass point corresponding to the workers earning the initial wage from unemployment.

Unemployment rate

Now we can solve for the employment rate in the steady state. We first integrate Equation 2.12 with respect to ϵ and rearranging gives the expression of the share of unemployed workers laid off from firm p ; $u(p)$ in the Equation 2.24 below.

$$u(p) = \frac{\delta e(p)E}{(\chi(p) + \phi(1 - \chi(p)))u_s} \quad (2.24)$$

Given $E + u_s + u_0 = 1$, integrating with respect to p from $\underline{p}$ to $\bar{p}$, and the expression for u_0 in Equation 2.13 yield the expression for the employment rate E in Equation 2.25. Let $\tilde{\chi}(x)$ be the arrival rate from unemployment expressed in terms of firm rank.

$$E = \frac{\lambda}{\lambda + \delta \int_0^1 \frac{\lambda + \phi(1 - \tilde{\chi}(x))}{\tilde{\chi}(x) + \phi(1 - \tilde{\chi}(x))} \tilde{l}(p) dx} \quad (2.25)$$

Finally, the measure of unemployed workers with connections is given by Equation

2.26.

$$u_s = \delta E \int_{\underline{p}}^{\bar{p}} \frac{1}{\chi(x) + \phi(1 - \chi(x))} dE(x) \quad (2.26)$$

The effect of coworker networks on the employment rate is governed by three factors. First, how well the unemployed workers keep track of their previous connections. Second, how strongly coworkers affect the unemployed job arrival rate. Third, indirectly through the effect of coworker networks job-to-job search efficiency on firm size thus the job arrival rate of the unemployed with connections.

2.3 Quantitative analysis

In this section, I estimate the model using German administrative data with indirect inference (Gourieroux et al., 1993). I will first present the data, then discuss the empirical strategy, the results, and finally I will present a counterfactual exercise where search through networks is not permitted to evaluate their effects on mobility, wage dispersion, and match quality.

2.3.1 Data

I use the dataset Sample of Integrated Employer-Employee Data (SIEED), published by the Institute for Employment Research (IAB) of the Federal Employment Agency (BA). It is a linked employer-employee administrative dataset containing information about wages, employment characteristics, and skill levels, establishment characteristics as well as establishment identification for a large sample of German workers between 1975 and 2018. These workers are sampled as follows; 1.5% of all establishments of Germany are chosen randomly, then, the whole employment histories of all the workers who have ever worked in these establishments are added (Schmidtlein et al., 2020).

I transform the dataset into a monthly worker panel spanning from 2012-2015. I consider a short time period because the model assumes a fixed productivity distribution for firms and a fixed ability for workers (e.g., no human capital accumulation).

Table 2.1: Summary statistics by education group; SIEED 2012-2015

	Low education	Medium education	High education
Number of workers	156,201	459,914	117,219
Number of employment spells	178,305	518,089	130,345
Number of non-employment spells	40,383	89,541	21,107
Number of establishments	26,529	49,307	21,659
Observations	2,698,411	11,079,875	2,806,607

I include only the observations in panel establishment where we observe the whole worker population. I restrict the analysis to workers between the ages of 19 and 64 who worked in establishments smaller than 100 workers. The latter restriction is common in empirical coworker networks literature (e.g., Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021) because the focus is on workplaces rather than firms per se. I conduct the analysis by level of education (low, medium, high).¹²

The worker is employed if they are observed in an establishment on a cut-off date (the first day of the month). They are considered non-employed if they are not observed. Here, I approximate unemployment with non-employment. I consider the workers who were not employed for less than three months and then returned to their previous employer as part of the same employment spell.

A job-to-job transition (J2J) occurs when a worker moves from an employer to another within 30 days of leaving the employer, and is considered an employment to non-employment (E2N) transition otherwise. For non-employment to employment transition (N2E), I consider the non-employment spells that follow a separation from a panel establishment. This is because later I will use the size of the previous employer as a determinant of unemployment duration.

2.3.2 Parameterisation and identification

In this section, I will present the set of parameters I estimate, the parametric assumptions, and discuss the identification process for the model. I will motivate the choice of estimation targets by deriving their closed-form expressions when possible.

¹² Table 2.1 presents the sample breakdown.

For cases where closed forms are not achievable, I will make heuristic arguments to link the relevant features of the data to the model.

Worker reallocation and firm size

The job destruction rate δ is identified by the employment to non-employment rate (E2N) in the data. I fix $\eta = \frac{\delta}{\delta + \gamma}$. To identify the job arrival rate through other channels and network strength – γ and α , respectively – I use the job-to-job transition rate (J2J) and the second and third moment of the firm size distribution. In the model, we can find a closed form to these measures. First, J2J is given by $J2J = \int_{\underline{p}}^{\bar{p}} e(x) \xi(x) \bar{F}(x) dx$ which can be expressed in terms of the model parameters

$$J2J = \frac{\delta\gamma(\gamma + \delta + \alpha)}{(\gamma + \delta)(\gamma + \alpha)^2} \left[(\gamma + \delta + \alpha) \log\left(\frac{\gamma + \delta + \alpha}{\delta}\right) - (\gamma + \alpha) \right] \quad (2.27)$$

As for firm size, I fix the ratio of the number of workers (employed and non-employed) by the number of firms to match the data. Further, I normalise the firm size by the average firm size, variance of the normalised firm size is simply $Var(l(p)) = \int_{\underline{p}}^{\bar{p}} (l(x) - 1)^2 f(x) dx$ and the solution yields.

$$Var(l(p)) = \frac{\gamma^2(\gamma + \delta + \alpha)}{3\delta(\gamma + \delta)^2} \quad (2.28)$$

I target the coefficient of variation of the average firm size in the data. Similarly, we can find a closed form to the skewness of the average firm size in the model

$$Sk(l(p)) = \frac{\sqrt{27}}{5} \frac{(\gamma + \alpha)}{\sqrt{\delta(\gamma + \alpha + \delta)}} \quad (2.29)$$

Note that I over-identify these parameters; since two of the three moments above can fully identify them. As discussed above $J2J$, $Var(l(p))$, and $Sk(l(p))$ are all increasing in α and γ .¹³

¹³ The reason I did not use the coefficient of the regression of J2J probability on firm size to identify α is that α does not monotonically change the slope of the relationship between the two variables under the present parametric assumption on $\xi(p)$. Instead, it changes its curvature.

As for non-employment, I assume that the unemployed keeps their connections until they find a new job and fix $\phi = 0$. For simplicity, I assume that network size increases (potentially) the efficiency of search through other channels multiplicatively, taking the form

$$\chi(p) = \lambda \left(\frac{EN}{B} l(p) \right)^{\alpha_u} \quad (2.30)$$

Where λ is the arrival rate through other channels for the unemployed, α_u is the scale of the effect of networks, where N is the number of workers, B is the number of firms (i.e., $\frac{EN}{B} l(p)$ is the firm size when the measures of workers and firms are not 1, recall $\frac{N}{B} = 1$ in the model). I identify λ and α_u using the non-employment-to-employment rate (N2E) in the model and the coefficient $\hat{\kappa}$ from the following regression.

$$\log(y_{i,\tau}) = \kappa \ln(S_{j(i,\tau)}) + \kappa_1 X_i + \kappa_2 B_{j(i,\tau)} + v_{i\tau} \quad (2.31)$$

Where $y_{i,\tau}$ is the number of months spent in non-employment by worker i in spell τ . $S_{j(i,\tau)}$ is the average firm size of the previous employer (firm $j(i, \tau)$). Note that the coefficient κ is decreasing in α_u . X_i are worker characteristics (age, age squared, gender, and whether German), and $B_{j(i,\tau)}$ are the characteristics of the previous employer (location, industry, and average wage).

Productivity and ability distribution

Firm productivity is distributed according to a beta distribution, where productivity $p(\theta_p) = \underline{p} + \psi \theta_p$, and $\theta_p \sim \text{Beta}(\theta_1, \theta_2)$. Where θ_1 and θ_2 are shape parameters, ψ is the scale parameter and $\underline{p}$ is the location parameter (which satisfies Equation 2.9). Worker ability distribution follows a lognormal distribution with a mean 0 and standard deviation σ_ϵ .

I use the wage information to identify these parameters. I will include the first four moments of the log-daily wage distribution (mean, standard deviation, skewness, and Kurtosis) following Burdett et al. (2016).¹⁴ Additionally, I will incorporate another

¹⁴ An alternative choice to identify the productivity and ability parameters are to use an AKM (Abowd et al., 1999) like regression or log wage regression outcomes as in (e.g., Bagger et al., 2014; Bagger and

Table 2.2: Parameters and moments

Parameters		Moments
Job destruction rate	δ	Average E2N
Employed job search		
Employed job arrival rate through other channels	γ	Average J2J
Employed network strength	α	SD firm size Skewness firm size
Unemployed job search		
Unemployed job arrival rate through other channels	λ	Average N2E
Unemployed network strength	α_u	Coefficient of the regression of the duration of unemployment on $\ln(\text{firm size})$ ($\hat{\kappa}$)
Productivity and ability distributions		Wage information
Productivity shape parameter 1	θ_1	Mean log wage
Productivity shape parameter 2	θ_2	SD log wage
Productivity scale parameter	ψ	Skewness log wage
Standard deviation of ability distribution	σ_ϵ	Kurtosis log wage Mean log wage-min log wage ratio
Worker bargaining power	β	Average wage growth Average wage growth following J2J
Utility flow from unemployment	z	Mean log wage following N2E SD log wage following N2E

measure of wage dispersion: the *mean log wage-min log wage ratio*. This is calculated as the mean log wage divided by the mean of the log of the wages of the bottom 5% of the wage distribution.¹⁵ The worker bargaining power β is informed by the average wage growth and average wage growth following J2J. Conditional on the other parameters, a higher worker bargaining power indicates that workers capture more rent after encountering a new firm, leading to higher wage growth.

Since the initial wage following a N2E follows a different distribution, I include the first and second moments (mean and SD) of log initial wage following a N2E to identify the utility flow from unemployment z .

Table 2.2 summarises the set of parameters to estimate, and the corresponding moments to identify them.

Lentz, 2019; Burdett et al., 2020). Another alternative that could be considered is within and between firm variances as moments to identify the productivity parameters.

¹⁵ Bagger and Lentz (2019) use the average of the bottom 5% as the minimum to calculate the min-mean ratio.

2.3.3 Estimation

I estimate jointly the 11×1 vector of parameters P using indirect inference. The parameters are estimated at a monthly frequency, and I focus on the steady state. The vector of estimated parameters $\hat{P}$ minimise the weighted square difference between a set of empirical statistics (moments) and their simulated counterpart; and given by

$$\hat{P} = \underset{P}{\operatorname{argmin}} (M_e - M_s(P))' W (M_e - M_s(P)) \quad (2.32)$$

M_e is a 15×1 vector of empirical moments and $M_s(P)$ is a 15×1 vector of simulated moments. The 15×15 weight matrix W is a matrix with the inverse of the moment variance from the data on the diagonal and zero off diagonal. The variances are bootstrapped from the data. Additional information on the estimation procedure is found in Appendix B.2.

2.3.4 Results

Table 2.3 shows the value of the targeted moments in the data against those in the estimated model, and Table 2.4 shows the parameter estimates. The model fits the data generally well. In the next subsections, I will discuss the moment and parameter values.

Flows and firm size

The model produces estimates close to that for the flow and firm size moments for all education groups. We see that the parameter for network effect α is highest for the low education group (at 0.42) followed by the high education group (at 0.19), and finally the medium education group (at 0.0877). This however, does not give an idea on how networks affect mobility since the other flow parameters and the firm productivity play a role. The effect of networks on J2J mobility is composed of a direct effect on the job finding rate, and the indirect effect on change the firm size distribution (i.e., the shape of the network). We will discuss the indirect effect in the next section.

Table 2.3: Model fit

Moment	Low education		Medium education		High education	
	Data	Model	Data	Model	Data	Model
N2E	0.0949 (0.0005)	0.0786	0.1261 (0.0004)	0.1090	0.1174 (0.0008)	0.1189
E2N	0.0178 (0.0001)	0.0119	0.0088 (0.0000)	0.0063	0.0082 (0.0001)	0.0077
J2J	0.0142 (0.0001)	0.0186	0.0135 (0.0000)	0.0099	0.0131 (0.0001)	0.0137
Duration of unemployment regression coefficient ($\hat{\kappa}$)	-0.0102 (0.0038)	-0.0272	-0.0217 (0.0029)	-0.0203	-0.0154 (0.0049)	-0.0156
SD firm size	2.0334 (0.0004)	2.0309	1.7561 (0.0002)	1.7632	2.1865 (0.0005)	2.1645
Skewness firm size	6.1483 (0.0015)	6.1314	4.1079 (0.0005)	4.0748	5.2697 (0.0014)	5.2849
Mean log wage	3.1819 (0.0007)	3.1789	3.7012 (0.0003)	3.6910	4.3109 (0.0007)	4.2782
SD log wage	1.0648 (0.0005)	1.0693	1.0619 (0.0003)	1.0471	1.1470 (0.0006)	1.1134
Skewness log wage	-0.4578 (0.0019)	-0.0684	-0.9578 (0.0009)	-0.4043	-0.9951 (0.0020)	-0.1997
Kurtosis log wage	3.0267 (0.0077)	3.3643	3.5817 (0.0044)	4.6343	3.9795 (0.0111)	4.8053
Mean log wage-min log wage ratio	4.0598 (0.0110)	3.3863	3.4688 (0.0034)	2.7502	3.0153 (0.0060)	2.2453
Mean log wage N2E	2.8661 (0.0046)	2.7467	3.3519 (0.0035)	2.6486	3.7198 (0.0081)	3.5062
SD log wage N2E	0.9866 (0.0035)	1.3026	1.0683 (0.0024)	1.3932	1.1877 (0.0059)	1.4515
Average wage growth	0.0096 (0.0001)	0.0051	0.0109 (0.0001)	0.0066	0.0088 (0.0001)	0.0060
Average wage growth J2J	0.4546 (0.0073)	0.0441	0.6600 (0.0038)	0.1781	0.5227 (0.0076)	0.0954

Notes: Bootstrapped standard errors in parenthesis.

The direct effect of networks on J2J mobility is significant. The contribution networks to J2J mobility, keeping the firm size constant,¹⁶ is 79% for low education, 64.2% for medium education and 73.4% for high education.

The effect of network search on unemployment duration and N2E mobility is somewhat limited for all education group. The data and model give estimates between -0.0156 and -0.0272 for the regression coefficient $\hat{\kappa}$. This is somewhat similar to those found in Cingano and Rosolia (2012) for the relationship between network size and the duration of unemployment for displaced workers in Italy (coefficients between -0.02 and -0.06). This gives a similar picture as those in Glitz (2017) for Germany and

¹⁶ The contribution is the one minus the mass of workers finding a job through other means divided by J2J and given by $\frac{\gamma \int_0^1 (1-x) \tilde{l}(x) dx}{J2J}$.

Table 2.4: Parameter values

	Low education	Medium education	High education
γ	0.0154 (0.0000)	0.0170 (0.0000)	0.0177 (0.0002)
α	0.4172 (0.0020)	0.0877 (0.0004)	0.1906 (0.0031)
δ	0.0118 (0.0001)	0.0065 (0.0000)	0.0077 (0.0000)
λ	0.0729 (0.0006)	0.1020 (0.0006)	0.1133 (0.0019)
α_u	0.0329 (0.0011)	0.0241 (0.0012)	0.0207 (0.0040)
θ_1	4.1233 (0.1626)	1.1958 (0.0157)	6.9050 (0.4099)
θ_2	0.6640 (0.0200)	4.9368 (0.0280)	4.9117 (0.0559)
σ_ϵ	1.0598 (0.0006)	0.9433 (0.0005)	1.0815 (0.0013)
ψ	0.8960 (0.0184)	4.2670 (0.0063)	3.6628 (0.0560)
z	3.7386 (0.0012)	4.0237 (0.0016)	4.7979 (0.0088)
β	0.0821 (0.0007)	0.4715 (0.0020)	0.4704 (0.0137)

Notes: Standard errors in parenthesis. Further details on the standard error computation are found in Appendix B.2.

Saygin et al. (2021) for Austria.¹⁷ The direct effect of networks to N2E mobility (given by $1 - \frac{\lambda}{N2E}$) is higher for the low education group. It is 7.2%, 6.4% and 4.7% for the low, medium, high education groups, respectively.

Wages

For the three education groups, the model estimates match very well the mean and standard deviation of log wage, but somehow underestimates the skewness of log wage. It also produces a slightly lower estimate of the mean log wage-min log wage ratio than suggested by the data.

The model allows for the wage variance decomposition, similar to Postel-Vinay and Robin (2002). The variance of the wage is decomposed into the person effect $V[\ln(\epsilon)]$, search frictions $EV[\ln(w)|p]$, and the firm effects $VE[\ln(w)|p]$. The results

¹⁷ The latter two papers have the U2E probability as the dependant variable not the duration of unemployment.

Table 2.5: Variance decomposition

	Person effect $V[\ln(\epsilon)]$		Search frictions $EV[\ln(w) p]$		Firm effect $VE[\ln(w) p]$	
	Value	% of $V[\ln(w)]$	Value	% of $V[\ln(w)]$	Value	% of $V[\ln(w)]$
Low education	1.114	98.614	0.014	1.262	0.001	0.124
Medium Education	0.883	88.989	0.079	7.940	0.030	3.072
High Education	1.160	96.005	0.038	3.149	0.010	0.846

are shown in Table 2.5. The estimates give a similar picture for the contribution of each component for all levels of education.

Wage dispersion is mainly explained by individual ability. For low education, individual ability explains 98.6% of the wage variance, the 89% for medium education and 96% for high education.¹⁸ We also see that the firm productivity distribution is narrow. This means that output (i.e., ϵp) distribution and wage distribution are very similar (see Figure B.1). Frictional wage distribution ($EV[\ln(w)|p]$) is small, representing at most 7.9% of the log wage variance for the medium education group. This implies that, even having a larger number of coworkers improves search efficiency, the effect on wages is minimal.

Note that the model underestimates wage growth, especially following a job-to-job move. Wage growth is low given that the productivity distribution. Generally, a higher value for the worker bargaining power β could lead to higher wage growth, but we see that the value is especially low for the low education group at 0.08, and moderate for the medium and high education groups, at 0.47. It is possible that because the model does not feature other sources a wage growth, such as human capital accumulation (as in, for example Burdett et al., 2011; Bagger et al., 2014; Lise and Postel-Vinay, 2020), it does not succeed in reproducing the wage growth values in the data.

¹⁸ Jarosch (2023) found a similar pattern (narrow firm productivity spread) using similar data in Germany.

Table 2.6: Counterfactual exercise; no search through coworkers when employed

	Low education			Medium education			High education		
	SIM	CFL	% Δ	SIM	CFL	% Δ	SIM	CFL	% Δ
Wages									
Average log wage	3.179	3.181	0.054	3.691	3.669	-0.590	4.278	4.269	-0.210
Var log wage	1.143	1.116	-2.366	1.097	0.921	-16.050	1.240	1.172	-5.512
Mean log wage-min log wage ratio	3.386	3.194	-5.668	2.750	2.175	-20.919	2.245	2.101	-6.435
Average wage growth	0.005	0.001	-89.748	0.007	0.002	-77.005	0.006	0.001	-81.188
Mobility									
N2E	0.079	0.077	-1.862	0.109	0.108	-1.297	0.119	0.117	-1.400
J2J	0.019	0.006	-69.798	0.010	0.005	-49.588	0.014	0.006	-59.678
Employment rate	0.869	0.867	-0.250	0.944	0.944	-0.070	0.939	0.938	-0.084
Firm size									
SD firm size	2.031	0.630	-68.967	1.763	0.856	-51.434	2.165	0.838	-61.310
Skewness firm size	6.131	1.045	-82.959	4.075	1.498	-63.249	5.285	1.436	-72.834
Match quality									
Average log output	3.220	3.219	-0.030	3.865	3.799	-1.718	4.356	4.341	-0.338

Notes: SIM is simulation and CFL is counterfactual.

2.3.5 Implications on mobility and wage dispersion

In this section, I run a counterfactual exercise where I set the parameter α to zero, meaning that workers cannot search through their networks when they are employed. I then discuss the effect on mobility, dispersion and match quality. The results of this exercise is shown in Table 2.6. The first column of each group (SIM for simulation) is the base case, the second column is the counterfactual case (CFL), and the third column is the percentage difference.

In the model, setting α to zero will make the firm size distribution less dispersed, a significantly lower SD and skewness. This is because less productive firms become better able to retain workers and more productive firms become less able to attract workers. Here, the more productive firms are smaller in the base case, leading to lower dispersion.

The effect of coworkers on N2E mobility in this case will only go through the effect on the firm size distribution. We see the effect N2E mobility and employment is small. This is because the effect of firm size on N2E, governed by α_u , is small. However, the effect on J2J mobility is significant, setting α to zero reduces the average J2J in the economy by 69.8% for low education, 49.6% for medium education, and 59.7% for high education.

As for wages, the effect of coworkers is minimal for all groups because firm productivity distribution is narrow and the most of wage dispersion is due to individual ability. While the level of wage growth as predicted by the model is small, coworkers will have a significant effect on wage growth. The average wage growth in the counterfactual exercise is significantly lower than the base case by 89.7%, 77.0%, and 81.2% for the low, medium, and high education groups, respectively.

In theory, information transmission through coworkers should improve match quality (measured by the mean log output per worker) by channelling workers from less productive firms to more productive ones. However, since productivity differences between firms are small, the effect on match quality is minimal.

2.4 Discussion and extensions

This paper makes two main assumptions on the relationship between coworker network size and search efficiency. The first is that the job arrival rate depends only on the firm size and the second is homogeneous social ability (workers with the same number of connections have equal job arrival rates). In this section, I will discuss and present variants of the model where I relax these assumptions. I will first, derive the law of motion of the number of connections with tenure. Then, I will show how that affects wage determination in a sequential auction framework. Finally, I will discuss the impact of heterogeneous social ability and how it leads to sorting.

2.4.1 Social capital accumulation

This paper assumes that search efficiency depends on the firm's average size as a proxy for the worker's network size. The benefit of this approach is that the model is stationary, and that we only need to take into account the incumbent and poaching firm sizes which significantly simplifies the computation of the model. However, when a worker joins a firm, they continuously form new coworkers because of churn. The empirical literature stresses the importance of past coworkers in job finding. Addition-

ally, from an empirical strategy point of view, it allows for within firm identification because network size will be heterogeneous within a given firm.

The evolution of network size with tenure

Consider a firm at its steady state size $l(p)$. At the steady state, the firm size is constant as the inflow and outflow into the firm are equal. This means we can find a simple expression of network size in terms of firm size. Here, the network size is the number of current coworkers plus the number of past coworkers. Past coworkers are those who separated from the firm during the given worker's tenure.

Formally, Let $N(p, \tau)$ be the measure of links (network size) a worker in a firm p with tenure τ possesses. Let $\tilde{N}(p, \tau)$ be the measure of past coworkers, so the network size is $N(p, \tau) = l(p) + \tilde{N}(p, \tau)$. Let $J(p)$ be the separation rate (J2J or E2N separation), and ϕ be the rate at which the worker loses the link with past coworkers. The law of motion of measure of connected past coworkers is

$$\frac{\partial \tilde{N}(p, \tau)}{\partial \tau} = -\phi \tilde{N}(p, \tau) + (1 - \phi)J(p)l(p) \quad (2.33)$$

With the initial condition being that $N(p, 0) = 0$, i.e., workers who just moved to the firm do not have any past coworkers. The solution to the differential equation above is

$$\tilde{N}(p, \tau) = J(p)l(p) \frac{1 - \phi}{\phi} (1 - e^{-\phi \tau}) \quad (2.34)$$

The change in the measure of past coworkers over time is the inflow which is the measure of separations who stay in contact minus the outflow which is the measure of past coworkers who lose contact. The size of a worker's network will be

$$N(p, \tau) = l(p) \left[1 + J(p) \frac{1 - \phi}{\phi} (1 - e^{-\phi \tau}) \right] \quad (2.35)$$

Network size is increasing in the firm size, tenure, and the separation rate, all else held constant. Now, the alternative to main approach of the paper is to have the job

arrival rate $\xi(p, \tau)$ directly depending on the network size rather than the firm size. Crucially, the job arrival rate within the same firm will be heterogeneous. The key measures (firm size, network size, employment, etc) can be then solved for following a similar approach to that in Section 2.2.3 (which will be left for future work).

Within firm social capital accumulation and wage determination

Now I turn to the case where the number of coworkers a worker has increases with tenure leading to an increase in the job arrival rate when employed over time. I assume that workers lose connections with workers from past employers as soon as they join a new employer. So the arrival rate depends on the connections they formed while employed at the current employer (past or current).

For simplicity, I consider the sequential auction case where the firm holds all the bargaining power (i.e., $\beta = 0$) following Postel-Vinay and Robin (2002). In addition, I assume workers are homogeneous with respect to ability. The value of employment of a worker with wage w in firm p and tenure τ is $V(w, p, \tau)$ and satisfies

$$(r + \delta)V(w, p, \tau) = w + \xi(p, \tau) \int_{q(w, p, \tau)}^p \bar{F}(x) \frac{\partial V(x, x, \tau)}{\partial x} dx + \delta U + \frac{\partial V(w, p, \tau)}{\partial \tau} \quad (2.36)$$

The value of employment depends on the future growth opportunities that evolves with tenure as they accumulate more social capital.¹⁹

This differential equation does not have a convenient closed form solution as in, for example, Postel-Vinay and Robin (2002) or the base model in this paper. It can however be solved numerically by assuming stationarity after some large $\tau > T$, then solved backward.

Now this model is not qualitatively very different from main model; as workers are willing to take larger cuts if they work in larger firms. The main difference is that the reservation wage will depend on the tenure, thus the network size, in the poaching firm (then $\tau = 0$), or if wages are renegotiated, it will depend on the timing of the meeting.

¹⁹ Note that because every firm the worker meets will extract all the surplus, the maximum value a worker can get in firm p does not depend on tenure. Thus, $V(p, p, \tau) = \frac{p + \delta U}{r + \delta}$.

The advantage of this version of the model is that it is closer to how coworker networks are represented in the empirical literature (which include past and current coworkers). It also allows for within firm identification of the direct effect of coworkers on mobility and wage growth.

2.4.2 Social ability and sorting

So far, I assumed that social ability plays no role in job finding. The implication is that, conditional on firm characteristics, workers will have the same job finding rates and leading to the absence of sorting. More formally, the share of workers in firm p with ability ϵ will be the product of the share of workers in firm p and the share of workers with ability ϵ ; i.e., $e(\epsilon, p) = e(p)h(\epsilon)$. What if workers differ in their capacity to use coworker networks to search for jobs? What are the implications in terms sorting? In this section I will provide a simple example.

For simplicity, I assume workers have no bargaining power (workers will always move to more productive firms), that firm size only affect job search for employed workers, and I focus on the case with homogeneous production ability. Let s to be the worker's social ability and $\zeta(p, s)$ is the job arrival rate at firm p for a worker with social ability s . Let $h_s(s)$ be the share of workers with social ability s .

Now the process of deriving the equilibrium firm size and arrival rate is similar to that in Section 2.2.3. For brevity I will move to the final expressions. The nature of sorting will depend on how the job arrival rate depends on the firm size given ability. For example, it could depend on the total number of coworkers or the number of coworkers in the same group (homophily).

I will consider the latter, where the job finding rate depends on the measure of coworkers with the same social ability. This will give a simple closed form solution the key features of the model. The expression to the job arrival rate is similar to Equation 2.16 and is given by

$$\zeta(s, p) = \gamma + \alpha_s \left(1 - \frac{\eta_s f(p)}{e(s, p)} \right) \quad (2.37)$$

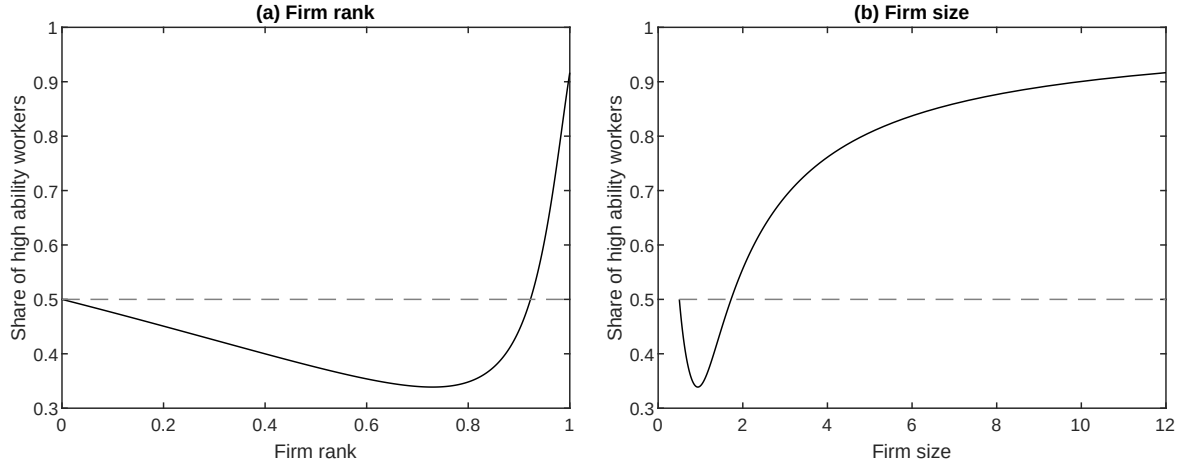


Figure 2.1: Share of high ability workers in employment by firm rank and firm size

Notes: The parameters are $\alpha_1 = 0.4$, $\gamma = 0.01$, $\delta = 0.01$, $\eta = \frac{\delta}{\delta + \gamma}$.

Where $\frac{e(s,p)}{f(p)}$ is the average measure of coworkers with ability s in firm p , α_s is the strength of the effect of firm on the job finding for a worker with ability s . Solving for the steady state leads to an expression for $e(p, s)$ that is very similar to that in Equation 2.18

$$e(p, s) = f(p) \left[\frac{\delta \left(h(s)(\delta + \gamma + \alpha_s) + \alpha_s \eta_s (1 - 2F(p)) \right) + \alpha_s \eta_s (\gamma + \alpha_s) \bar{F}(p)^2}{\left(\delta + \bar{F}(p)(\gamma + \alpha_s) \right)^2} \right] \quad (2.38)$$

Now it is clear sorting exists, as $e(s, p) \neq h(s)e(p)$. Some types of workers will be over or under represented in some types of firms. To showcase it, let us consider an example where there are two types of workers; half of the workers can use networks ($\alpha_1 > 0$) and the other cannot ($\alpha_0 = 0$).

Let us consider the share of workers who can use networks in employment in firm p ; $\frac{e(1,p)}{e(p)}$. Figure 2.1 shows the relationship between the share of high ability workers and firm size and rank. We see that the relationship has a U shape; sorting is non-monotonic. The share of high ability workers is initially decreasing in firm rank up until a point where it increases. High ability workers are generally under represented in smaller and medium firms and over represented in larger firms and more productive

firms. The mechanism is that for small and medium firms, as the level of productivity increases, the firm size increases and thus the job arrival rate for high ability workers. Given that they are relatively low on the productivity ladder, high ability workers are more likely to get poached by more productive firms than low ability workers. Therefore, the share of high ability workers will be low, and decrease with the level of productivity. As the level of productivity reaches a certain threshold, firms will be better able to keep the workers they have, and given that they are more likely to poach high ability workers from below, the share of workers in more productive firms will be large.

This will have an implication on wages via two channels. First, through the job arrival rate as it determines the amount of wage cuts the worker is willing to take. Second, the case productive ability and social ability are heterogeneous. The distribution of productive ability within and between firms will depend on the sign and strength of correlation between the productive and social ability and will be an additional determinant of wage dispersion.

2.5 Conclusion

This paper studied the link between coworker networks, firm size, and labour market outcome. I developed an on-the-job search model where workers in larger firms search more efficiently for jobs when employed and when unemployed. The model predicts that in equilibrium, more productive firms are larger and workers in larger firms search more efficiently for jobs. This links to wage bargaining as workers are willing to take larger wage cuts in exchange for better wage growth opportunities through their coworker networks. An empirical exercise on administrative data for Germany showed that coworkers play an important role in worker mobility, but limited effects on wages, mainly because generally firms in Germany are very similar in their level of productivity.

This paper has some shortcomings that can be addressed in future work. For example, that workers can keep track of one employer, implying that the effect of networks

can be summarised by the average firm size. While this assumption simplifies the model significantly, it falls short capturing the entirety coworker networks. Empirical literature (e.g., Cingano and Rosolia, 2012; Saygin et al., 2021; Glitz and Vejlin, 2021) often defines coworkers as the total number of worker has encountered in a given time period. That includes current coworkers, past coworkers from the current employer and previous employers. This allows for within firm identification as workers in the same firm will differ in their network sizes. Incorporating this definition in the model opens the possibility of linking previous empirical work on the subject to the model mechanisms. Some of these mechanisms were discussed in the previous section, but future work can delve further into the formation of coworker networks, and the implications of alternative definitions and assumptions on what constitutes a coworker.

The model abstracts from several interesting mechanisms, such as the process behind information transmission between coworkers (e.g., Calvo-Armengol and Jackson, 2004; Calvó-Armengol and Zenou, 2005; Calvó-Armengol and Jackson, 2007), peer effects (Cornelissen et al., 2017) and learning from coworker (Jarosch et al., 2021; Herkenhoff et al., 2024). It, however, provides insights into the implications of information transmission on coworker network formation, mobility, and wage dispersion.

Chapter 3

Coworker Networks and Occupational Allocation

3.1 Introduction

The type of job or occupation a worker holds is determined by their ability to perform the job, and their access to it. In the classical Roy (1951) model, workers self-select into occupations based on their comparative and absolute advantages. On the other hand, in an environment characterised by search frictions, the allocation of workers to occupations will also depend on their capacity to find such jobs. As argued in Chapter 1, coworker networks are an important mechanisms with which workers find jobs, and the types of jobs found through networks may depend on the characteristics of these networks.

This paper analyses the implications of the composition of coworker networks on occupational choices, job mobility, and wages. I propose a model that combines differences in worker ability, firm productivity, and occupational specialisation with a network-based job search process. I estimate the model using data from Germany, and assess the effect of coworker networks on labour market outcomes, and compare their efficiency to the non-network channels.

The model in this paper features two key elements: heterogeneity in occupations and firms, and job search through coworker networks. The first part of the model builds on an on-the-job search framework with heterogeneous workers, firms, and occupations. Worker productivity depends jointly on the firm and the occupation and is

influenced by the worker's own ability.¹ Firms differ in productivity, and workers have occupation-specific comparative advantages. Workers can switch jobs across firms and occupations, and may even move to lower-productivity occupations if matched with sufficiently productive firms.²

The second part of the model incorporates job search through both general and network-based channels following Mortensen and Vishwanath (1994). General search is modelled as random matching, while network search uses connections formed with coworkers. These connections are occupation-firm specific and are created when a worker joins a job. If the worker leaves for another job, these links are lost, but if the worker becomes unemployed, the links are preserved and can be used in future job searches. When searching through networks, a worker randomly selects among their retained links, with the likelihood of choosing a particular job determined by the relative strength of their connections. I follow the *Labour flow networks* approach (Guerrero and Axtell, 2013) to the definition of a coworker link. Firms are linked if a worker moves from one firm to the other. So the *labour flow networks* are composed of the links between firms that are generated by the movement of workers between them.

The model produces several predictions. High-ability workers are more likely to sort into highly specialised occupations. Workers with more network connections to a given occupation are more likely to move into that occupation. Furthermore, as workers move up the job ladder, their connections to better occupations increase, making upward occupational mobility easier. This reflects the fact that specialised occupations are more concentrated higher on the job ladder.

This framework allows for the characterisation of how the occupational composition of coworkers influences occupational choice. In the empirical implementation, I identify the network effect by exploiting variation in coworker composition within the

¹ The model is similar to the Roy model with search frictions in Taber and Vejlín (2020), but without the compensating differentials and human capital accumulation present in their model. But here, I introduce job search through coworker networks.

² In Germany, using the SIEED, I calculate that 15% of occupational mobility is downwards, i.e., to an occupation with a lower skill requirement. This is also consistent with the case of the United States, where Forsythe (2023) documents that a significant part of occupational movers move down the occupational ladder.

same job, given the worker ability, firm and occupation productivity.

I evaluate the model using administrative data from Germany. I use a broad definition of occupation based on skill requirements (unskilled tasks, skilled tasks, complex tasks), and I proxy the worker ability with the level of education (low, medium, high). I use the wage information in an AKM (Abowd et al., 1999) style regression to estimate the firm effects, worker ability, and the occupation productivity. The results suggest that high ability workers have a comparative advantage in high skilled tasks compared to low ability workers.

I then estimate the acceptance rate of each occupation based on the current firm, occupation and worker ability, which I use as controls for the regression of occupational choice on the occupational composition of coworkers. The results show that workers with a higher share of coworkers with a certain occupation are more likely to end up in that occupation following a job-to-job transition or a non-employment to employment transition.

I use these results as inputs to estimate the model using indirect inference (Gourieroux et al., 1993). The results suggest a significant role for coworker networks in job search. It account for 18% of search intensity for employed workers and 36% for the unemployed.

I run a counterfactual exercises to evaluate the role of the coworker networks on wages, occupational choices and mobility. I simulate the case where workers dedicate their entire search effort to the general channels. The results indicate that network-based search leads to higher wages, more wage dispersion, higher levels of transitions, and the reallocation of labour to highly skilled complex tasks.

There is a large body of literature on the mechanics of occupational choice. One strand focuses on the effects of occupation-specific human capital accumulation; as a worker's productivity improves with experience, this shapes their future occupational choices (Guvenen et al., 2020; Lise and Postel-Vinay, 2020). A second strand studies environments in which workers are uncertain about their own abilities in a given occupation but learn about them through experience (Miller, 1984; Moscarini and Vella,

2008). This paper focuses on how differences in access to opportunities induced by coworker networks shape occupational choices. This occurs in an environment where workers possess comparative advantages in certain occupations due to their ability (education or training). Coworkers direct a worker's search effort toward more productive jobs. Since more productive, high-skilled occupations are concentrated higher up the job ladder, as workers advance, their social capital improves, thereby enhancing access to these occupations.³

The mechanism in this paper is closely related to Arbex et al. (2019), whose model suggests that jobs obtained through networks pay higher wages and that workers with stronger connections climb the job ladder faster. The key difference is that, in this paper, occupational mobility is an integral part of a worker's upward movement on the job ladder. Moreover, this movement is driven by endogenously formed networks through coworker interactions, rather than by exogenously given networks as in Arbex et al. (2019).

The findings of this paper contrast the literature on the effects of stronger ties — such as family networks — on occupational choice and mismatch (e.g., Bentolila et al., 2010; Horváth, 2014; Zaharieva, 2018; Alaverdyan and Zaharieva, 2022). That literature suggests that such ties may lead to occupational mismatch, if social connections do not necessarily link workers to jobs suited to their abilities. By contrast, in this paper, networks channel workers to better jobs.⁴ The framework features a clear hierarchy of productivity across broad occupational categories⁵, and—as suggested by the data—this ordering holds for all workers. This implies that the standard notion of

³Another mechanism with which coworkers affect occupational choice is team learning (e.g., Herkenhoff et al., 2024; Jarosch et al., 2021), i.e., the occupation specific human capital improving by learning from coworkers. This may lead workers to choose occupations given the human capital acquired through their interaction with coworkers.

⁴The main reason is that, in this paper, coworker networks are homophilous, i.e., workers associate with others with the same ability. Since I consider past coworkers who separated to another job, the connections the workers have are necessarily to better jobs (given the workers with the same ability will have the same productivity ordering).

⁵This is unlike the literature that look at the productivity at the occupational level as the interaction between the set of skills that a worker possesses (for example, cognitive, physical, etc.) and the skill requirement of the job (e.g., Guvenen et al., 2020; Lindenlaub, 2017; Lindenlaub and Postel-Vinay, 2023; Lise and Postel-Vinay, 2020). I assume a simple reduced form worker-occupation component similar to Carrillo-Tudela and Visschers (2023) and Taber and Vejin (2020).

occupational mismatch does not necessarily apply here. Instead, the results suggest that coworkers help workers obtain jobs in which they have a productive advantage.

The rest of the paper is structured as follows. Section 3.2 presents the theoretical model. Section 3.3 describes the estimation procedure of the model, presents the model estimates, and runs counterfactual experiments to evaluate the implications of coworker network search on labour market outcomes. Section 3.4 concludes.

3.2 Model

3.2.1 Environment

Time is continuous and goes forever. We consider an economy at the steady state, and populated by an infinitely lived unit mass of risk neutral workers and an infinitely lived unit mass of risk neutral firms. Workers can be either employed or unemployed, and can search for jobs in both states. The exogenous separation rate (job destruction) is unique for all workers and given by δ . The agents operate in an environment with search friction, but with complete information otherwise. There is free entry and the cost of posting a vacancy is zero.

There are discrete types of workers in terms of ability, denoted ϵ , and the share of workers with ability ϵ is given by μ_ϵ . Worker types are observed by the researcher (for example education, vocational training, etc.).

Firms, occupations, and jobs

This model features a two dimensional job ladder.⁶ The first dimension is firms, and the second is occupations.

Firms are heterogeneous, with general productivity levels p drawn from a distribution $\Gamma(p)$ with a density function $\gamma(p)$, and supported on the interval $[\underline{p}, \bar{p}]$. The definition of the firm here is similar to that in Chapter 2; a unit that produces p , and

⁶ This approach is similar to that of Robin et al. (2014) and Taber and Vejlín (2020).

where its employees can interact.⁷

The economy features multiple occupations, indexed by $k = 1, 2, \dots, K$, and each occupation has a specific skill requirement or return to ability. The share of job vacancies in the economy of type k is independent of p and given by η_k .⁸ When a worker with ability ϵ is employed in occupation k by a firm with productivity p , the match produces an output equal to $y(\epsilon, k, p)$. I assume that the frictions that firms face to movements within the firm, i.e., within firm occupational change, is similar to the frictions they face when hiring from the outside. Here, the internal job ladder is the same as the external job ladder.⁹

Now, it is useful to work in terms of jobs not in terms of firms and occupations. We collapse the two dimensions of the job ladder (firms and occupations) into a single dimension (jobs). A job is a firm-occupation pair (p, k) that produces y (short for $y(\epsilon, k, p)$). The density y of is given by $\phi(y) = \eta_k \gamma(p)$. Given that workers are heterogeneous, they might have different orderings of job productivities. Although the probability of finding a job y is the same for all workers, their preference ordering depends on their type. Thus, the cumulative distribution function over jobs depends on worker type, and is given by $\Phi_\epsilon(y)$. The model will be solved separately for each worker type, so for brevity, I will drop the ϵ subscript, and the cumulative distribution will be $\Phi(y) = \int_{\underline{y}}^y \phi(x) dx$.

Production function

I make some assumptions on the functional form of $y(\cdot)$ that produce two features. First, some occupations are more productive than others. Second, some worker types have comparative advantages in some occupations. This could be achieved by assuming that the production function $y(\cdot)$ is multiplicatively separable into a worker-

⁷ In the quantitative analysis in Section 3.3, the firm in the model corresponds to the establishment in the data.

⁸ The notion of the firm offering a number of occupation which are drawn with a given probability follows (Papageorgiou, 2018).

⁹ In Robin et al. (2014)'s framework, there are distinctions between the internal and the external job ladders, and the barriers to internal mobility are different to those to external mobility. This is an interesting mechanism to consider for future work.

occupation component and a firm component. Specifically, the worker-occupation component (or occupation specific human capital) is time invariant¹⁰ given by π_{ek} , and the production function is $y(\epsilon, k, p) = \pi_{ek}p$.¹¹

To showcase the implications of this functional form, assume we have two occupations; s represents a skilled occupation and n represents an unskilled occupation, and two worker types; h is high skilled workers and l low skilled workers. The first feature means that high skilled occupations are more productive than low skilled occupations; $\pi_{es} > \pi_{en}$. Second, if high skilled workers have a comparative advantage in high skilled occupation, we will have $\frac{\pi_{hs}}{\pi_{hn}} > \frac{\pi_{ls}}{\pi_{ln}}$.¹²

Wage setting

I make two simplifying assumptions on wage setting. First, I assume that the worker holds all the bargaining power so they are paid the marginal productivity. The wage of a worker with ability ϵ and occupation k working in firm p is $R(\epsilon, k, p) = y(\epsilon, k, p)$. As the focus of the paper is the occupational allocation in a frictional labour market, this assumption provides a clearer link between wages on one hand, and the interplay between the skill of the worker and the skill requirements of the occupation they hold, on the other.¹³ Under a more standard wage setting protocols such as the sequential auctions framework of Postel-Vinay and Robin (2002), or the bargaining framework of Cahuc et al. (2006), the productivity of the job offer from a poaching firm (or the previous employer) play a role in wage determination. I present and discuss a version of this model with wages determined under the Postel-Vinay and Robin (2002) sequential auctions framework in Appendix C.1.

Second, I assume that agents are myopic; they make a decision if it improves their

¹⁰ I assume no human capital accumulation.

¹¹ This formulation of the occupation specific human capital corresponds to the job types in Taber and Vejlin (2020). It is also a time invariant version of the reduced form formulation in Carrillo-Tudela and Visschers (2023).

¹² The limit cases here are if $\pi_{es} = \pi_{en}$ for all ϵ , we are in a model with homogenous occupations and heterogenous workers. If $\pi_{hk} = \pi_{lk}$ for all k , we have homogenous workers and heterogenous occupations.

¹³ Firms under this assumption, and given costless vacancy posting, are indifferent between posting a vacancy and not posting it. The only barrier for the worker to joining the firm is search frictions.

current state. This assumption implies that the worker's movement from one job to another depends only on the job pay, which is exogenous, and it means that the worker problem and the steady state can be solved separately.¹⁴

In an on-the-job search context, workers will move to a higher paying job, conditional on finding a job. They can move from a generally high paying firm to a low paying firm if they move to a more fitting occupation yielding a better pay. In addition, this will lead to sorting, both on the occupation level and the firm level, since the workers' preference ordering over jobs depends on the worker-occupation component.

Utility flow from unemployment

The utility flow from unemployment is ability specific and given by z_ϵ . The model setup implies that the least productive firms in equilibrium will only be able to attract workers in their most productive occupations; this is $\pi_\epsilon^* \underline{p} = z_\epsilon$, where π_ϵ^* is the occupation specific human capital for the most productive occupation for the ability group ϵ .

I impose a simplifying assumption on the form of z_ϵ , by making it proportional to π_ϵ^* , such that $z_\epsilon = \pi_\epsilon^* \bar{z}$, where $\bar{z}$ is a common factor for all ability groups. Under this assumption, the lower bound of the productivity distribution $\underline{p}$ is the same for all ability groups, and the least productive firm can attract workers from all ability groups.

Since I solve the model for each group separately, I drop the ϵ subscript and now the utility flow from unemployment will be given by z .

3.2.2 Job search and coworker networks

Workers engage in job search both when unemployed and on the job. They search through two types of channels: general channels and coworker-based networks. I follow the Mortensen and Vishwanath (1994) approach with a mix of random matching for the general channels, and balanced matching, where the worker samples a job

¹⁴ This feature exists in the sequential auctions framework (Postel-Vinay and Robin, 2002). The sequential auctions version Appendix C.1 relaxes the assumption of myopic agents.

among the firm's employees. In this paper, the worker sample jobs from among their connections they form when they start a job.

First, let's consider on the job search. Through the general channels, the matching process is random. The sampling distribution of the job is the same as the productivity distribution. Let λ_e be the search intensity when employed. Let α be the share of the worker's job search time dedicated to coworker networks – with $1 - \alpha$ spent on search through the general channels. The arrival rate of a job y through the general channels is $(1 - \alpha)\lambda_e\phi(y)$.

When a worker goes to unemployment, they keep the connections from their previous firm, which can help them find a job. The same as before, the workers can search over general channels with probability $1 - \alpha_u$ and over networks with a probability α_u , and with a search intensity λ_u .

Coworker networks

I follow a *labour flow networks* approach (Guerrero and Axtell, 2013). Labour flow networks are those that are composed of links generated by the movement of workers between firms (or jobs in this case). Coworker networks are the collection of these links. In this section, I solve for the characteristics of the network at the steady state.

I assume that the coworker networks are job specific; the worker forms these networks when they join a job and lose them when they move to another. Additionally, I assume these networks are homophilous; the links are formed with workers of the same type holding the same job (occupation). The links are also directed; the connection can go from job i to j , they do not necessarily go the other way around. These links are generated by direct job-to-job transitions (not unemployment to employment transitions).

I make two additional key assumptions on how networks affect job search, and on the formation of the network;

Assumption 1. The sampling probability of a job v for a worker with a job y through the network is equal to the proportion of the measure of links to v in the total measure

of links emanating from the job y .

Assumption 2. The rate of loss and gain of links from job v to job y are equal.

These assumptions yield endogenous job sampling distributions. Here, matching through networks is not random; it is generated by the movement of workers between jobs.¹⁵

By Assumption 1, the probability a given job is sampled is equal to the ratio between the measure of links to that job and the total measure of links the worker possesses. The idea is that the worker is more likely to encounter a job opportunity if a larger proportion of their connections are to that job.

Consider a link between two jobs producing y and v , respectively, as the measure of workers who move from job y to job v . In this case we must have $y < v$. Let $Q(y, v)$ be the measure of links between y and v at the steady state. The total measure of links for a worker in job y is $Q(y)$, and obtained by integrating over all available links, $Q(y) = \int_{\bar{y}}^y Q(y, v)dv$.

Then, assume the links between the two jobs are destroyed with a constant rate h . At the steady state, the measure of links gained is equal to the measure of links lost (by Assumption 2). The balance equation is given by;

$$l(y)(1 - \alpha)\lambda_e\phi(v) + l(y)\alpha\lambda_e\frac{Q(y, v)}{Q(y)} = hQ(y, v) \quad (3.1)$$

where $l(y)$ is the average measure of workers in with a job y . The measure of links gained is the measure of workers in y who move to v through the general channels $[l(y)(1 - \alpha)\lambda_e\phi(v)]$ plus those who move through the network $[l(y)\alpha\lambda_e\frac{Q(y, v)}{Q(y)}]$. Now, to solve for the share of links to v for those who work in y , $\frac{Q(y, v)}{Q(y)}$, we integrate Equation 3.1 from y to $\bar{y}$, which will give us the expression for $Q(y)$, and then we replace it back

¹⁵ This is in contrast to matching through the general channels, which is random.

into 3.1.¹⁶ The solution is shown in Equation 3.2 below.

$$\frac{Q(y, v)}{Q(y)} = \frac{\phi(v)}{1 - \Phi(y)} \quad (3.2)$$

we find that the sampling probability of v through the network (the share of the links to v) is simply the density of jobs v over the share of jobs more productive than y . We see in the solution that the relevant characteristics of the network depend only on the fundamentals of the model –namely the productivity distribution–, but not on the general state of the economy in terms of employment or the distribution of firm size. We can solve for the other equilibrium objects in the next sections by plugging in the solution in Equation 3.2 into the sampling distribution of jobs, the job arrival rates, and the various employment balance equations in the next sections.

Sampling distribution of jobs

The setup of the model implies that matching is not random; the sampling distribution of jobs is not identical to the productivity distribution. The way workers sample jobs depends on their current job. Let $f(v|y)$ be the sampling probability density function for a job v conditional on working in job y at the steady state. The expression for $f(v|y)$ is given by;

$$\begin{aligned} f(v|y) &= (1 - \alpha)\phi(v) + \mathbb{1}[v > y]\alpha \frac{Q(y, v)}{Q(y)} \\ &= (1 - \alpha)\phi(v) + \mathbb{1}[v > y]\alpha \frac{\phi(v)}{1 - \Phi(y)} \end{aligned} \quad (3.3)$$

When $v \leq y$, the worker can only find jobs using the general channels, and thus the density of jobs found does not depend on their current job, and $f(v|y) = (1 - \alpha)\phi(v)$. On the other hand, because the worker can find a more productive job through networks, the probability of density of the job depends on the current job and becomes

¹⁶ Specifically, integrating Equation 3.1 over v from y to $\bar{y}$ yields $l(y)(1 - \alpha)\lambda_e(1 - \Phi(y)) + l(y)\alpha\lambda_e = hQ(y)$, which is the balance equation for the total measure of links formed in a job y . The left hand side is the measure of link formed, which are the measure of coworkers who move to a more productive firms, and the right hand side is the measure of links lost.

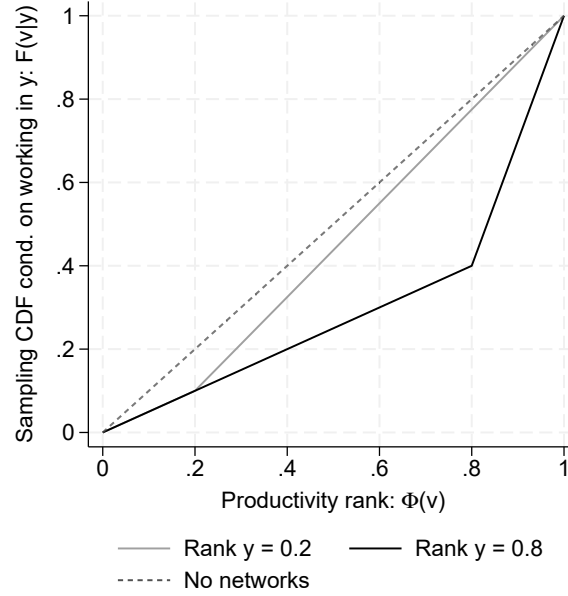


Figure 3.1: The productivity CDF of a job v and the sampling CDF of v conditional on working in y

Notes: This graph visualises equation 3.4. The share of search effort directed to networks α is set to 0.5.

$f(v|y) = \phi(v) \left[(1 - \alpha) + \frac{\alpha}{1 - \Phi(y)} \right]$ when $v > y$. This expression comes from replacing the share of links to job v from job y which is given by $\frac{Q(y,v)}{Q(v)}$ by its steady state value $\frac{\phi(v)}{1 - \Phi(y)}$ from Equation 3.2.

To arrive at the cumulative distribution $F(v|y)$, we integrate $f(v|y)$ from $\underline{y}$ to v . It is the probability of finding a job less productive than v for a worker in a job y . which is given by

$$F(v|y) = \Phi(v)(1 - \alpha) + \mathbb{1}[v > y] \alpha \frac{\Phi(v) - \Phi(y)}{1 - \Phi(y)} \quad (3.4)$$

The conditional sampling density of jobs v for an unemployed worker who lost their job y follows a similar formulation to that of the employed. But here, we replace α with the probability of searching through the network for the unemployed α_u . The conditional cumulative sampling distribution of the an unemployed worker who lost their job y is $F^u(v, y)$, which satisfies the equation below

$$F^u(v|y) = \Phi(v)(1 - \alpha_u) + \mathbb{1}[v > y] \alpha_u \frac{\Phi(v) - \Phi(y)}{1 - \Phi(y)} \quad (3.5)$$

Figure 3.1 illustrates the relationship between the productivity distribution (on the

horizontal axis) and the sampling distribution of a job v for workers with a job y (on the vertical axis). This pattern applies to both the employed and the unemployed. We see that, compared to the baseline without networks – where the sampling distribution of v is equal to the productivity distribution – the model with networks produces a weakly convex relationship between the two variables. Here, networks increase the sampling probability of jobs higher on the job ladder. Moreover, the effect of networks is stronger the higher the worker goes on the job ladder. If we consider the case of two workers, one employed in a job with a rank 0.2, and another in job ranked 0.8, and where the share of search effort dedicated to networks α is set to 0.5. For the former, the probability of sampling a job more productive than the 80th percentile is about 21%, whereas for the latter, it is 60%. So network quality is better up the job ladder as it directs the worker's search towards more productive jobs.

As for the role of networks in occupational choice, the relationship is more complex, as it depends on the relative productivities of the different occupations (π_{ek}) and the shape of the firm productivity distribution. But generally, we can deduce that if a worker is more productive in a given occupation, then jobs with that occupation will be concentrated higher on the job ladder. Thus, workers in the most productive jobs in their most productive occupation (or lost a job that is productive) will find it easier to encounter jobs in the occupation in which they are more productive, and will seem more attached to that occupation.

Job arrival rates

Now we turn to the derivation of the job arrival rates for the employed and the unemployed at the steady state. Let $\xi(y, v)$ be the arrival rate of a job v for a worker employed in y . The arrival rate is simply the conditional density times search intensity;

$$\begin{aligned}\xi(y, v) &= f(v|y)\lambda_e \\ &= (1 - \alpha)\lambda_e\phi(v) + \mathbb{1}[v > y]\alpha\lambda_e\frac{\phi(v)}{1 - \Phi(y)}\end{aligned}\tag{3.6}$$

We arrive at this expression by replacing the sampling density by its solution from Equation 3.4. The first term is the arrival rate through the general channels, and the second term is the arrival rate through coworker networks.

Similarly, for an unemployed worker who lost their job y , the arrival rate of a job v is given by $\chi(y, v)$;

$$\chi(y, v) = (1 - \alpha_u)\lambda_u\phi(v) + \mathbb{1}[v > y]\alpha_u\lambda_u\frac{\phi(v)}{1 - \Phi(y)} \quad (3.7)$$

We then plug in the arrival rate results from Equations 3.6 and 3.7 into the balance equations in the next section.

3.2.3 Steady state

Let E be the employment rate and U be the unemployment rate. Let $e(y)$ be the share of employment in y and $u(y)$ be the share of unemployed workers who lost their job in y . The steady state balance equation for $u(y)$ is given by

$$U(u(y)(1 - \alpha_u)\lambda_u(1 - \Phi(z)) + u(y)\alpha_u\lambda_u \int_y^{\bar{y}} \frac{\phi(t)}{1 - \Phi(y)} dt) = \delta E e(y) \quad (3.8)$$

The inflow is $\delta E e(y)$ and the outflow is the measure of workers who find the job through the general channels $[Uu(y)(1 - \alpha_u)\lambda_u(1 - \Phi(z))]$ and the measure of workers who find a job through networks $[Uu(y)\alpha_u\lambda_u \int_y^{\bar{y}} \frac{\phi(t)}{1 - \Phi(y)} dt] = Uu(y)\alpha_u\lambda_u$. Notice a worker can only find a job through networks in a job more productive than theirs. Solving the equation above gives the unemployment rate

$$U = \frac{\delta}{\delta + \lambda_u(1 - \alpha_u\Phi(z))} \quad (3.9)$$

Then, by plugging Equation 3.9 back into Equation 3.8, we see that the density of employed workers in y is equal to the density of unemployed workers who lost their job y ; i.e., $e(y) = u(y)$.

Now, we turn to the balance equation for employment, which given by

$$\begin{aligned}
 Ee(y)(\delta + \lambda_e(1 - \alpha)(1 - \Phi(y)) + \lambda_e\alpha \int_y^{\bar{y}} \frac{\phi(t)}{1 - \Phi(y)} dt) \\
 = \phi(y) \left[U(1 - \alpha_u)\lambda_u + U\alpha_u\lambda_u \int_z^y \frac{u(t)}{1 - \Phi(t)} dt \right. \\
 \left. + E(1 - \alpha)\lambda_e \int_z^y e(t) dt + E\alpha\lambda_e \int_z^y \frac{e(t)}{1 - \Phi(t)} dt \right]
 \end{aligned} \tag{3.10}$$

Using the unemployment balance equations this simplifies to

$$\begin{aligned}
 e(y)(\delta + \lambda_e(1 - \alpha)(1 - \Phi(y)) + \lambda_e\alpha) \\
 = \phi(y) \left[\frac{(1 - \alpha_u)\delta}{1 - (1 - \alpha_u)\Phi(z)} + \frac{\alpha_u\delta}{1 - (1 - \alpha_u)\Phi(z)} \int_z^y \frac{e(t)}{1 - \Phi(t)} dt \right. \\
 \left. + (1 - \alpha)\lambda_e \int_z^y e(t) dt + \alpha\lambda_e \int_z^y \frac{e(t)}{1 - \Phi(t)} dt \right]
 \end{aligned} \tag{3.11}$$

The differential equation above has a convenient closed form solution. We first define the average job size as a function of the productivity rank as $l(\phi(y)) = \frac{e(y)}{\phi(y)}$ and then we perform a change of variable from $\Phi(y)$ to x , which gives;

$$\begin{aligned}
 l(x)(\delta + \lambda_e(1 - \alpha)(1 - x) + \lambda_e\alpha) \\
 = \frac{(1 - \alpha_u)\delta}{1 - (1 - \alpha_u)\Phi(z)} + \frac{\alpha_u\delta}{1 - (1 - \alpha_u)\Phi(z)} \int_0^x \frac{l(t)}{1 - t} dt \\
 + (1 - \alpha)\lambda_e \int_0^x l(t) dt + \alpha\lambda_e \int_0^x \frac{l(t)}{1 - t} dt
 \end{aligned} \tag{3.12}$$

The solution to the differential equation is

$$l(x) = C(1 - x)^{\frac{-(k + \alpha\lambda_e)}{\delta + \alpha\lambda_e}} \left[\delta + \alpha\lambda_e + (1 - \alpha)\lambda_e(1 - x) \right]^{\frac{k - \alpha\lambda_e - \delta}{\alpha\lambda_e + \delta}} \tag{3.13}$$

where $k = \frac{\alpha_u\delta}{\alpha_u\Phi(z)}$, and the constant C satisfies $l(z) = \frac{\delta(1 - \alpha_u)}{(1 - (1 - \alpha_u)\Phi(z))(\delta + \lambda_e(1 - \alpha)(1 - \Phi(z)) + \lambda_e\alpha)}$.

From this, we see that more productive job cells are larger ($\frac{\partial l(x)}{\partial x} > 0$). Additionally, more productive firms are larger because of the independence of the sampling probability of occupations and the sampling density of firms, and the assumption that job productivity is increasing in firm productivity.

We can calculate the share of employed workers in jobs less productive than y and given by $E(y)$. We integrate $l(x)$ in Equation 3.13 with respect to x from 0 to 1, and then we change the variable x back into $\Phi(y)$. Since the density of employed workers in y is equal to the density of the of unemployed workers who lost their jobs y , then the equation above also describes the share of unemployed workers who lost a job less productive than y , and given by $U(y)$. Thus, $E(y)$ and $U(y)$ are given by

$$E(y) = U(y) = \frac{C}{k - \delta} \left[\frac{1 - \Phi(y)}{\delta + \alpha\lambda_e + (1 - \alpha)\lambda_e(1 - \Phi(y))} \right]^{\frac{\delta - k}{\delta + \alpha\lambda_e}} - 1 \quad (3.14)$$

Now, we can define steady state equilibrium. It is characterised by the balance of links gained and lost within the firm, and the usual employment flow balance. Specifically;

Definition. The steady state equilibrium is a collection $\{Q(y, v), \xi(y, v), \chi(y, v), F(v|y), F^u(v|y), E(y), U(y), U\}$ such that

1. The measure of links for a worker with a job y to a job v at the steady state – $Q(y, v)$ – satisfies Equation 3.2.
2. The steady state job arrival rate for the employed $\xi(y, v)$ and the job arrival rate for the unemployed $\chi(y, v)$ satisfy Equations 3.6 and 3.7, respectively.
3. The sampling distribution of jobs conditional on working in a job y is given by $F(v|y)$ and satisfies Equation 3.4.
4. The sampling distribution of jobs for an unemployed worker who lost a job y is given by $F^u(v|y)$ and satisfies Equation 3.5.
5. The share of employed workers in jobs less productive than y and the share of unemployed workers who lost a job less productive than y , $E(y)$ and $U(y)$ respectively, satisfy equation 3.14.
6. The unemployment rate U satisfies Equation 3.9.

3.3 Quantitative analysis

In this section, I estimate and evaluate the model using administrative data from Germany. I use the wage information to estimate the productivity parameters offline,

Table 3.1: Summary statistics

	Low education	Medium education	High education	All
Individual characteristics (share)				
Female	0.51	0.50	0.45	0.49
Occupation (share)				
Unskilled task	0.42	0.17	0.04	0.15
Skilled task	0.48	0.67	0.26	0.57
Complex task	0.10	0.17	0.70	0.28
Number of observations	1,373,711	18,049,408	5,366,628	24,789,747
Number of workers	57,485	599,704	189,224	846,313
Number of establishments	15,092	44,063	21,416	48,254

Notes: Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990.

which then are going to be used to estimate the transition and network parameters internally using indirect inference. I then run counterfactuals experiments where workers dedicate the entirety of their search effort to the general channels to evaluate the efficiency of coworker networks, and their implications on wages, mobility, and occupational allocation.

3.3.1 Data

I use the dataset Sample of Integrated Employer-Employee Data (SIEED), published by the Institute for Employment Research (IAB) of the Federal Employment Agency (BA) (Schmidtlein et al., 2020). I conduct the analysis on a monthly worker panel spanning the years 2014 and 2017.¹⁷ I restrict the analysis to German workers born between 1955 and 1990.¹⁸ For occupations, I use the broad task classification of the occupation (unskilled task, skilled task, complex tasks, highly complex tasks). I group the complex and highly complex tasks together to form the complex tasks group. I drop the observations where the occupation data is unavailable. Table 3.1 shows the details and composition of the final sample.

When the wage information is used, I further restrict the sample to full-time work-

¹⁷ When constructing the network variables, I use observations from 2012 onwards.

¹⁸ The purpose is to use working age individuals, but to fix the sample of workers based on a unvarying characteristic.

ers subject to social security. As wages are censored above the social security contribution ceiling, I impute the censored observations using the two-step procedure in Dustmann et al. (2009) and Card et al. (2013).¹⁹

3.3.2 Parameterisation and identification

In this section, I will describe the parametric assumptions and identification of the model. I first use the moments of the firm and education-occupation fixed effects from a log wage regression to identify and estimate the productivity and ability parameters. I then use these estimates as inputs for the estimation of the transition and network parameters using indirect inference.

Productivity and ability

I identify the productivity and ability parameters using the wage information. Since workers are paid their marginal product, and given the multiplicative production function, the log of wages can be decomposed into a firm component and an occupation-ability component. The parameters of the firm productivity distribution can be recovered using the moments of the firm fixed effects, and the ability parameters using the ability-occupation fixed effects. The details of the parameterisation and identification of the productivity and ability parameters are described below.

Recall the production function is given by $y(\epsilon, p, k) = \pi_{\epsilon k} p$, where $\pi_{\epsilon k}$ governs the relative productivity of each occupation for each ability. I assume the firm productivity distributed p is such that $p - \underline{p} = \theta_p \sim \text{Beta}(\theta_1, \theta_2)$. I also assume that there are three types of occupation and three types of workers. The occupations correspond to the data, and are represented by the unskilled, skilled, and complex occupations. Similarly, the worker ability is represented by the level of education; low, medium, and high.

Since the workers are paid the marginal product, and given the functional assumption on the production function, the log of marginal product (and wages) is given by $\log(y(\epsilon, p, k)) = \log(\pi_{\epsilon k}) + \log(p)$. We can thus recover the productivity and ability

¹⁹ The code is based on that of Dauth and Eppelsheimer (2020).

Table 3.2: Estimates of ability and task productivity

Task	Education		
	(1) Low	(2) Medium	(3) High
Unskilled	0.628	0.643	0.644
Skilled	0.741	0.779	0.922
Complex	1	0.997	1.299

Notes: Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990. It includes only full-time workers subject to social security. Controls are the interaction between gender and year of birth. Wages above the social security contribution ceiling are top coded. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009); Card et al. (2013). See appendix A.3.2 for the imputation details.

parameters with an AKM Abowd et al. (1999) style regression, with firm fixed effects, and occupation and education interacted

$$\log(w_{it}) = \psi(i, t) + \sum_{\epsilon} \sum_k P_{\epsilon k} \mathbb{1}[\kappa(i, t) = k] \mathbb{1}[\epsilon_i = \epsilon] + X_i \beta + u_{it} \quad (3.15)$$

Where w_{it} of the wage of worker i at time t . $\psi(i, t)$ are firm fixed effects, $\kappa(i, t)$ is the occupation, and ϵ_i is education. u_{it} is the error term. X is a group of controls that is the interaction between gender and year of birth. $P_{\epsilon k}$ is the worker-occupation component to estimate and is equal to $\log(\pi_{\epsilon k})$. To arrive at a consistent estimate of the model the error term u_{it} must be uncorrelated with any worker-firm specific characteristics, i.e., there should not be sorting on unobservable worker characteristics.

I replace the coefficients for the education-occupation fixed effects, which are nine parameters denoted $\hat{\pi}_{\epsilon k}$, directly into the model. I normalise the parameters such that the base category is low education workers with complex tasks is one.

Table 3.2 shows the estimates of the education-occupation components ($\hat{\pi}_{\epsilon k}$). We see that, first, wages increase with the task complexity for all levels of education. Second, workers with a higher level of education have a comparative advantage in high skilled tasks. For example, highly educated workers earn on average twice as much in complex tasks compared to unskilled tasks (Column 3), whereas the ratio for workers

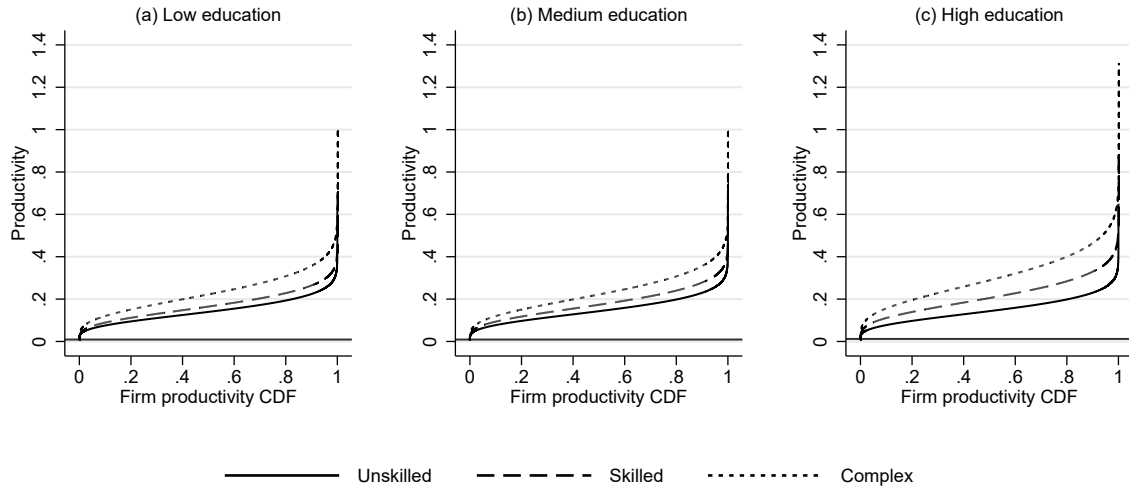


Figure 3.2: Worker productivity by firm productivity rank and occupation

Notes: Author's calculations using SIEED. The estimates are based on Tables 3.2. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990. It includes only full-time workers subject to social security. Wages above the social security contribution ceiling are top coded. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009); Card et al. (2013). See appendix A.3.2 for the imputation details.

with a low level of education is 1.6 (Column 1). Overall, the highly educated workers exhibit the highest level of specialisation. On the other hand, low and medium educated workers are less specialised and have a similar estimates in terms of the worker ability components.

I use the firm fixed effects estimates, denoted ψ_j to calculate the productivity distribution parameters. In the model, the estimates should give us the log of productivity plus a constant; we therefore need to recentre then exponentiate the estimates. Following the parametric assumption on the firm productivity distribution, the upper bound of θ_p is one and the lower bound is zero (since it is generated from a Beta distribution). We can solve for $\underline{p}$ by taking the difference between the maximum and the minimum; $\overline{\psi_j} - \underline{\psi_j} = \log(\underline{p} + 1) - \log(1)$. Thus the estimate for the productivity of the least productive firm is $\underline{\hat{p}} = \frac{1}{\exp(\overline{\psi_j} - \underline{\psi_j}) - 1}$. Then we can recentre by adding $\log(\underline{\hat{p}}) - \underline{\psi_j}$ to ψ_j and then we exponentiate, which will give us the desired scalar for the firm productivity, denoted $\hat{\psi_j} = \exp(\log(\underline{\hat{p}}) - \underline{\psi_j} + \psi_j)$.

I then solve for the productivity parameters (θ_1, θ_2) , matching the mean and standard deviation of $\hat{\psi}_j$ to those produced by the Beta distribution, for which the solution is unique.

Figure 3.2 provides a picture of the firm productivity distribution estimate and how it relates to (productivity) wages. On the horizontal axis we have the firm productivity rank and on the vertical axis the corresponding productivity by education and occupation. As well as showing the degree of specialisation for each education group (shown by the gap between the occupation lines), the figure maps the possible transitions for each group (conditional on finding a job). For example, a low or medium educated worker with a complex job in a firm at the 40th percentile of the productivity distribution will be willing to move to an unskilled job at the 80th percentile. On the other hand, a highly educated worker in the same position will only be willing to move to an unskilled job in a firm in the 99th percentile. The setup of the model and the corresponding estimates suggest that highly educated workers are more selective on the occupation dimension.

Transition and network parameters

I fix the exogenous separation rate δ to the employment to non-employment (E2N) rate in the data. I jointly estimate the remaining parameters $(\lambda_u, \lambda_e, \eta_u, \eta_s, \alpha, \alpha_u)$ using indirect inference (Gourieroux et al., 1993). I identify the search effort in unemployment λ_u using the non-employment (N2E) rate. The higher the λ_u , the higher the N2E rate. Similarly, the job-to-job (J2J) transition rate informs the job search effort during employment λ_e . Then, the sampling probability of the unskilled and skilled tasks – η_u and η_s – identify the share of workers in the unskilled and skilled tasks.²⁰

Network parameters. The model allows us to directly characterise the probability of choosing a given occupation, conditional on ability, current firm, occupation, and the share of coworkers who switch to the occupation. We can thus identify the network parameters α and α_u using the coefficient of the regression of occupational choice on

²⁰ The share in the third occupation (complex tasks) is redundant.

the share of coworkers who switch to a given occupation.

We first consider the case of employed workers. Given the expression for the arrival rate for job v for a worker with a job y (that is $\zeta(y, v)$ and Equation 3.6), and given that the sampling probability of the occupation is independent from that of the firm, we can derive the rate with which a worker with occupation k , working in firm p moves to an occupation k' – henceforth $M(p, k, k')$. We arrive at the expression by integrating Equation 3.6 with respect to p' (where p' is the productivity of firms other than p) from p to $\bar{p}$;

$$M(p, k, k') = (1 - \alpha)\lambda_e\eta_{k'}\bar{\Gamma}\left(\frac{y(\epsilon, p, k)}{\pi_{ek'}}\right) + \alpha\lambda_e\tilde{Q}(p, k, k') \quad (3.16)$$

Where $\bar{\Gamma}\left(\frac{y(\epsilon, p, k)}{\pi_{ek'}}\right)$ is the share of firms that can offer a worker earning y a higher wage in an occupation k' ($\bar{\Gamma}(\cdot) \equiv 1 - \Gamma(\cdot)$). $\tilde{Q}(p, k, k')$ is the measure of links from a job (p, k) to a job with occupation k' over the total measure of links from job (p, k) . The first term is the general channels component, and the second term is the network component. Conveniently, both variables $\bar{\Gamma}\left(\frac{y(\epsilon, p, k)}{\pi_{ek'}}\right)$ and $\tilde{Q}(p, k, k')$ are linear in the combination of parameters $(1 - \alpha)\lambda_e$ and $\alpha\lambda_e$, respectively. This motivates the use of a linear probability model to inform the network parameter α .

First, let G_{it}^m be the share of firms of firms to which worker i at time t would be willing to move and work in occupation m , which is equivalent to $\bar{\Gamma}\left(\frac{y(\epsilon, p, k)}{\pi_{ek'}}\right)$. I use the estimates for firm, worker, and occupation fixed effects as well as the productivity parameters to derive the data counterpart. Specifically,

$$G_{it}^m = 1 - \text{Beta}\left(\frac{\hat{\pi}_{ik(i,t)}}{\hat{\pi}_{im}}\psi_{j(i,t)} - \hat{p}, \hat{\theta}_1, \hat{\theta}_2\right) \quad (3.17)$$

Where $\text{Beta}(\cdot)$ is the regularised incomplete Beta function.

I approximate the network variable $\tilde{Q}(p, k, k')$ with the share of workers with the same occupation and education who separated to occupation m in the total number of workers with the same occupation and education who separated to another job (J2J

separations) in the past three years, denoted S_{it}^m .²¹

I estimate the following linear probability model for each destination occupation;

$$q_{it}^m = \Omega_0^m G_{it}^m + \Omega^m S_{it}^m + e_{it} \quad (3.18)$$

Where q_{it}^m is the conditional probability an employed worker J2J separates to an occupation m . e_{it} is the error term.

The coefficient Ω^m corresponds to the probability of finding any job through the network for the employed (i.e., $\alpha\lambda_e$), so it is monotonically increasing in α . In the model, if α is positive, then the effect of the share of coworkers who switch to a given occupation on the probability of changing to that occupation (Ω^m) is also positive, conditional on the firm, occupation, and the worker ability. If workers direct all their effort to the general channels and α is zero, then the coefficient Ω^m will be zero.

The model is estimated for the three types of occupations; unskilled, skilled, and complex (denoted u, s, c, respectively). I match the average of the three coefficients ($\hat{\Omega}^u, \hat{\Omega}^s, \hat{\Omega}^c$), which is increasing α .

The exercise is simpler for the unemployment to employment transition. This is because the movement to a given occupation through the general channels does not depend on the productivity of the previous job; it only depends on the worker ability and their network. Particularly, the rate with which a worker who lost their job (p, k) moves to a job with occupation k' ($M^u(p, k, k')$) is simply;

$$M^u(p, k, k') = (1 - \alpha_u) \lambda_u \eta_{k'} \bar{\Gamma} \left(\frac{\pi_{ek^*}}{\pi_{ek'}} \hat{p} \right) + \alpha_u \lambda_u \tilde{Q}(p, k, k') \quad (3.19)$$

which can be approximated with the following linear probability model;

$$g_{it}^m = \sum_{\epsilon} \omega_{0,\epsilon}^m \mathbb{1}[\epsilon_i = \epsilon] + \omega^m S_{it}^m + e_{it}^u \quad (3.20)$$

Where g_{it}^m is the conditional probability that an unemployed worker who transitioned

²¹ Since a large majority of workers stay in the same job for more than three years, this is equivalent to the share of past coworkers who chose occupation k .

Table 3.3: Coworker networks and occupational choice

	Job-to-job			Non-employment-to-employment		
	(1) Unskilled	(2) Skilled	(3) Complex	(4) Unskilled	(5) Skilled	(6) Complex
Share of separations (S_{it}^m)	0.007*** (0.000)	0.006*** (0.000)	0.008*** (0.000)	0.407*** (0.000)	0.236*** (0.000)	0.424*** (0.000)
Share of acceptable firms (G_{it}^m)	0.009*** (0.000)	0.0143*** (0.000)	0.002*** (0.000)			
Observations	18,596,863	18,596,863	18,596,863	149,473	149,473	149,473
Education fixed-effects				X	X	X

Notes: Dependent variable is the probability of moving to a given occupation. Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990. Part time-workers take the values of the productivity estimates (firm, worker-occupation) from their full-time counterparts when available. Wages above the social security contribution ceiling are top coded. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009); Card et al. (2013). See appendix A.3.2 for the imputation details. Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to employment moves to a jobs with occupation m . e_{it}^u is the error term. The independent variables are education fixed effects (to take into account differences in acceptance rates by ability) and the share of coworkers that move to occupation m . As before, I match the average of $(\hat{\omega}^u, \hat{\omega}^s, \hat{\omega}^c)$, which is be increasing in the network parameter α_u .

Table 3.3 reports the results for J2J and N2E transitions. Columns (1)-(3) shows the estimates for the J2J exercise. We see that the probability of J2J transition to a given occupation is increasing in the share of coworkers moving to that occupation for all occupational classifications. The coefficients for the three regressions are statistically significant and remarkably close in magnitude, at 0.007, 0.006, and 0.008, for the unskilled, skilled, and complex tasks, respectively. I match the average of the three coefficients, which is 0.007.²² Columns (4)-(6) shows the estimates for the N2E exercise. The results suggest that the probability of choosing an occupation is increasing in the share of coworkers moving to that occupation conditional on an N2E transitions. Similar to the J2J case, the coefficients are statistically significant and similar in magnitude, at 0.407, 0.236, and 0.424, for the unskilled, skilled, and complex tasks, respectively. The

²² We also see that the probability of moving an occupation is increasing in the share of firms where the worker is willing to accept a job in that occupation. The magnitudes differ by occupation because the coefficients capture the probability of finding a job in that occupation. The results suggest that probability of finding a job in skilled tasks is the highest, followed by the unskilled tasks, and finally the complex tasks. This is corroborated by the estimates of the respective structural parameters in Table 3.4.

statistic to match is the average of these coefficients, which is 0.336.

3.3.3 Estimation

In total, we have a vector of 6 parameters to estimate using indirect inference, and 6 moments to match, so the model is just identified. The vector of estimated parameters $\hat{Z}$ minimises the weighted square difference between a set of empirical statistics (moments) and their simulated counterpart; and given by

$$\hat{Z} = \underset{Z}{\operatorname{argmin}} (M_e - M_s(Z))' W (M_e - M_s(Z)) \quad (3.21)$$

M_e is a 6×1 vector of empirical moments and $M_s(Z)$ is a 6×1 vector of simulated moments. The 6×6 weight matrix W is a matrix with the inverse of the values of the empirical moments on the diagonal and zero otherwise, i.e., the objective is to minimise the sum of the squares of the percentage change from the empirical moments.

Simulation. I simulate an economy of 100,000 workers for 50 periods. I fix the ratio of workers to firms to 16. The shares of workers with a given level of education follow the data (see Table 3.4). I run a separate simulation for each type of workers and then combine them to compute the moments.

I draw from the steady state in the first period. We can recover the job productivity rank by inverting Equation 3.14 and solving for $\Phi(\cdot)$. Let $\tilde{w}$ be a random draw from from the worker pool, the job productivity rank is given by;

$$\Phi(\tilde{w}) = \frac{1 - (\lambda_e + \delta) \left[\frac{k-\delta}{C} (\tilde{w} - 1) \right]^{\frac{\delta+\alpha\lambda_e}{\delta-k}}}{1 - (1-\alpha)\lambda_e \left[\frac{k-\delta}{C} (\tilde{w} - 1) \right]^{\frac{\delta+\alpha\lambda_e}{\delta-k}}} \quad (3.22)$$

This equation also gives the productivity rank of the last job of the unemployed workers.

Then, in the subsequent periods, we can determine the productivity rank of the new job $\Phi(v)$ for an employed worker given a random draw from the job distribution $\tilde{F}$ (replacing $F(v|y)$ for brevity) and the productivity rank of the current job $\Phi(y)$ by

Table 3.4: Estimation results

Parameter		Value	Moment	Data	Model
Estimated offline					
Worker ability distribution	μ_ϵ		Education		
Low		0.053	Low	0.053	0.053
Medium		0.723	Medium	0.723	0.723
Productivity distribution	p	0.009	$\min\{\hat{\psi}_j\}$	0.009	0.009
	θ_1	4.312	$\mathbb{E}[\hat{\psi}_j]$	0.232	0.232
	θ_2	15.034	$\text{sd}[\hat{\psi}_j]$	0.092	0.092
Worker ability/task productivity	π_{ek}	see Table 3.2	$\hat{\pi}_{ek}$	see Table 3.2	
Exogenous separation rate	δ	0.008	E2N rate	0.008	0.008
Estimated internally					
Search intensity (unemployed)	λ_u	0.119	N2E rate	0.120	0.120
Search intensity (employed)	λ_e	0.039	J2J rate	0.011	0.011
Sampling probability of tasks					
Unskilled tasks	η_1	0.344	Share unskilled	0.146	0.146
Skilled tasks	η_2	0.585	Share skilled	0.564	0.564
Network strength					
Employed	α	0.179	$\frac{\hat{\Omega}^u + \hat{\Omega}^s + \hat{\Omega}^c}{3}$	0.007	0.007
Unemployed	α_u	0.356	$\frac{\hat{\omega}^u + \hat{\omega}^s + \hat{\omega}^c}{3}$	0.336	0.336

Notes: Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990.

solving for $\Phi(v)$ in Equation 3.4, which is given by

$$\Phi(v) = \begin{cases} \frac{\tilde{F}}{1-\alpha} & \text{if } \tilde{F} \leq (1-\alpha)\Phi(y) \\ \frac{\tilde{F}(1-\Phi(y)) + \alpha\Phi(y)}{1-(1-\alpha)\Phi(y)} & \text{if } \tilde{F} > (1-\alpha)\Phi(y) \end{cases} \quad (3.23)$$

We repeat the same procedure to arrive at the productivity rank for a job v found by an unemployed worker who lost their job y and a random draw from the sampling distribution of jobs for the unemployed (similar to Equation 3.23 above, but replace $\tilde{F}$ with $F^u(v|y)$ and α with α_u).

3.3.4 Results

Table 3.4 shows the estimated parameters, along with the data and the simulated moments. The model fits the key labour market moments very closely. The search

intensity estimate for the unemployed workers λ_u is 0.119 and that of the employed workers λ_e is 0.039. The observed non-employment-to-employment (N2E) and job-to-job (J2J) transition rates replicate their simulated equivalents almost exactly (0.120 and 0.011, respectively). The sampling probabilities for different task types — unskilled (0.344) and skilled (0.585) — produce task shares in the model (0.146 for unskilled and 0.564 for skilled) that align perfectly with the data.

Regarding network strength, the model estimates that employed workers allocate about 18% of their search effort through networks, and unemployed workers about 36%. The simulated coefficient averages for the regression of the share of coworkers switching to a given occupation on the probability of moving to that occupation match those in the data (0.007 for employed and 0.336 for unemployed).

These results indicate that networks contribute significantly to mobility. The probability of a J2J transition through networks for the employed is $\alpha\lambda_e$.²³ The ratio of a J2J transition through network to that of overall J2J transition is 0.63. This is comparable to the contribution of networks to J2J mobility in Chapter 2, and within the range of 25% to 80% of jobs found through networks suggested by Topa (2019). Despite the fact that only 18% of the search effort is dedicated to network search, it accounts for 63% of J2J mobility. This is because network search is more efficient; it directs the workers to job opportunities that are more productive, thus acceptable to the worker.

As for the unemployed, the contribution of networks to N2E transitions is 35%, which is $\frac{\lambda_u\alpha_u}{N2E}$. This rate is higher than that found in Chapter 2 (between 5%-10%), but within the bounds suggested by the literature. Here, the network contribution is close to that of the share dedicated to network search because almost all jobs are acceptable to the unemployed worker, since the utility flow from unemployment z is low. Note that, although general search is almost as efficient network channels, the network channels lead to better jobs, since the jobs found through networks in the model are almost as productive as that of the previous employer of the unemployed worker.

²³ Which is obtained by integrating $\zeta(y, v)$ in Equation 3.6 over v from y to $\bar{y}$.

Table 3.5: Counterfactual exercise; dedicate search effort to general channels

	(1)	(2)	(3)	(4)
	Baseline	J2J $\alpha = 0$	N2E $\alpha_u = 0$	J2J & N2E $\alpha = 0 \text{ \& } \alpha_u = 0$
Average log wage	-1.142	-1.273	-1.323	-1.411
Var log wage	0.258	0.162	0.203	0.148
J2J rate	0.011	0.007	0.013	0.009
N2E rate	0.119	0.118	0.117	0.119
Task distribution				
Unskilled	0.146	0.165	0.185	0.212
Skilled	0.564	0.621	0.611	0.640
Complex	0.290	0.214	0.204	0.148

Notes: Except for the network parameters, the model parameters are the estimates from Table 3.4. Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990.

3.3.5 Implications on wages, mobility, and occupational allocation

In this section, I conduct counterfactual exercises to evaluate the implications of coworker networks on wages, mobility, and occupational allocation. I simulate scenarios where workers dedicate all their search effort towards general search. This allows us to compare the effectiveness of the network search to relative general channels, and to illustrate the mechanisms through which networks shape labour market outcomes.

Table 3.5 shows the results of exercises.²⁴ I examine 3 specifications and compare them to the baseline case (Column 1). The first scenario has employed workers dedicate all search effort to general search ($\alpha = 0$, Column 2). In the second, only unemployed workers dedicate their search to general channels ($\alpha_u = 0$, Column 3). The third is the case where both rely solely on general search ($\alpha = \alpha_u = 0$, Column 4). The model becomes similar to a Cahuc et al. (2006) model with workers holding all the bargaining power, and with a job arrival rate equal to λ_e for the employed and/or λ_u for the unemployed.

First, we consider the case where the employed workers dedicate their search to the general channels in Column 1. We see that the effect on average wages is small;

²⁴ Table C.2 in Appendix C.2 breaks down the effect by level of education; the main conclusions do not change.

the average log wages (or productivity) decreases slightly from -1.142 to -1.273. As in Chapter 2, networks increase wage dispersion, and the variance of log wages decreases from 0.258 at the baseline to 0.162. The model suggests that network search becomes more efficient up the job ladder. It is harder for workers at the bottom of the job ladder to breakout, while workers in the middle will find it easier to move up. Thus, jobs in the middle of the productivity distribution are smaller and jobs at the top are larger due to networks, which increases the variance of the wage.

We see that J2J mobility decreases by 36%, from 0.011 in the baseline case to 0.007 in the case without networks. This is because, in the latter, workers sample randomly from jobs, whereas with networks, workers – depending on their position on the job ladder – can sample more heavily from jobs higher on the job ladder. Thus networks are more efficient than general channels in terms of facilitating J2J mobility. On the other hand, the effect on the N2E rate is small (the rate falls from 0.119 to 0.118); the parameter α affects the N2E mobility only indirectly through the changes in the distribution of workers across firms.

In terms of occupational allocation, we see that networks channel workers away from the unskilled and skilled tasks and towards complex tasks. When workers allocate their effort towards the general search, the share of workers in complex tasks decreases from 29% to 21%. This is because, for all educational groups, complex tasks are more productive than the other occupations, thus higher on the job ladder. Since networks increase the sampling probability of jobs up the job ladder (as discussed in Section 3.2.2), removing the network channel decreases the share of workers with complex occupations.

Now we turn to the case where unemployed workers allocate all their search effort to general channels in Column (3). Similar to the employed case, networks lead to slightly higher wages and more wage dispersion. However, we see that the J2J rate increases from 0.007 to 0.013 when we shut down the network channel. This is because, in a sense, when workers lose their jobs, they do not restart from the bottom of the job ladder. Instead, through networks, they can direct their search efforts toward jobs

higher on the job ladder. This implies lower overall J2J mobility.

In terms of N2E mobility, we see a similar pattern to the first case; networks do not affect the unemployed overall N2E mobility significantly. They only affect their direction of search. This is why we see that the share of workers with complex tasks going from 0.29 in the case with networks, to 0.204 for the case without networks.

Finally, we consider the case where workers cannot search through networks, whether employed or unemployed (Column 4). Overall, shutting down network-based search results in lower wages, less dispersion, fewer J2J and N2E transitions, and a reallocation of labour away from complex tasks.

This showcases the significant implications that networks could have on the economy. We see that network search is more efficient. This is because general search is random (as the case in many search models), and search through coworker networks can help workers direct their search to job opportunities they are willing to take.

3.4 Conclusion

This paper develops and estimates a structural model to study how the composition of coworker networks shape workers' occupational choices, job mobility, and wages. The model combines occupational specialisation, firm and worker heterogeneity, and a job search process that includes both general and network-based channels.

Using administrative data from Germany, the model is estimated with indirect inference. The results show that coworker networks play a significant role in job search and occupational mobility, accounting for 18% of search effort among employed workers and 36% among the unemployed. Counterfactual exercises show that removing the network-based search leads to lower average wages, reduced wage dispersion, fewer J2J transitions, and a shift away from complex occupations—indicating the efficiency of networks in improving match quality and enabling upward mobility.

This paper highlights how networks serve as an important mechanism through which workers access new opportunities, with significant implications on the alloca-

tion of workers into occupations, mobility, and wage dispersion. The past literature emphasise the importance of coworker networks for the unemployed (Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2021; Glitz and Vejlin, 2021) and the employed (Caldwell and Harmon, 2019). This paper adds to this literature by showing how the effects of these networks are mediated by the worker's ability and their fit to their job and occupation.

Future research could explore different setups for network transmission, such as in Calvó-Armengol and Zenou (2005); Calvó-Armengol and Jackson (2007). This paper provides some rationale, as the type of job the worker finds is correlated with the type of job their coworkers find. However, it is far from the only or most realistic setup. Future research could also study how learning from coworkers affect other job market outcomes. If, for example, having coworkers with the same occupation improves the worker's occupation specific human capital, it could perhaps affect their future choices or their job search strategies.

Bibliography

- Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.
- Alaverdyan, S. and Zaharieva, A. (2022). Immigration, social networks and occupational mismatch. *Economic Modelling*, 114:105936.
- Antoninis, M. (2006). The wage effects from the use of personal contacts as hiring channels. *Journal of Economic Behavior & Organization*, 59(1):133–146.
- Arbex, M., O’Dea, D., and Wiczer, D. (2019). Network search: climbing the job ladder faster. *International Economic Review*, 60(2):693–720.
- Arnoud, A., Guvenen, F., and Kleineberg, T. (2019). Benchmarking global optimizers. Working Paper No.26340, National Bureau of Economic Research.
- Bachmann, R., Bechara, P., and Vonnahme, C. (2020). Occupational mobility in europe: Extent, determinants and consequences. *De Economist*, 168(1):79–108.
- Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, 104(6):1551–1596.
- Bagger, J. and Lentz, R. (2019). An empirical model of wage dispersion with sorting. *The Review of Economic Studies*, 86(1):153–190.
- Bentolila, S., Michelacci, C., and Suarez, J. (2010). Social contacts and occupational choice. *Economica*, 77(305):20–45.
- Blossfeld, H.-P. (1987). Labor-market entry and the sexual segregation of careers in the federal republic of germany. *American Journal of Sociology*, 93(1):89–118.
- Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, 150(1):41–55.
- Burdett, K., Carrillo-Tudela, C., and Coles, M. (2016). Wage inequality: A structural decomposition. *Review of Economic Dynamics*, 19:20–37.
- Burdett, K., Carrillo-Tudela, C., and Coles, M. (2020). The cost of job loss. *The Review of Economic Studies*, 87(4):1757–1798.
- Burdett, K., Carrillo-Tudela, C., and Coles, M. G. (2011). Human capital accumulation and labor market equilibrium. *International economic review*, 52(3):657–677.

- Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.
- Burks, S. V., Cowgill, B., Hoffman, M., and Housman, M. (2015). The value of hiring through employee referrals. *The Quarterly Journal of Economics*, 130(2):805–839.
- Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364.
- Caldwell, S. and Harmon, N. (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Unpublished manuscript, Univ. Copenhagen*.
- Calvo-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American economic review*, 94(3):426–454.
- Calvó-Armengol, A. and Jackson, M. O. (2007). Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, 132(1):27–46.
- Calvó-Armengol, A. and Zenou, Y. (2005). Job matching, social network and word-of-mouth communication. *Journal of urban economics*, 57(3):500–522.
- Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015.
- Carrillo-Tudela, C., Clymo, A., Comunello, C., Jäckle, A., Visschers, L., and Zentler-Munro, D. (2023). Search and reallocation in the covid-19 pandemic: Evidence from the uk. *Labour Economics*, page 102328.
- Carrillo-Tudela, C., Hobijn, B., She, P., and Visschers, L. (2016). The extent and cyclicity of career changes: Evidence for the uk. *European Economic Review*, 84:18–41.
- Carrillo-Tudela, C., Menzio, G., and Smith, E. (2011). Job search with bidder memories. *International Economic Review*, 52(3):639–655.
- Carrillo-Tudela, C. and Smith, E. (2017). Search capital. *Review of Economic Dynamics*, 23:191–211.
- Carrillo-Tudela, C. and Visschers, L. (2023). Unemployment and endogenous reallocation over the business cycle. *Econometrica*, 91(3):1119–1153.
- Carrillo-Tudela, C., Visschers, L., and Wiczer, D. (2022). Cyclical earnings, career and employment transitions. Working Paper No. 9990, CESifo.
- Casella, A. and Hanaki, N. (2008). Information channels in labor markets: On the resilience of referral hiring. *Journal of economic behavior & organization*, 66(3-4):492–513.
- Cingano, F. and Rosolia, A. (2012). People i know: job search and social networks. *Journal of Labor Economics*, 30(2):291–332.
- Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, 107(2):425–56.

- Dauth, W. and Eppelsheimer, J. (2020). Preparing the sample of integrated labour market biographies (siab) for scientific analysis: a guide. *Journal for Labour Market Research*, 54(1):1–14.
- Dustmann, C., Glitz, A., Schönberg, U., and Brücker, H. (2016). Referral-based job search networks. *The Review of Economic Studies*, 83(2):514–546.
- Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the german wage structure. *The Quarterly journal of economics*, 124(2):843–881.
- Eliason, M., Hensvik, L., Kramarz, F., and Skans, O. N. (2022). Social connections and the sorting of workers to firms. *Journal of Econometrics*.
- Fitzenberger, B., Lickleder, S., and Zwiener, H. (2015). Mobility across firms and occupations among graduates from apprenticeship. *Labour Economics*, 34:138–151.
- Forsythe, E. (2023). Occupational job ladders within and between firms. Discussion paper No. 16682, IZA.
- Galenianos, M. (2013). Learning about match quality and the use of referrals. *Review of Economic Dynamics*, 16(4):668–690.
- Galenianos, M. (2014). Hiring through referrals. *Journal of Economic Theory*, 152:304–323.
- Gathmann, C. and Schönberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1):1–49.
- Glitz, A. (2017). Coworker networks in the labour market. *Labour Economics*, 44:218–230.
- Glitz, A. and Vejlin, R. (2021). Learning through coworker referrals. *Review of Economic Dynamics*, 42:37–71.
- Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118.
- Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1360–1380.
- Groes, F., Kircher, P., and Manovskii, I. (2015). The u-shapes of occupational mobility. *The Review of Economic Studies*, 82(2):659–692.
- Guerrero, O. A. and Axtell, R. L. (2013). Employment growth through labor flow networks. *PloS one*, 8(5):e60808.
- Guvenen, F. (2011). Macroeconomics with heterogeneity: a practical guide. *Economic Quarterly*, 97(3Q):255–326.
- Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, 12(1):210–244.
- Gyetvai, A. and Zhu, M. (2025). Coworker networks and the role of occupations in job finding. *Labour Economics*, 94:102699.

- Heath, R. (2018). Why do firms hire using referrals? evidence from bangladeshi garment factories. *Journal of Political Economy*, 126(4):1691–1746.
- Hellerstein, J. K. and Neumark, D. (2020). Social capital, networks, and economic well-being. *The Future of Children*, 30(2020):127–152.
- Hensvik, L. and Skans, O. N. (2016). Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics*, 34(4):825–867.
- Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2024). Production and learning in teams. *Econometrica*, 92(2):467–504.
- Hethey-Maier, T. and Schmieder, J. F. (2013). Does the use of worker flows improve the analysis of establishment turnover evidence from german administrative data. *Journal of Contextual Economics–Schmollers Jahrbuch*, 133(4):477–510.
- Hong, L. and Lattanzio, S. (2022). The peer effect on future wages in the workplace. *Available at SSRN 4052587*.
- Horváth, G. (2014). Occupational mismatch and social networks. *Journal of Economic Behavior & Organization*, 106:442–468.
- Jarosch, G. (2023). Searching for job security and the consequences of job loss. *Econometrica*, 91(3):903–942.
- Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning from coworkers. *Econometrica*, 89(2):647–676.
- Kossinets, G. and Watts, D. J. (2009). Origins of homophily in an evolving social network. *American journal of sociology*, 115(2):405–450.
- Kugler, A. D. (2003). Employee referrals and efficiency wages. *Labour economics*, 10(5):531–556.
- Lazarsfeld, P. F. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18.
- Lengyel, B. and Eriksson, R. H. (2017). Co-worker networks, labour mobility and productivity growth in regions. *Journal of Economic Geography*, 17(3):635–660.
- Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, 84(2):718–789.
- Lindenlaub, I. and Postel-Vinay, F. (2023). Multidimensional sorting under random search. *Journal of political Economy*, 131(12):3497–3539.
- Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–2376.
- Longhi, S. and Brynin, M. (2010). Occupational change in britain and germany. *Labour Economics*, 17(4):655–666.
- Longhi, S. and Taylor, M. (2013). Occupational change and mobility among employed and unemployed job seekers. *Scottish Journal of Political Economy*, 60(1):71–100.

- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.
- Marin, A. (2012). Don't mention it: Why people don't share job information, when they do, and why it matters. *Social Networks*, 34(2):181–192.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.
- Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political economy*, 92(6):1086–1120.
- Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American economic review*, 81(5):1408–1418.
- Mortensen, D. T. and Vishwanath, T. (1994). Personal contacts and earnings: It is who you know! *Labour economics*, 1(2):187–201.
- Moscarini, G. and Vella, F. G. (2008). Occupational mobility and the business cycle. Technical report, National Bureau of Economic Research.
- Papageorgiou, T. (2018). Large firms and within firm occupational reallocation. *Journal of Economic Theory*, 174:184–223.
- Ponzo, M. and Scoppa, V. (2010). The use of informal networks in Italy: Efficiency or favoritism? *The Journal of Socio-Economics*, 39(1):89–99.
- Postel-Vinay, F. and Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350.
- Rege, M., Telle, K., and Votruba, M. (2012). Social interaction effects in disability pension participation: evidence from plant downsizing. *The Scandinavian Journal of Economics*, 114(4):1208–1239.
- Robin, J.-M., Postel-Vinay, F., and Kramarz, F. (2014). Occupational mobility and wage dynamics within and between firms. In *2014 Meeting Papers*, number 168 in Meeting Papers. Society for Economic Dynamics.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146.
- Saygin, P. O., Weber, A., and Weynandt, M. A. (2021). Coworkers, networks, and job-search outcomes among displaced workers. *ILR Review*, 74(1):95–130.
- Schmidtlin, L., Seth, S., Vom Berge, P., et al. (2020). Sample of integrated employer employee data (SIEED) 1975-2018. Technical report, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].
- Taber, C. and Vejlín, R. (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica*, 88(3):1031–1069.
- Topa, G. (2011). Labor markets and referrals. In *Handbook of social economics*, volume 1, pages 1193–1221. Elsevier.

- Topa, G. (2019). Social and spatial networks in labour markets. *Oxford Review of Economic Policy*, 35(4):722–745.
- Zaharieva, A. (2013). Social welfare and wage inequality in search equilibrium with personal contacts. *Labour Economics*, 23:107–121.
- Zaharieva, A. (2018). On the optimal diversification of social networks in frictional labour markets with occupational mismatch. *Labour Economics*, 50:112–127.

Appendix A: Chapter 1

A.1 Additional descriptive statistics

Table A.1: Summary statistics; worker, coworker, IV samples

	Workers	Coworkers	IV
Number of workers	41,300	592,085	437,855
Share female	0.45	0.47	0.47
Share German	0.89	0.85	0.84
Age (in 2015)			
Mean	39.6	38.8	37.5
10th percentile	27	22	22
50th percentile	39	37	35
90th percentile	52	58	56
Average daily wage (2012-2017)			
Mean	89.7	69.2	65.4
10th percentile	30.2	14.4	14.2
50th percentile	74.3	58.1	55.8
90th percentile	159.7	126.5	118

Notes: The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. Coworkers are defined as those who worked in the same establishment at the same time for at least 30 days in the last 3 years. The instrumental variable is the number of coworkers with same occupation who were hired after the worker.

Table A.2: Summary statistics; year, location and industry

Observations	41,300
Year (%)	
2012	15.8
2013	16.3
2014	17.0
2015	16.5
2016	17.1
2017	17.2
Location of origin establishment (%)	
Schleswig-Holstein	3.2
Hamburg	2.9
Niedersachsen	9.1
Bremen	1.1
Nordrhein-Westfalen	20.7
Hessen	7.5
Rheinland-Pfalz	4.2
Baden-Wuerttemberg	12.2
Bayern	14.5
Saarland	1.3
Berlin	5.0
Brandenburg	3.2
Mecklenburg-Vorpommern	2.0
Sachsen	5.9
Sachsen-Anhalt	3.2
Thueringen	3.9
Industry of origin establishment (%)	
Agriculture	0.9
Mining	0.3
Manufacturing	16.8
Utilities	1.3
Construction	4.5
Commerce	18.8
Hospitality	4.6
Transport	5.9
Finance	2.1
Real estate	21.7
Public administration	2.5
Education	2.7
Social services	13.1
Other services	4.8

Notes: The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

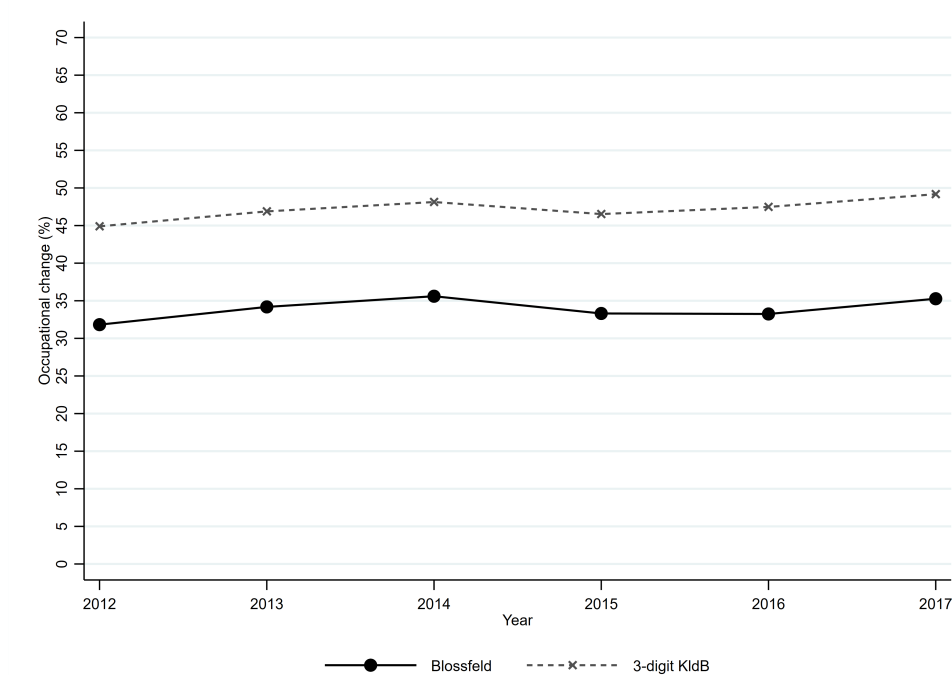


Figure A.1: Share of job transitions resulting in occupational change, by year

Notes: Author's calculations using SIEED. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

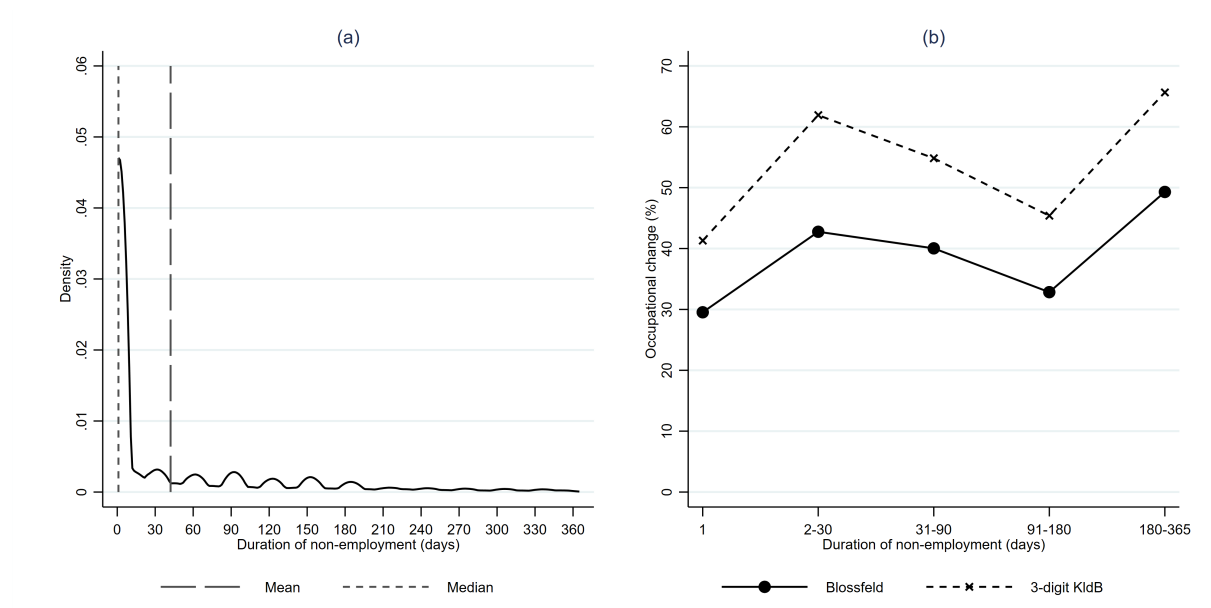


Figure A.2: Duration of non-employment and occupational change

Notes: Author's calculations using SIEED. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

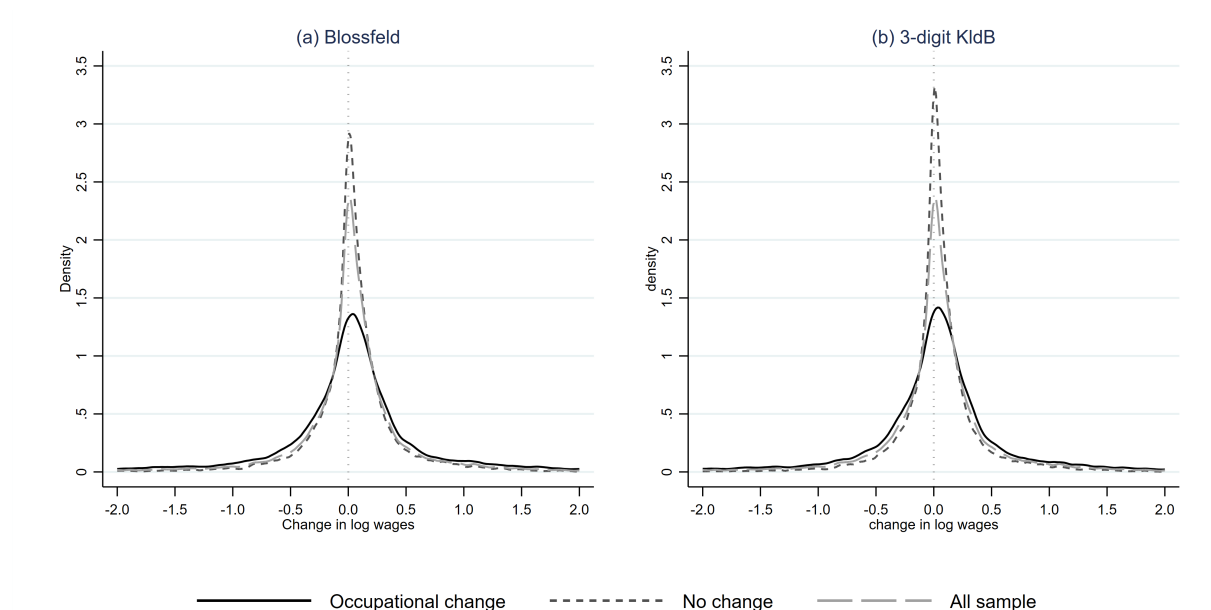


Figure A.3: Distribution of change in log wages

Notes: Author's calculations using SIEED. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

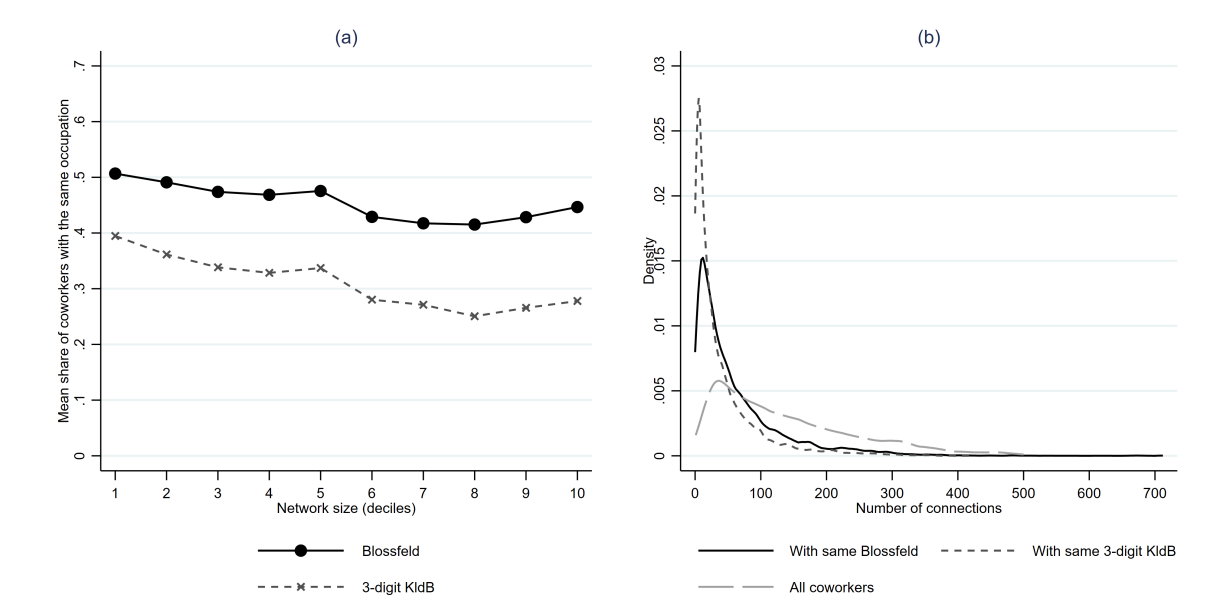


Figure A.4: Degree and homophily distributions

Notes: Author's calculations using SIEED. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment.

A.2 Occupational definitions

Table A.3: KldB–Blossfeld correspondence; administration occupations

Unskilled commercial and administrative	Skilled commercial and administrative		Managers
Salespersons	Managers in agriculture and animal b	Publicity occupations	Entrepreneurs, managing directors,
Commercial agents, travellers	Wholesale and retail trade buyers,	Cost accountants, valuers	Management consultants, organisers
Post masters	Publishing house dealers, bookselle	Accountants	Chartered accountants, tax advisers
Postal deliverers	Bank specialists	Data processing specialists	Members of Parliament, Ministers, e
Telephonists	Building society specialists	Office specialists	Senior government officials
Cashiers	Health insurance specialists (not s		Association leaders, officials
Stenographers, shorthand-typists, t	Life, property insurance specialist		
Data typists	Forwarding business dealers		
Office auxiliary workers	Tourism specialists		

Notes: The conversion code is provided by Dauth and Eppelsheimer (2020).

Table A.4: KldB–Blossfeld correspondence; service occupations

Unskilled services	Skilled services	Semiprofessions	Professions
Pharmacy aids	Druggists/chemists (pharmacy)	Journalists	Arbitrators
Service-station attendants	Brokers, property managers	Interpreters, translators	Legal representatives, advisors
Mobile traders	Landlords, agents, auctioneers	Librarians, archivists, museum spec	Physicians
Cash collectors, cashiers, ticket s	Railway engine drivers	Nurses, midwives	Dentists
Other traffic controllers, conducto	Railway controllers, conductors	Social workers, care workers	Veterinary surgeons
Motor vehicle drivers	Soldiers, border guards, police off	Home wardens, social work teachers	Pharmacists
Coachmen	Firefighters	Work, vocational advisers	University teachers, lecturers at h
Street attendants	Safety testers	Nursery teachers, child nurses	Gymnasium teachers
Deck seamen	Chimney sweeps	Primary, secondary (basic), special	Economic and social scientists, sta
Inland boatmen	Judicial administrators	Technical, vocational, factory inst	Humanities specialists, n.e.c.
Other water transport occupations	Judicial enforcers	Music teachers, n.e.c.	Ministers of religion
Warehouse managers, warehousemen	Musicians	Sports teachers	
Transportation equipment drivers	Artists' agents	Other teachers	
Stowers, furniture packers	Visual, commercial artists		
Stores, transport workers	Scenery, sign painters		
Factory guards, detectives	Artistic and assisting occupations		
Watchmen, custodians	Interior, exhibition designers, win		
Doormen, caretakers	Photographers		
Domestic and non-domestic servants	Non-medical practitioners		
Health-protecting occupations	Masseurs, physiotherapists and rela		
Performers, professional sportsmen,	Nursing assistants		
Restaurant, inn, bar keepers, hotel	Dietary assistants, pharmaceutical		
Waiters, stewards	Medical receptionists		
Others attending on guests	Members of religious orders without		
Other housekeeping attendants	Religious care helpers		
Employees by household cheque proce	Hairdressers		
Laundry workers, pressers	Other body care occupations		
Textile cleaners, dyers and dry cle	Housekeeping managers		
Household cleaners	Consumer advisors		
Glass, buildings cleaners			
Street cleaners, refuse disposers			
Vehicle cleaners, servicers			
Machinery, container cleaners and r			

Notes: The conversion code is provided by Dauth and Eppelsheimer (2020).

Table A.5: KldB–Blossfeld correspondence; production occupations

Agriculture	Unskilled Manual		
Farmers	Miners	Other metal-cutting occupations	Meat, sausage goods makers
Winegrowers	Mechanical, electrical, face workers	Metal polishers	Fish processing operatives
Animal breeders	Stone crushers	Engravers, chasers	Ready-to-serve meals, fruit, vegeta
Fishermen	Earth, gravel, sand quarriers	Metal finishers	Other beverage makers, tasters
Land workers	Oil, natural gas quarriers	Galvanisers, metal colourers	Tobacco goods makers
Milkers	Mineral preparers, mineral burners	Enamellers, zinc platers and other	Milk, fat processing operatives
Family-member land workers, n.e.c.	Stone preparers	Welders, oxy-acetylene cutters	Flour, food processors
Animal keepers and related occupatio	Jewel preparers	Solderers	Sugar, sweets, ice-cream makers
Gardeners, garden workers	Stoneware, earthenware makers	Riveters	Concrete workers
Florists	Shaped brick, concrete block makers	Metal bonders and other metal conne	Roofers
Forestry managers, foresters, hunter	Ceramics workers	Container builders, coppersmiths an	Scaffolders
Forest workers, forest cultivators	Frit makers	Pipe, tubing fitters	Paviors
	Hollow glassware makers	Precision fitters n.e.c.	Road makers
	Flat glass makers	Electric motor, transformer fitters	Tracklayers
	Glass processors, glass finishers	Electrical appliance, electrical pa	Land improvement, hydraulic enginee
	Chemical plant operatives	Other assemblers	Other civil engineering workers
	Rubber makers, processors	Metal workers (no further specifica	Building labourer, general
	Plastics processors	Spoolers, twisters, ropemakers	Earth movers
	Paper, cellulose makers	Weaving preparers	Other building labourers, building
	Packaging makers	Weavers	Insulators, proofers
	Other paper products makers	Tufted goods makers	Screed, terrazzo layers
	Copiers	Machined goods makers	Other wood and sports equipment mak
	Printer's assistants	Felt makers, hat body makers	Goods painters, lacquerers
	Wood preparers	Textile processing operatives (brai	Wood surface finishers, veneerers
	Wood moulders and related occupatio	Clothing sewers	Ceramics, glass painters
	Wood products makers	Laundry cutters, sewers	Goods examiners, sorters, n.e.c.
	Iron, metal producers, melters	Embroiderers	Packagers, goods receivers, despatc
	Rollers	Hat, cap makers	Assistants (no further specificatio
	Metal drawers	Sewers, n.e.c.	Other machinists
	Semi-finished product fettlers and	Other textile processing operatives	Crane drivers
	Sheet metal pressers, drawers, stam	Textile dyers	Earthmoving plant drivers
	Wire moulders, processors	Textile finishers	Construction machine attendants
	Other metal moulders (non-cutting d	Leather makers, catgut string maker	Machine attendants, machinists' hel
	Drillers	Footwear makers	Stokers
	Planers	Fine leather goods makers	Machine setters (no further specifi
	Borers	Leather clothing makers and other l	
	Metal grinders	Hand shoemakers	
Skilled Manual		Technicians	Engineers
Glass blowers (lamps)	Spinners, fibre preparers	Dental technicians	Agricultural engineers, agriculture
Chemical laboratory workers	Cutters	Ophthalmic opticians	Garden architects, garden managers
Vulcanisers	Shoemakers	Mechanical engineering technicians	Mechanical, motor engineers
Book binding occupations	Coarse leather goods finishers, tru	Electrical engineering technicians	Electrical engineers
Type setters, compositors	Skin processing operatives	Building technicians	Architects, civil engineers
Printed goods makers	Bakery goods makers	Measurement technicians	Survey engineers
Printers (letterpress)	Confectioners (pastry)	Mining, metallurgy, foundry technic	Mining, metallurgy, foundry enginee
Printers (flat, gravure)	Butchers	Chemistry, physics technicians	Other manufacturing engineers
Special printers, screeners	Cooks	Remaining manufacturing technicians	Other engineers
Moulders, coremakers	Wine coopers	Other technicians	Chemists, chemical engineers
Mould casters	Brewers, maltsters	Foremen, master mechanics	Physicists, physics engineers, math
Turners	Bricklayers	Biological specialists	Air transport occupations
Steel smiths	Carpenters	Physical and mathematical specialis	Scientists n.e.c.
Sheet metal workers	Explosives men (except shotfiring)	Chemical laboratory assistants	
Plumbers	Stucco workers, plasterers, rough c	Photo laboratory assistants	
Locksmiths, not specified	Tile setters	Technical draughtspersons	
Building fitters	Furnace setter, air heating install	Navigating ships officers	
Sheet metal, plastics fitters	Glaziers	Technical ships officers, ships eng	
Engine fitters	Room equippers	Radio operators	
Plant fitters, maintenance fitters	Upholsterers, mattress makers	Medical laboratory assistants	
Steel structure fitters, metal ship	Carpenters		
Motor vehicle repairers	Model, form carpenters		
Agricultural machinery repairers	Cartwrights, wheelwrights, coopers		
Aircraft mechanics	Painters, lacquerers (construction)		
Precision mechanics	Generator machinists		
Other mechanics	Winding engine drivers, aerial rope		
Watch-, clockmakers	Radio, sound equipment mechanics		
Toolmakers	Electrical appliance fitters		
Precious metal smiths	Telecommunications mechanics, craft		
Musical instrument makers	Electrical fitters, mechanics		
Doll makers, model makers, taxiderm			

Notes: The conversion code is provided by Dauth and Eppelsheimer (2020).

A.3 Variable construction

A.3.1 Mapping coworker networks

The SIEED dataset is structured with entries $\{N, f, l\}$, where $i \in N = \{1, \dots, n\}$ represents workers, $f = \{f_1, \dots, f_F\}$ establishments, and $l = \{l_1, \dots, l_N\}$ spells, where $l_i = \{l_{i1}, \dots, l_{i\bar{L}_i}\}$, and $\bar{L}_i$ is the final spell of worker i . Each spell l_i is characterised by a starting day $\underline{t}_{li}$ and an end day $\bar{t}_{li}$, such that $\underline{t}_{li} < \bar{t}_{li}$, drawn from $t \in \mathbb{N}$, meaning that t is discrete and each element represents a day. Originally, multiple worker spells can have the same establishment even if they are consecutive. I collapse these spells into one as to have each spell representing a continuous employment episode in an establishment.

Using this format, it is possible to construct $\bar{T} - \underline{T}$ networks, i.e., a coworker network G_t for each day within the period of analysis. Let $G_t = \{N, g_t\}$ be a graph where the nodes N is the set of workers. Let g_t be a $n \times n$ unweighted and undirected adjacency matrix, where $g_{ij,t} = 1$ if two workers are connected in t , and 0 otherwise.

To define $g_{ij,t} = 1$, let $B : f \rightarrow l \times N$ be a function that maps establishments to workers and spells. Worker i is connected to worker j , in a spell $\hat{l}$, with $\underline{t}_{\hat{l}} \leq t \leq \bar{t}_{\hat{l}}$ if $\exists l_j$ and $l_i \leq \hat{l}$ such that

- (1) $B(i, l_i) = B(j, l_j)$; i and j worked in the same establishment
- (2) $\exists l_i, l_j$ such that $\min\{\bar{t}_{li}, \bar{t}_{lj}\} - \max\{\underline{t}_{li}, \underline{t}_{lj}\} + 1 > 30$; their spells in the establishment overlapped for at least 30 days
- (3) for l_i, l_j satisfying (1) and (2), $\underline{t}_{\hat{l}} - 1096 \leq \min\{\bar{t}_{lj}, \bar{t}_{li}\}$; the overlap occurred within 3 years in the past

Each g_t is formed by accumulating the information from $g_{\underline{T}}, \dots, g_{t-1}$ and removing the links where condition (3) is not satisfied. If $B(i, \hat{l}) = B(j, l_j)$, i and j are current coworkers, and for $l_i < \hat{l}$, they are former coworkers.

With this procedure, I create adjacency lists, where two workers are connected if they work or worked in the same establishment at the same time for at least 30 days

in the past 1096 days (3 years). Then, I select the links in networks G_t to construct the homophily variables; these are the of worker i 's links that exist on their last day in the establishment.

A.3.2 Wage imputation

Wages in the SIEED dataset above the social security contribution ceiling are top coded. In my sample, only 3% of the observations are censored, but the incidence of top coding is higher for male and educated workers. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009) and Card et al. (2013). I run the procedure for all workers observed in the SIEED dataset after 2004, and the models are estimated for each cell grouping high/medium/low level of education, East/West Germany, and year. The reported daily wages in SIEED are nominal; I deflate the values using CPI and run the procedure on real wages.

The first step fits a Tobit log wage equation for each cell, controlling for Gender, age (squared), a dummy for workers above the age of 40, tenure in establishment (squared), experience in occupation (squared), Blossfeld occupation, German and part-time. I then replace the censored log wages with $X\beta + \sigma\Phi^{-1}[k + u(1 - k)]$, where $X\beta$ are the predicted values, σ is the standard deviation of the residual, Φ is the standard normal density function, u is a random draw from a uniform distribution ranging from zero to one, $k = \Phi(\frac{c - X\beta}{\sigma})$, and c is the censoring point.

In the second step, I take the estimates from the first step and calculate leave-one-out wage means for workers and establishments. The worker leave-one-out mean is the average wage of the worker over the years excluding the observation in question. It is similar to worker fixed effects. The establishment leave-one-out mean is the average wage in the establishment excluding the worker in question. It is similar to establishment fixed effects. I then include repeat the first step procedure including the leave-one-out means in the controls. For more details, see Dauth and Eppelsheimer (2020).

A.4 Additional results

Table A.6: Occupational homophily and occupational change; heterogeneity by year

	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)
Occupational homophily	-0.266*** (0.081)	-0.167* (0.091)	-0.222*** (0.084)	-0.245*** (0.087)	-0.258*** (0.074)	-0.185** (0.079)
ln(degree)	0.023 (0.024)	-0.016 (0.024)	0.004 (0.020)	-0.041** (0.021)	0.018 (0.023)	-0.013 (0.022)
FS F-stat	114.874	38.387	135.603	87.250	105.018	94.343
Observations	6525	6746	7034	6825	7051	7119

Notes: Dependent variable is the probability of changing occupation. Estimated with 2SLS following the specification in column (5) Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The instrumental variables is the number of coworkers with the same occupation who were hired after the worker in question (entrants). Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Occupational homophily and occupational change; heterogeneity by network size, duration of non-employment, and establishment size change

	Number of coworkers			Duration of non-employment		Establishment size change		
	1st tercile (1)	2nd tercile (2)	3rd tercile (3)	< 30 days (4)	≥ 30 days (5) * (5)	< -0.3 (6)	-0.3 ≤ ≤ 0.3 (7)	> 0.3 (8)
Occupational homophily	-0.157*** (0.059)	-0.228*** (0.032)	-0.235*** (0.045)	-0.209*** (0.046)	-0.246*** (0.076)	0.059 (0.375)	-0.262*** (0.041)	-0.089 (0.325)
ln(degree)	-0.017 (0.016)	-0.036 (0.029)	-0.031 (0.026)	-0.008 (0.011)	0.014 (0.016)	-0.031 (0.040)	0.004 (0.010)	0.060 (0.059)
FS F-stat	770.837	1427.046	269.388	142.716	76.939	23.169	150.678	9.353
Observations	13924	13661	13715	28453	12847	3748	36163	1389

Notes: Dependent variable is the probability of changing occupation. Estimated with 2SLS following the specification in column (5) Table 1.3. Sample of worker aged 24-55 who changed establishment between 2012-2017. They only worked panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The instrumental variables is the number of coworkers with the same occupation who were hired after the worker in question (entrants). Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Occupational homophily and Blossfeld occupational change; heterogeneity by skill and education

	Occupation skill		Education		
	Blue Collar (1)	White Collar (2)	Low (3)	Medium (4)	High (5)
Occupational homophily	-0.288*** (0.054)	-0.170*** (0.038)	-0.305** (0.132)	-0.188*** (0.034)	-0.340*** (0.058)
ln(degree)	-0.020** (0.009)	0.003 (0.011)	-0.004 (0.023)	-0.012 (0.008)	-0.022 (0.018)
FS F-stat	1145.410	1029.303	187.611	1712.891	299.775
Observations	24326	16974	5556	28518	7226

Notes: Dependent variable is the probability of changing the Blossfeld occupation. Estimated with 2SLS following the specification in column (5) of Panel A, Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. The instrumental variables is the number of coworkers with the same occupation who were hired after the worker in question (entrants). Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Occupational homophily and occupational change; alternative specifications to occupational homophily

	Number		logs	
	(1)	(2)	(3)	(4)
Panel A: Blossfeld				
# of coworkers with the same occupation	-0.0007*** (0.0000)	-0.0004*** (0.0001)		
# of coworkers with a different occupation	0.0005*** (0.0000)	0.0003*** (0.0001)		
log(# of coworkers with the same occupation)			-0.0564*** (0.0019)	-0.0459*** (0.0046)
log(# of coworkers with a different occupation)			0.0581*** (0.0018)	0.0089 (0.0079)
Observations	41235	41235	40670	40670
Number of establishments		3301		3292
Panel B: 3-digit KldB				
# of coworkers with the same occupation	-0.0009*** (0.0000)	-0.0005*** (0.0001)		
# of coworkers with a different occupation	0.0004*** (0.0000)	-0.0000 (0.0001)		
log(# of coworkers with the same occupation)			-0.0438*** (0.0019)	-0.0203*** (0.0037)
log(# of coworkers with a different occupation)			0.0585*** (0.0020)	-0.0144 (0.0092)
Observations	41235	41235	39570	39570
Number of establishments		3301		3292
Controls		X		X
<i>Fixed effects</i>				
Occupation×year		X		X
State×year		X		X
Industry×state×year		X		X
Establishment		X		X

Notes: Dependent variable is the probability of changing occupation. Estimated with OLS following the specification in column (6) Table 1.3. The sample is workers aged 24-55 who changed establishment between 2012-2017, worked only in panel establishments in the last three years, and found a new job within one year of leaving the origin establishment. Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Chapter 2

B.1 Mathematical appendix

B.1.1 Monotonicity of the value of employment in productivity

In this section, I show $V(1, p, p)$ is increasing in p under the following three assumptions

Assumption 1. $V(1, p, p)$ is differentiable.

Assumption 2. Let $\xi(V(1, p, p))$ be the arrival rate at firm p at the steady state, and let $\xi(V(1, p, p))$ be strictly monotonic and differentiable in $V(1, p, p)$.

Assumption 3. the worker moves from p to $\tilde{p}$ if the value of employment in $\tilde{p}$ is higher.

Following assumption 1, we can write a more general Bellman equation for employment in firm p at wage p as

$$(r + \delta)V(1, p, p) = p + \xi(V(1, p, p))\beta \int_{V(1, p, p)}^{\bar{V}} \bar{F}_V(x) dx + \delta \bar{U} \quad (\text{B.1})$$

Where $F_V(\cdot)$ is the cumulative distribution of firm values $V(1, p, p) \in [\underline{V}, \bar{V}]$ at the steady state.

Proposition 1. Under assumptions 1,2, and 3 the value of employment is strictly increasing with productivity

Proof.

Lemma 1. Under assumptions 1,2, and 3 $V(1, p, p)$ is strictly monotonic.

Proof. The first step is to show that $V(1, p, p)$ is one-to-one injective. By way of contradiction, assume $V(1, p, p)$ is not one-to-one injective; there exists some p and $\tilde{p}$ such that $p \neq \tilde{p}$ and $V(1, p, p) = V(1, \tilde{p}, \tilde{p})$. But by equation B.1 we see that $V(1, p, p) = V(1, \tilde{p}, \tilde{p})$ holds only when $p = \tilde{p}$. Which is the contradiction implying $V(1, p, p)$ is one-to-one injective. Since $V(1, p, p)$ is differentiable, thus continuous, by assumption 1, and one-to-one injective, then it must be strictly monotonic. $\square$

Notice that by equation B.1 and given $V'(1, p, p)$ is differentiable (assumption 1), we have

$$V'(1, p, p) = \frac{1}{r + \delta - \frac{\partial \xi(V(1, p, p))}{\partial V(1, p, p)} \beta \int_{V(1, p, p)}^{\bar{V}} \bar{F}_V(x) dx + \beta \xi(V(1, p, p)) \bar{F}_V(V(1, p, p))} \quad (\text{B.2})$$

Since we know that $V(1, p, p)$ is strictly monotonic, it is sufficient to show that $V'(1, p, p) > 0$ at some p to imply that $V'(1, p, p) > 0$ everywhere. Let p^* be such be such that $V(1, p^*, p^*) = \bar{V}$, then $V'(1, p^*, p^*) = \frac{1}{r + \delta} > 0$. Therefore, $V'(1, p, p) > 0 \forall p$. $\square$

B.1.2 Derivation of the value of employment

Recall $V(\epsilon, \epsilon p, p)$ is given by

$$(r + \delta)V(\epsilon, \epsilon p, p) = \epsilon p + \xi(p)\beta \int_p^{\bar{p}} V'(\epsilon, \epsilon x, x) \bar{F}(x) dx + \delta \bar{U}(\epsilon) \quad (\text{B.3})$$

Now let $\omega(\epsilon, p) = \int_p^{\bar{p}} V'(\epsilon, \epsilon x, x) \bar{F}(x) dx$. Differentiate with respect to p , rearrange, and integrate from $\bar{p}$ to p yielding

$$V(\epsilon, \epsilon p, p) = V(\epsilon, \epsilon \bar{p}, \bar{p}) + \int_p^{\bar{p}} \frac{\omega'(\epsilon, x)}{\bar{F}(x)} dx \quad (\text{B.4})$$

Replace $V(\epsilon, \epsilon p, p)$ with the above expression, then differentiate with respect to p and rearrange. $\omega(\epsilon, x)$ satisfies the following first order linear differential equation

$$\omega'(\epsilon, p) + \frac{\beta \zeta'(p) \bar{F}(p)}{r + \delta + \beta \zeta(p) \bar{F}(p)} \omega(\epsilon, p) = -\epsilon \frac{\bar{F}(p)}{r + \delta + \beta \zeta(p) \bar{F}(p)} \quad (\text{B.5})$$

Where the initial condition is $\omega(\epsilon, \bar{p}) = 0$. Let $T'(p) = \frac{\beta \zeta'(p) \bar{F}(p)}{r + \delta + \beta \zeta(p) \bar{F}(p)}$, then the integrating factor is $e^{T(p)}$, where $T(p) = -\int_p^{\bar{p}} \frac{\beta \zeta'(y) \bar{F}(y)}{r + \delta + \beta \zeta(y) \bar{F}(y)} dy$. Then the solution to the differential equation is

$$\omega(\epsilon, p) = \epsilon e^{-T(p)} \int_p^{\bar{p}} \frac{\bar{F}(x)}{r + \delta + \beta \zeta(x) \bar{F}(x)} e^{T(x)} dx \quad (\text{B.6})$$

Then replace $\omega(\epsilon, p)$ back in equation B.3 to arrive at the closed form solution to $V(\epsilon, \epsilon p, p)$.

B.2 Estimation procedure

I simulate an economy of 100,000 workers For 50 periods. The number of firms satisfies the ratio of workers to firms in each education group. In the initial period we start at the steady state. A draw from the employment distribution $E(p)$, gives the position of the firm on the productivity distribution $F(p)$ (by inverting equation 2.17). A draw from the wage distribution will give the position of the poaching firm on the productivity distribution (inverting equation 2.23).

To estimate the parameters of interest, I follow the global minimisation procedure following Guvenen (2011) and Arnoud et al. (2019) which goes as follows.

First, create $N \times K$ vector of quasi-random numbers – from a Sobol sequence. Where $N = 40$ and K is the number of parameters to estimate. Then, I evaluate the function N times and choose the 10% best guesses. Then best guesses are ordered from best to worst, and iterate over the guesses. In each iteration, the j th row is chosen as initial values for the local minimisation (Nelder–Mead method). Take the output from local minimisation as a guess for the next iteration with the following update rule:

$x_{j+1} = \theta x_j^* + (1 - \theta)_{j+1}$. Where a x_j^* is the best outcome of the local minimisation as of j th iteration.

Now, we turn to the calculation of the standard errors in table 2.3. The variance-covariance matrix Σ is

$$\Sigma = (J'WJ)^{-1}J'WSWJ(J'WJ)^{-1} \quad (\text{B.7})$$

Where $J = \frac{\partial M_s(\hat{P})}{\partial \hat{P}}$, S is a matrix with the variance of the data moments on the diagonals and zero otherwise. The standard errors are square root of the diagonal.

B.3 Additional tables and figures

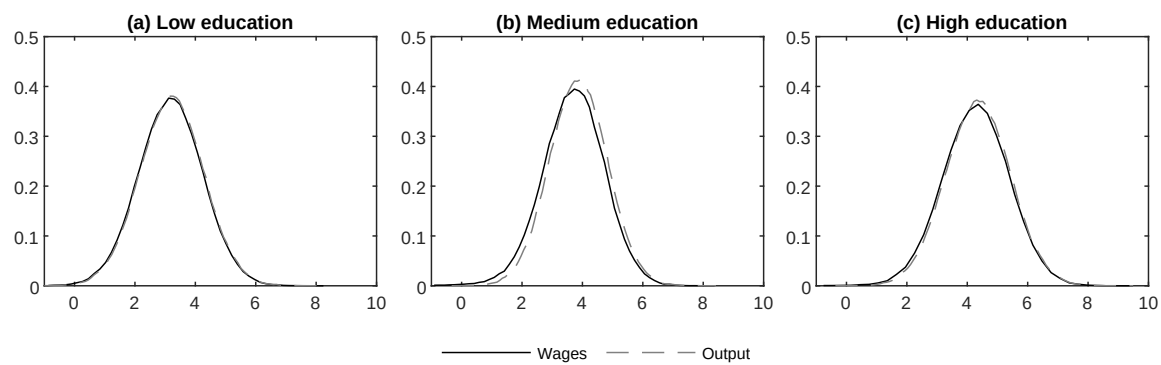


Figure B.1: Log wage and log output distribution

Notes: Simulated data.

Appendix C: Chapter 3

C.1 Sequential auctions framework

In the baseline model in Section 3.2, I assumed that agents are myopic for simplicity. This makes the model block recursive; the worker problem and the steady state can be solved separately. However, the shortcoming of this baseline is that it does not allow for identical workers with identical jobs to have different wages. Another way to achieve this while keeping the worker's dynamic optimization problem is to follow the sequential auction framework (Postel-Vinay and Robin, 2002). In this section, I will present this alternative setup.

Let's consider the case of homogenous workers, but heterogenous occupations and firms. Let r be the discount rate. Assume the firms hold all the bargaining power. When an unemployed worker comes across a job, the firm offers them a wage such that the value of employment is equal to the value of unemployment. The value of unemployment of a worker who lost their job y is $U(y) = U = \frac{z}{r}$. The value of unemployment is constant for all workers since any new employer will take all the surplus.

Let $V(w, y)$ be the value of employment with a wage w at a job y . When an employed worker in y meets a firm q , there are two cases. If the marginal productivity of the job q is higher than the job y , the worker moves to y and is paid a wage $R(q, y)$, such that $V(R(q, y), q) = V(y, y)$. If the job the worker encounters is less productive than y but more productive than a threshold $q(w, y)$, the worker stays, but negotiates a new wage $R(y, q)$, such that $V(R(y, q), y) = V(q, q)$. The Bellman equation for the employed is

$$\begin{aligned}
rV(w, y) = & w + (1 - \alpha)\lambda_e \int_{q(w, y)}^y V(x, x) - V(w, y) d\Phi(x) \\
& + (1 - \alpha)\lambda_e(1 - \Phi(y))(V(y, y) - V(w, y)) \\
& + \alpha\lambda_e(V(y, y) - V(w, y)) \\
& + \delta(\bar{U} - V(w, y))
\end{aligned} \tag{C.1}$$

This equation is similar to that of a standard Postel-Vinay and Robin (2002) model, with one difference. The term $\alpha\lambda_e(V(y, y) - V(w, y))$ is the result of the network. The arrival rate of a job q through networks, is $\alpha\lambda_e \frac{\phi(q)}{1 - \Phi(y)}$. We get the network expression by integrating over q from y to $\bar{y}$.

We can then use the wage setting conditions to solve for the wage of a worker employed at y who met q ;

$$\begin{aligned}
R(y, q) = & q - \frac{\lambda_e(1 - \alpha)}{r + \delta} \int_q^y 1 - \Phi(x) dx - (y - q) \frac{\lambda_e \alpha}{r + \delta} \\
= & q - \frac{\lambda_e}{r + \delta} \int_q^y \alpha + (1 - \alpha)(1 - \Phi(x)) dx
\end{aligned} \tag{C.2}$$

We see this setup allows for frictional wage dispersion. The worker will be willing to take wage cuts in exchange for future wage growth opportunities. Otherwise, the model is identical to the baseline.

C.2 Additional tables

Table C.1: Wage regression results

Dependent: Log daily wage	Coefficients
Task	
Unskilled	-0.466*** (0.001)
Skilled	-0.300*** (0.001)
Education	
Medium	-0.003*** (0.001)
High	0.262*** (0.001)
Task × Education	
Unskilled × Medium	0.027*** (0.002)
Unskilled × High	-0.236*** (0.002)
Skilled × Medium	0.053*** (0.001)
Skilled × High	-0.043*** (0.001)
R-squared	0.726
Observations	11,489,997
Number of establishments	27256
Establishment fixed effects	X
Controls	X

Notes: Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990. It includes only full-time workers subject to social security. Controls are the interaction between gender and year of birth. Wages above the social security contribution ceiling are top coded. I impute the censored observation using the code in Dauth and Eppelsheimer (2020), following the two-step procedure in Dustmann et al. (2009); Card et al. (2013). See Appendix A.3.2 for the imputation details. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Counterfactual exercise; dedicate search effort to general channels, by level of education

	(1) Baseline	(2) J2J $\alpha = 0$	(3) N2E $\alpha_u = 0$	(4) J2J & N2E $\alpha = 0 \text{ \& } \alpha_u = 0$
Low				
Average log wage	-1.209	-1.341	-1.391	-1.479
Var log wage	0.257	0.158	0.199	0.142
J2J rate	0.012	0.007	0.014	0.009
N2E rate	0.121	0.120	0.118	0.121
Task distribution				
Unskilled	0.166	0.201	0.203	0.232
Skilled	0.526	0.574	0.580	0.611
Complex	0.308	0.225	0.218	0.157
Medium				
Average log wage	-1.180	-1.308	-1.357	-1.443
Var log wage	0.242	0.150	0.191	0.138
J2J rate	0.011	0.007	0.013	0.009
N2E rate	0.119	0.118	0.117	0.119
Task distribution				
Unskilled	0.157	0.177	0.196	0.225
Skilled	0.568	0.623	0.612	0.634
Complex	0.275	0.200	0.192	0.141
High				
Average log wage	-1.004	-1.142	-1.198	-1.292
Var log wage	0.284	0.179	0.225	0.163
J2J rate	0.011	0.007	0.013	0.009
N2E rate	0.118	0.117	0.116	0.118
Task distribution				
Unskilled	0.103	0.118	0.144	0.164
Skilled	0.560	0.626	0.616	0.666
Complex	0.337	0.256	0.240	0.170

Notes: Except for the network parameters, the model parameters are the estimates from Table 3.4. Author's calculations using SIEED. The sample covers the years 2014-2017 and is composed of German workers born between 1955 and 1990.