On the Effects of Career Choice: Matching Efficiency of Different Occupations and Education Levels*

René Fahr
IZA, University of Bonn

Uwe Sunde
IZA, University of Bonn

Abstract

This paper investigates the differences in the matching process of job seekers and vacancies to be filled between different educational and occupational groups. To investigate this issue, matching functions are estimated across different occupations and educational cohorts, that is, on an even lower level of aggregation than previously investigated in the literature, and along different dimensions. In order to rule out spurious results, the panel data are tested for stationarity applying novel panel testing techniques based on bootstrapping methods. The data used also allow to distinguish inflows into jobs by previous job status of new hires.

JEL-classification: C15, C2, C3, E24, J21, J41, J42, J62, J63

Keywords: Matching, empirical matching functions, vacancies, unemployment, occupational structure, education, stationarity in panels

*We are indebted to Hielke Buddelmeyer, Melvyn Coles, and Monika Merz for very helpful comments on an earlier draft. Contact address: IZA, Schaumburg-Lippe Str. 7, 53113 Bonn, Tel. +49-0228-3894-533, email: fahr@iza.org, sunde@iza.org. All errors are of course our own.

1 Introduction

Labor market research has made extensive use of the concept of a matching function in recent years. In particular, the matching function appears as a shortcut to introduce frictions on the labor market in models investigating equilibrium unemployment. However, surprisingly enough, the bulk of research to date was theoretical. Moreover, the small existing literature investigating the empirical relevance of the concept rarely goes beyond the aggregate level. Consequently, the matching function is rarely used as an empirical concept to find out more about the structure of labor markets, in particular at less aggregate levels, which would allow to compare the outcomes on different labor markets. This paper will argue that empirical matching functions can serve as a helpful instrument to get more information about the functioning of labor markets, in particular on disaggregate levels.

Most of the empirical literature on matching functions supports the relevance of the concept also in the data. However, little is known about sectoral or occupational differences with respect to matching and job creation and the closely related questions about labor market efficiency, relative supply shortages of specific skills and the like, issues which are particularly relevant questions for individuals when planning their career as well as for politicians in designing labor market or migration policies. This paper is a first attempt to use the concept of a matching function to find out more about the structure of labor markets and thus to fill this gap.

The findings suggest that the relation between new hirings and stocks of job seekers and vacancies described by the matching function is indeed empirically relevant. The presumption of constant returns to scale of the aggregate matching function is not confirmed by the data. Not surprisingly, the findings for disaggregate matching functions differ substantially from aggregate results. Labor markets for different occupations seem to exhibit fundamentally different structures concerning the creation of new jobs. These differences are even more pronounced once labor markets are defined by different educational attainment instead of occupation. While some markets exhibit constant returns to scale, others are characterized by increasing or decreasing returns to scale. All in all, the results obtained from looking at disaggregate levels provide a very detailed picture of the functioning of labor markets. In particular with respect to policy advice, the approach followed in this paper is therefore superior to previous evaluations of empirical matching functions.

The following section briefly presents the theoretical concept of a matching function and surveys some related literature. Section 3 presents the empirical strategy followed in this paper and discusses some econometric issues of importance, in particular the procedure used to test for stationarity of the panels used. Section 4 contains a detailed description of the data used in the empirical investigation and addresses interesting features of the data. In section 5 we present results of estimations of aggregate matching functions. Section 6 studies data disaggregated by occupations. Eventually, in section 7 we investigate matching functions for different levels of educational attainment. Section 8 concludes.

2 Empirical Matching Functions

Most of the research dealing with matching functions to date was theoretical. This section introduces the economic rationale behind the concept of a matching function and surveys some recent contributions attempting to evaluate the empirical content of the matching function, especially on disaggregate levels.

The matching function is essentially a short-cut to introduce frictions and therefore to generate unemployment in models of the labor market. In a nutshell, because of imperfect information, trading frictions etc., matches m between workers looking for a job and firms looking for somebody to fill their vacancies do not arise instantaneously, but involve time consuming searching and finding of appropriate matches on both sides.[1] The larger the pool of people actively searching for employment, U, and the more posted available job vacancies, V, firms try to fill, the more matches are generated.[2] Essentially the matching function acts like a production function for new hires:

$$m = m(U, V) \tag{1}$$

with $\partial m/\partial U > 0$, $\partial m/\partial V > 0$, and $m(0, V) = m(U, 0) = 0$. The precise formal representation is flexible and depends on the problem one wants to tackle, but the vast majority of theoretical contributions involving a matching function assumes decreasing marginal returns and constant returns to scale.[3] The latter assumption is needed as to obtain a stationary unemployment rate (cf. Pissarides, 2000). This assumption is a hotly debated issue in the empirical literature. For example, Blanchard and Diamond (1989) find constant or slightly increasing returns to scale using aggregate US data. Broersma and Van Ours (1999) argue that the results for the returns to

[1] See Petrongolo and Pissarides (2001) and the references therein for microfoundations of the matching function.

[2] There are also alternative specifications of matching functions other than the dependence of matchings on stocks of job-seekers and vacancies. Coles and Smith (1998) suggest a stock-flow approach according to which the stock on one side of the market is only matched to new inflows on the other side.

[3] Exceptions exist, see e.g. Armengol-Calvo and Zenou (2001) who provide a microfoundation for an aggregate matching function which does not exhibit constant returns to scale. See also Storer (1994), Warren (1996) and Yashiv (2000) for empirical studies allowing for flexible functional forms of the matching function.

scale depend heavily on the data for active job seekers and posted vacancies used and emphasize the importance of looking at comparable measures for flows and explanatory stocks. For example, they show theoretically as well as empirically that the results for returns to scale are upward biased if only flows from unemployment to employment rather than all flows to employment are considered while using the same explanatory variables. See Petrongolo and Pissarides (2001) for an extensive overview of the empirical results on the question of constant returns.

In what follows, we retain the Cobb-Douglas representation of the matching function that is predominant in the empirical literature. Denote m_t as the total new hires at period t, or more precisely between $t-1$ and t, then

$$m_t = m(U_t, V_t) = AU_t^{\alpha} V_t^{\beta} \tag{2}$$

with α and β being the elasticities of matchings with respect to job seekers and posted vacancies, respectively, and A being a scale parameter capturing the overall efficiency of the matching process. Constant returns to scale implies $\alpha + \beta = 1$.

Empirically, the matching function can be estimated exploiting cross sectional variation across i entities of interest (like regions, industries or occupations) and time variation t. It is usually specified as a linear model containing a constant $CONS$, a time trend T and possibly other controls Z:

$$\ln m_{it} = CONS + \alpha \ln U_{it} + \beta \ln V_{it} + \gamma \ln Z_{it} + \zeta T + \epsilon_{it} \tag{3}$$

Due to the Cobb-Douglas formulation, the time trend and other controls like occupation, age and education group dummies enter the matching process in the form of augmenting 'total matching productivity'. Due to the log-linear form of the estimation equation, positive coefficients can therefore be interpreted as an additional increase in the efficiency of the labor market with respect to forming new matches stemming from the respective variable. The opposite is true for negative coefficients.

The relative sizes of the elasticities of the matching function with respect to the stock of unemployed and vacancies indicate the relative importance of labor supply and labor demand in the matching process.[4] For example, in a labor market characterized by a small β but a large α, an additional vacancy creates nearly no new hirings, while an additional job seeker leads to a new match with a high probability. In other words, there prevails a relative supply shortage on this labor market. The interpretation of a relative demand shortage is analogous.

[4]Formally, they are equivalent to the matching shares of the respective inputs in the matching process, similar to the income shares in an ordinary Cobb-Douglas production function.

In general, the empirical literature on aggregate matching functions, beginning with Pissarides (1979, 1986) and Blanchard and Diamond (1989), and recently Yashiv (2000) (see Petrongolo and Pissarides (2001) for an extensive survey), points at the empirical relevance of the concept and the strong implications of the labor market structure on the dynamic behavior of the economy.

There have been several attempts to empirically analyze matching functions on disaggregate levels. Blanchard and Diamond (1990) use U.S. data to estimate an aggregate matching function and a matching function for the manufacturing sector, and they find differences in the parameters estimated, shorter vacancy duration than for the economy as a whole and higher returns to scale. Van Ours and Ridder (1995) test for job competition between different skill groups and conclude that there seems to be competition for new hires among high skilled while there is no evidence for job competition at lower levels of education. Recent contributions like Coles and Smith (1996), Anderson and Burgess (2000), and Burgess and Profit (2001) have found similar results by using cross sectional data on regional levels compared to those stemming from aggregate time series. However, the results vary for different concepts of pools of searchers, and exhibit some regional spillovers. The analysis of Broersma and Van Ours (1999) who use an industry panel on the 1-digit level for the Netherlands, suggests some variation in the structure of labor markets of different industries.

For German data, aggregate matching functions have been estimated frequently, examples are Buttler and Cramer (1991), Entorf (1998) and references therein. However, with the exception of Entorf, disaggregate evidence, in particular, on the occupational level is scarce. Among other issues to be discussed next, one main contribution of this paper is to fill this gap in the literature.

3 Empirical Strategy and Econometric Issues

This section presents the empirical strategy followed in the remainder of the paper. Several conceptual and empirical problems deserve special attention and are therefore discussed briefly. A particularly important issue is that of spurious regression results if the data used are non-stationary. However, testing short panels for stationarity is rarely done in the respective literature. Therefore, the test strategy pursued in the empirical part is explained in some detail below.

3.1 Empirical Strategy

The primary aim of this paper is to deliver some comparable results on labor market structure across occupations and educational cohorts, that is, on an even lower level of aggregation than previously investigated in the literature, and along different dimensions. The lack of disaggregate evidence lamented by Hall (1989) has to do with the difficulties of obtaining appropriate data on disaggregated levels. Transferring the concept of a matching function to lower levels of aggregation aggravates problems with respect to the availability of appropriate data at appropriate frequencies which are already present at the aggregate levels. These problems aggravate once one starts looking at industry, regional, occupational or educational levels. Our main interest lies in the time invariant fundamental modes of functioning of the labor markets and their structural differences. Therefore, losing quite a lot of the dynamics as the frequency of the data used is annual has to be taken into account. However, the data used are longitudinal, so that intertemporal and cross sectional variation can be exploited. The investigations undertaken below should therefore be seen in a long term context. The data used are exceptionally rich and do not suffer from many data problems usually encountered in empirical studies on the matching function. In our view, this compensates weaknesses in the dynamics of the data, and justifies their use.

In particular, the data allow for disaggregation along occupations and levels of educational attainment, which are the relevant dimensions for analyzing different labor markets, rather than disaggregating along industries. Separating labor markets by occupation or education allows looking at the relevant comparable measures for flows and stocks, comparing "apples with apples". Separating by industries or regions blurs the frictional differences of labor markets arising from differences in qualificatory demands or in search intensity, screening problems etc. underlying the idea of matching functions, because industries or regions typically employ all sorts of occupations (albeit with potentially varying weights). Defining labor markets by occupations takes the structure of demand and supply and the differences in e.g. skill requirements and matching quality for certain jobs better into account. Moreover, switching industries within the same occupation might catch other effects than those interesting from the perspective of labor market structure as faced by policy makers deciding about training measures, immigration policies and the like.

With respect to the level of disaggregation, the paper is closest to the contribution by Entorf (1998) who uses panel data of yearly frequency for 40 occupational groups for the period 1971-1992. However, Entorf uses placements by the employment agency as dependent variable and concentrates on the dynamics without analyzing differences across occupations. Also Berman (1997) is able to group his data into seven occupation-industry groups which are classified by the Employment Agency and mutually exclusive. His findings vary considerably across groups and lead him to conclude that there is strong evidence of heterogeneity across, but also within groups. Van Ours and Ridder (1995) estimate matching functions for three different educational-occupational groups and find differences in competition and crowding out for the respective groups.

In every stage of the empirical analysis, we follow an identical plan: First, we present descriptive results and interesting features of the data described in the next section. In particular, because of the richness of the data it is possible to distinguish flows to employment, that is new matches, by their sources and for different dimensions of disaggregation. Then, we estimate different specifications of the benchmark matching equation (3). We also present results for regressing various measures of matches on stocks of unemployed and vacancies for reasons to be discussed next.

3.2 Conceptual and Econometric Issues

The empirical analysis of standard matching functions on disaggregate levels implies some conceptual and econometric problems which have to be addressed briefly.

Recent contributions also addressed the problem of biases in the parameters of interest due to *misspecified empirical matching functions*. In particular, Broersma and Van Ours (1999) and Mumford and Smith (1999) emphasize the importance of estimating flows with the correct corresponding stocks, which are unobservable in the case of employed job seekers. Mumford and Smith, as well as Anderson and Burgess (2000) find evidence for significant job competition between employed and unemployed job seekers and crowding out effects.

The data used in this study allow for estimations of matching functions for different measures of m and hence to replicate previous studies. Therefore we adopt the following notation in what follows. All hirings within a given period and a given definition of a labor market are denoted by m_{all}. New matches consisting of formerly employed (without a spell of unemployment between two employment relationships) read m_E. Likewise, let new matches from unemployment be written as m_U, and matches from outside the labor force m_{OL}. Taken together, new hirings from unemployment and from outside the labor force are defined as hirings from non-employment: $m_U + m_{OL} = m_X$.[5] Finally, new matches from registered vacancies, as measured by successful placements by employment agencies, are denoted as m_R. Moreover, the data allow disaggregation of different measures of active job seekers. A novelty of the paper is that matching functions of the form of equation (3) with different measures of flows, like m_{all}, m_X or m_U as dependent variable can be estimated using the same data set.

We are aware of problems in the interpretation of the results resulting from *job market competition* between employed and unemployed and *unobservable endogenous search behavior* on both sides as was suggested by Anderson and Burgess (2000) and discussed in a companion paper (Fahr and

[5]From this discussion it should also be clear that $m_X = m_{all} - m_E$.

Sunde, 2001a).[6] In this regard, we estimate several alternative specifications that try to regress flows of new hires on the correct corresponding stocks as explanatory variables. Therefore, in addition to the standard specification with all hirings as dependent variable, we present results estimations for hirings from non-employment and from unemployment to employment using unemployed job seekers as an explanatory variable, which can all be identified thanks to the richness of the data.

Disaggregation aggravates the problem of *time aggregation* because, as already mentioned before, the use of disaggregate data usually implies that the data exhibit lower frequencies than would be desirable.[7] The problem with low frequencies as the yearly data used in this study is that relevant changes in flows and stocks are misrecorded or simply not recorded: E.g. multiple matches of identical individuals leading to separations after a short employment spell show up in the flows but not in the stock of unemployed (see Petrongolo and Pissarides, 2001). In order to deal with this problem, stocks have been instrumented (e.g. Blanchard and Diamond, 1990, Berman, 1997), or the conventional matching concept has been extended along the lines of the stock-flow appoach in order to keep track of flows affecting the stocks within the observation period (e.g. Gregg and Petrongolo, 1997). As most of the disaggregated literature, we largely ignore these problems. However, in the appendix we describe an attempt to construct stock data that also take into account information on inflows to unemployment and vacancies within a year. Estimation results using these data serve as a robustness check.

A related problem is the *simultaneity bias* resulting from the use of stocks as explanatory variables which are themselves depleted by matches during the period of observation, the dependent variable. A straightforward solution to this problem is to use stocks that are measured at (or before) the beginning of the observation period of the flows. This is done in the empirical analysis.

From the preceding discussion it becomes clear that the main conceptual difficulty of empirical matching functions is to use the correct measures of matches as dependent variable and the correct corresponding stocks as explanatory variables. But besides the issues of time aggregation and simultaneity, one could also argue that the real challenge for empirical research is to only use those job seekers and vacancies that are actually *at risk* for being matched.[8] For example, not all registered unemployed (or vacancies) actually face a positive probability of being matched for various reasons, like lack of interest in finding employment, stigmatization effects etc. Therefore,

[6] See Pissarides (1994, 2000) for a theoretical model of on-the-job-search and endogenously determined search intensity.

[7] Already Blanchard and Diamond (1989) argue that the concept of a matching function would ideally require high frequency data as it is specified in continuous time.

[8] We are grateful to Melvyn Coles for bringing this point to our attention.

the usual measures of the relevant stocks imply a *problem of measurement error* leading to downward bias in the coefficient estimates: these stocks are merely a proxy of the correct relevant stocks of interest (those at risk). There is some evidence for considerable (unobserved) heterogeneity within the pools of job seekers and vacancies that affects the individual chances of being matched successfully (see Berman, 1997, Coles and Smith, 1998, and the debate on negative duration dependence). This has been interpreted as support for the validity of the stock-flow approach, because new inflows to these pools exhibit a significantly higher probability of being matched within a certain period. Thus, new inflows might be a better proxy for market participants at risk to find a match. Unfortunately, our data do not allow to identify new inflows into the pools on the levels of disaggregation necessary.

3.3 Testing for Spurious Matching Functions

The use of disaggregated data provides richer information due to additional variation across cross sectional units and allows to apply panel estimation techniques. However, there is still the danger of spurious regression results if the data are non-stationary. If both dependent variable and explanatory variables are non-stationary, estimation results might suggest a causal link which is just an artifact of the data. In this case, conventional critical values for t-statistics lead to misinterpretation since appropriate critical values would have to exceed the conventional ones.

Standard unit root tests to detect non-stationarities have low power in short time series. Testing for stationarity and unit roots in panel data is currently a field of intensive research (see Banerjee, 1999, Maddala and Wu, 1999 and Baltagi, 2000, for extensive surveys on the existing literature). In the empirical literature on matching functions, the problem of potentially spurious results on disaggregate levels has been largely ignored. A notable exception is Entorf (1998) who uses similar data like the data used in this paper.[9] He attaches considerable importance to the question whether estimates of aggregate matching functions could possibly be spurious, and indeed finds some evidence supporting this view. Citing evidence that builds on panel unit root tests he concludes that unit roots in his panel of occupational groups are unlikely.

In order to test for the possibility of spurious results, what is of interest is whether the data are stationary or not. The empirical analysis below tackles two new aspects in this respect. First, instead of testing a null of a unit root

[9] Entorf uses data from the same sources and for an almost identical period (1971-1992), but less detailed than those used in this study. For instance, he distinguishes only 40 occupational groups, while we have data on 83 occupations from 1980 to 1995. Moreover, while he uses data on placements as measure for hires, we use the arguably more appropriate social security notifications to identify matches.

in the panel data, we employ a panel test with the null of stationarity against the alternative of a unit root, proposed by Hadri (2000). The null implies that all tested series in the panel are stationary and is rejected if any one of the series has a unit root. A rejection therefore does not necessarily mean that the entire panel has a unit root. Testing for stationarity in the present context is therefore more informative than testing the panel for unit roots.[10]

Second, since the panels we use have a short time series dimension (16 observations), we cannot necessarily hope that asymptotic properties hold. Therefore, we use bootstrapping methods along the lines of Li and Maddala (1996) in order to obtain the empirical distributions of the test statistics and make inference.[11] The precise procedure is described in appendix B. In this context, testing for stationarity has the additional advantage over testing for unit roots that one has to make assumptions in order to construct bootstrap samples that are a lot less restrictive, since under a null of a unit root one would have to specify the precise time structure that is valid under the null.

4 Data

The data used for the analysis below are yearly data on unemployment, vacancies, employment levels and flows from registered vacancies to employment for Western Germany. The data are from official labor statistics and disaggregated at the occupational level. Occupations are defined by notifications by the current employer about the current activity or job and are part of an individual's social security record.[12]

Moreover, contrary to virtually all data sets used in the literature (e.g. Anderson and Burgess, 2000) also data on the pools of job seekers and vacancies are disaggregated.[13] In contrast to the help-wanted index frequently used in U.S. studies, the vacancy measure provides detailed information about all vacancies registered at local employment offices. The data where originally disaggregated by 83 occupational groups. For all estimations, the occupational group collecting all unspecified occupations, trainees and ap-

[10]In this case one would have to consider a null of unit roots in all series against stationarity in any one of the series. This might be a more conservative test in the present context, but a rejection would not indicate whether the entire panel is stationary and therefore be of little informational content.

[11]See also Maddala and Wu (1999) and Chang (2000) for contributions applying the boostrap to panel tests for unit roots. Yin and Wu (2000) propose a test for stationarity for panel data based on bootstrapped data which pools tests of the individual time series in the panel. In contrast, the methodology presented below obtains test statistics for the entire panel.

[12]For unemployed individuals or individuals out of the labor force, occupation is defined by the notification of their last employer about their last job.

[13]This is true for the main focus of disaggregation in this paper, the disaggregation by occupation

prentices (group 98) is dropped, which leaves us with 82 occupational groups throughout the empirical analysis.[14] In the definition of relevant aggregation levels of labor markets, there is a trade-off between disaggregating as much as possible on the one hand and representing labor markets as homogenous entities on the other hand.[15] Therefore, observations were clustered in nine broad occupational groups.[16] The groups were purely constructed on economic grounds. In particular, the criterion for the grouping was the proximity and relatedness of occupations with respect to skill requirements and the activities usually performed. The hirings are measured on the individual level and stem from an anonymized representative 1% sample of German social security records. The data are available for the years 1975-1995. The database is supplemented by data on unemployment benefits recipients and by establishment information (see Bender *et al.* (2000) for details.) Because there is some measurement error in the data before 1980 and to keep the data comparable to the aggregate data from labor statistics, the observations from 1980 to 1995 are retained for the empirical analysis.

The individual data include a firm identifier and information on the employment status. All in all, the data allow to identify hirings from one year to another for each occupation by source of hiring. Specifically, hirings from out of the labor force, from unemployment, and from employment can be distinguished.[17]

There is, however, a shortcoming in the data regarding the distinction of hirings from unemployment and hirings from out of the labor force. Due to measurement error in the data, our measure of hirings from out of the labor force in fact measures hirings from out of the labor force as well as hirings from unemployment (see data appendix for details). On the other hand, our measure of hirings from unemployment underestimates the true number of hirings from unemployment. Nevertheless we could clearly measure total hirings, flows from employment to employment and hirings from non-employment. Despite the measurement problems with our measure of hirings from unemployment, we present results for m_U along results for flows from non-employment m_X, because our measure of m_U serves as a better proxy for the true number of hirings from unemployment than comparable data used in the literature. The precise identification from the sources of flows to employment is important because, as already found by Broersma

[14]On average over the entire observation period, the members of group 98 make up for about 0.17 percent of total employment. The effect of dropping this group is therefore negligible.

[15]In particular, it is desirable to design occupational groups so as to represent labor markets in order to capture miscodings and to treat changes between germane occupations as within-group events.

[16]The data appendix provides a detailed description of the data used in the analysis including a list of occupational groups.

[17]However, due to use of social security data, holders of jobs which are exempt from social security payments are not recorded as employed.

and van Ours (1999) and Mumford and Smith (1999), estimates of matching functions vary considerably with the respective flows considered. However, the data are even richer than theirs in that they allow disaggregation to lower levels and along several dimensions. The central focus of disaggregation is on the occupational level. The data also allow to disaggregate flows to employment in total and by sources along other relevant dimensions like age and education. The richness of the data thereby allow to investigate all sources of flows to employment by three age groups and three educational levels, which is another innovation in the context of empirical matching.[18]

One problem of the data is that unemployment and vacancy rates are reported for certain reference dates only while the flows to employment are calculated for a given period, in the data set one year. In order to capture the relevant stocks better in this respect, some attempts have been made to create data that come closer to the theoretically desirable measures.[19]

5 Aggregate Matching

This section summarizes the results from estimations of different specifications of the matching function as specified in equation (3) using data aggregated up from 82 occupational groups for the period 1980 to 1995. These results are of particular interest since the literature on empirical matching functions is based almost exclusively on aggregate time series data although at much higher frequencies. The section discusses the benchmark results and different time structures and presents estimates of matching function for different measures of flows.

5.1 General Results

Descriptive statistics of the data are listed in Table 5-1. It is striking that the vacancy rate is so much smaller than the unemployment rate. This has to do with the fact that both measures are collected for reference dates, but that the average duration of a vacancy is a lot shorter than that of an unemployment spell. It is interesting to note that more than two thirds of all new hirings affect individuals who are not employed.[20] The numbers

[18] Information on the education attainment is reported by the employer and might therefore contain some measurement error. This should be taken into account when interpreting the results.

[19] Details on how these adjusted stocks are constructed are described in the Data Appendix.

[20] The numbers are roughly in line with those mentioned in the literature: Blanchard and Diamond (1989) conclude that in their U.S. data, m_E are about 15 percent of total matches while m_U and m_{OL} make up for 45 and 40 percent, respectively. According to Burda and Wyplosz (1994), German data reveal 16, 42 and 42 percent for m_E, m_U and m_{OL}.

obtained for hirings from unemployment and out of the labor force by disaggregating further reveal surprisingly high flows from out of the labor force and surprisingly low flows from unemployment. However this has to do with the coding of the data mentioned before, so these numbers have to be taken with a big grain of salt. Flows from non-employment seem to be a lot more reliable and the analysis below will thus concentrate on those. Flows from registered vacancies make almost 60 percent of total flows and are measured as successful placements by the employment agencies. The main problem with these flows is that according to official statistics up to 20 percent of successful placements result in employment relations that end within less than 8 days (see ANBA, 1997). Therefore, also results obtained using this flow have to be interpreted with care.

[INSERT TABLE 5-1 ABOUT HERE.]

Table 5-2 presents results from estimating several specifications of a standard Cobb-Douglas matching function of the form of equation (3).[21] The baseline specification (1) contains a linear time trend. Specification (2) additionally includes dummies for broad occupational groups. In both cases, the hypothesis of constant returns to scale can be rejected at the 1%-level. The stock of unemployed job seekers has a slightly higher weight in the creation of total new matches than the stock of registered vacancies. The linear time trend has a significant negative coefficient indicating an increase in the frictions on the labor market over time. This pattern is virtually unchanged qualitatively as well as quantitatively once time dummies are introduced instead of a linear time trend (specification 3), and when dummies for 9 broad occupational groups are added (specification 4). However, while the coefficients for all time dummies are consistently negative, they become highly significant only after 1985, and their absolute value increases over time. This indicates a deterioration in the efficiency of the functioning of the labor market over time, equivalent to an outward shift of the Beveridge Curve. The finding of negative time effects is consistent with what is found in the literature, typically starting with the early 1960s (see Pissarides and Petrongolo, 2001, for a survey.) Controlling for GDP (model 5), GDP growth (model 6), and both (model 7) leaves the results virtually unchanged

[21] The results are obtained by pooling the data from 82 occupations, resulting in $16 \times 82 = 1312$ observations. In contrast, Entorf (1998) aggregates the data before estimating which leaves him with only 22 observations.

while the coefficients for the controls are not significant.[22,23]

As was mentioned previously, in order to rule out the possibility of spurious results, we test the aggregate data for stationarity using the panel test suggested by Hadri (2000). For inference, we apply the bootstrap method described in appendix B. It turns out that we can neiter reject the null of level stationarity for all matches, nor the null of stationarity when a trend is included.[24]

[Insert Table 5-2 about here.]

The results obtained using panel data fit in the broad picture obtained by numerous studies of empirical matching functions obtained with aggregate time series data (see Broersma and Van Ours (1999) for an overview.) One striking difference to most previous results is that the matching elasticity with respect to unemployment is consistently higher than that for vacancies. One has to bear in mind, however, that the matching function estimated might be misspecified in the sense that the stocks used as explanatory variables might not be the relevant measures since for instance the pool of employed job seekers is neglected completely. We take a closer look at this problem below and in Table 5-3. These results are indeed particularly interesting in the light of the result of Broersma and Van Ours (1999). They argue that ignoring relevant parts of the explanatory stock of job seekers leads to an underestimation of the true value of the matching elasticity with respect to the number of job seekers. This point will be discussed in more detail below in section 5.2.

As was mentioned in section 3.2, the estimation results obtained so far could potentially be biased due to problems of time aggregation and simultaneity, because flows could affect contemporaneous stocks used as explanatory variables. This is unlikely as we use stock data collected before the beginning of the period for which matches are recorded.[25]

Another potential problem has been already mentioned before in the data section. Explanatory variables are reported only for reference dates

[22]Note that including macro variables in a regression like time trends, GDP or GDP growth leads to downwardly biased standard errors, see Moulton (1990). A correction of the standard errors would therefore render the coefficient estimates for time trend, GDP and GDP growth less significant. We refrain from a correction of the standard errors, since the results for time trends have very low standard errors and are virtually identical to those for time dummies, and because the coefficients for GDP and GDP growth are insignificant anyway.

[23]In specifications without time trend, coefficients for GDP and GDP growth, respectively, are negative and highly significant.

[24]The respective empirical $p-$values are 0.345 and 0.164, respectively. Also, the null of stationarity of the series of the stocks, unemployment and vacancies, could not be rejected. Details are available from the authors upon request.

[25]In a companion paper (Fahr and Sunde, 2001b) we check robustness for using lagged explanatory variables. The results remain qualitatively as well as quantitatively unchanged.

and therefore do not represent the correct measures for stocks of unemployed and vacancies relevant for the estimation of matching functions. In order to overcome this problem, the reference data have been adjusted using aggregate data on total yearly inflows into unemployment and registered vacancies. Estimations of the same specifications of the matching functions as in Table 5-2 using these adjusted stocks as explanatory variables deliver almost identical results.[26] We take this as an indication that the results obtained with the non-adjusted data are reasonably robust and accurate.

Due to the similarity of the results obtained by including time trends or time dummies, in what follows, matching functions will be estimated including a time trend.

The results presented in Table 5-2 indicate that the sum of unemployment and vacancy elasticities of new matches sum up to less than one which implies decreasing instead of constant returns to scale. The results of F-tests for the null that the sum of the coefficients equals one, that is constant returns, clearly reject the constant returns hypothesis at any reasonable significance level. This finding is confirmed by the results of using more flexible functional forms of the matching function (see Fahr and Sunde, 2001b).[27] Moreover, results for specifications with different measures of matches, in particular all hirings, hirings from non-employment and from unemployment, consistently lead to similarly strong rejections of the hypothesis of constant returns to scale in favor of decreasing returns.

5.2 Matching Functions by Sources of Flows

The data set allows to disentangle new matches by the source of either the successful job seekers or the type of successful vacancies. That is, successful matches of formerly unemployed can be distinguished from successful matches of individuals switching from another job to their new match, etc. Table 5-3 presents the results for matching functions defined by the sources of flows. The models presented replicate all flow specifications used in the literature as presented by Broersma and Van Ours (1999) using one single data set and therefore allow for direct comparisons of the different results.

[Insert Table 5-3 about here.]

The first model (model (8)) repeats the results for taking total hirings per occupation and period as was already displayed in Table 5-2. Across all specifications, the time effect is negative and, with exception of model 13, significant, which is a common result in the literature (see above).

[26]The estimation results for adjusted explanatory variables can be found in Fahr and Sunde (2001b).

[27]See also Gross (1997) for similar findings for German aggregate data.

Looking at hires from non-employment, which is the dependent variable in model (9) we find very similar coefficients for the matching elasticities.[28] Since the literature comes to mixed conclusions about the relative size of the matching elasticities we take our outcomes as a consequence of using an identical data set also for the benchmark model. This might also be an indication for the robustness of our results. Due to the fact that flows from non-employment are very reliable, the results indicate that the biases due to incompatibility of flows and stocks used (Broersma and Van Ours, 1999) or endogenous job competition (Mumford and Smith, 1999, Anderson and Burgess, 2000) are rather moderate.

The main point of Broersma and Van Ours (1999) is to stress the importance of specifying the matching function such that the measure of flows (new matches) corresponds to the measures of stocks of job seekers and vacant positions. From a conceptual point of view, the models presented in this section are problematic in this respect since, strictly speaking, one would have to estimate e.g. model (10) with employed job seekers instead of the stock of unemployed. However, the stock of employed actively searching for a new job is not observable. Therefore, the stock of unemployed might serve as an instrument for the stock of employed job seekers. But in this case, problems regarding the composition of the pool of job seekers and crowding out effects, as investigated by Anderson and Burgess (2000) are neglected. In turn, the estimated parameters might be biased due to the misspecification. Moreover, as was already mentioned in section 3.2, what one would actually want to look at is the stocks of seekers *at risk* of being matched, so one faces an additional measurement problem leading to biased coefficient estimates.

To sidestep these problems, and in order to be able to interpret the results for disaggregate matching functions in the remainder of the paper in the context of the existing literature, in what follows we present results for models of type (8), (9) and (11). We pay particular attention to models of type (9). In this specification, the outflows from non-employment into work are estimated using the stock of unemployed job seekers and the stock of registered vacancies. This implicitly assumes that the stock of unemployed represents the relevant stock of non-employed job seekers sufficiently well and that non-employed, that is unemployed or individuals (re-)entering the labor market, predominantly apply for registered instead of non-registered vacancies. Arguably, this specification minimizes problems related to misspecification and mis-measurement.

An alternative would be to use the same explanatory variables for regressing flows from registered vacancies (model 13), assuming that these

[28]Bootstrapped panel stationarity tests reveal that the hypothesis that m_X is level stationary cannot be rejected, with the p-value of a type-I-error being 0.353. Likewise, we cannot reject stationarity when a trend is included, with a p-value of 0.313.

are (almost) exclusively filled with unemployed. However, the results in Table 5-3 reveal that this specification exhibits the strongest discrepancy from the main pattern of results from all other specifications, in particular with respect to the estimated matching elasticities with respect to the pool of unemployed. The coefficient for unemployment is higher, the coefficient for vacancies lower than in all other cases. The significantly higher coefficient in the model for flows from registered vacancies could indeed be an indication for endogenous search behavior of employed individuals which is contained in this specification but not in the other one for outflows from non-employment: An additional job seeker has a higher impact in this model because implicitly also some employed job seekers have the chance to lead to a successful match. Therefore, the emphasis on models of type (9) in the analysis of disaggregated matching functions is warranted.

The dependent variable in model (10) is hires from employment. Although conceptually the model is misspecified at least with respect to the pool of job seekers, which is the stock of unemployed, the results are very similar to the ones obtained for the benchmark model. In contrast to the results obtained by others, the coefficient for vacancies is lower than the coefficient for the stock of unemployed.

Model (11) regresses matches from unemployment on the stock of unemployed and the stock of vacancies. The elasticity of matches with respect to the stock of job seekers is higher, the elasticity with respect to vacancies lower than in the benchmark model.[29] The intuition behind this result is that total flows into a job capture also jobs filled with employed job seekers. Therefore, an identical increase in the stock of unemployed job seekers leads to a smaller increment in total matches than if only unemployment outflows into a new job are looked at, simply because there is crowding out by employed job seekers. A similar argument can be given to explain the smaller elasticity of vacancies for outflows from unemployment: Since the pool of potential applicants actually is higher for total flows to jobs than for unemployment outflows, posting one more vacancy has a higher effect on total hirings than on unemployment outflows. Qualitatively the same results - a higher matching elasticity with respect to the stock of unemployed than with respect to vacancies - is consistently found in the literature, albeit with somewhat higher coefficients for unemployment than obtained in the present study. However, as mentioned before, the results for flows from unemployment as well as for flows from out of the labor force to be inspected below have to be interpreted with care as the data contain a potentially large number of miscodings. Interestingly, the observation by Broersma and Van Ours (1999) that estimates for returns to scale are upward biased when

[29] Bootstrap panel stationarity tests of the null of stationarity of m_U deliver probabilities for committing a type-I-error of 0.697 and 0.622 when a trend is excluded or included, respectively. Hence we cannot reject stationarity.

regressing m_U instead of m_{all} is not supported in our data.

If new matches of individuals from outside the labor force are taken as the object of study (model 12), again the estimated elasticities are in line with the previous findings. However, the coefficient for vacancies is slightly higher than in the other models and also higher than the coefficient for the stock of unemployed, indicating that the availability of job opportunities might affect the participation decision positively, as one would expect.[30]

For completeness, all models contained in Table 5-3 are estimated in a two-way fixed effects panel specification. The results are contained in Table A4-3-4 in the appendix. The results change somewhat. In particular, the differences between the alternative models become more pronounced. The most striking difference is in the model estimated with flows from employment as dependent variable (model 10'), where the elasticity with respect to the stock of unemployed goes up from 46 percent to 57.5 percent in the panel specification, while the vacancy elasticity goes down from 41.7 percent to 27.1 percent with comparable explanatory powers of the models. Constant returns to scale are unanimously rejected also in all panel models.

6 Matching by Occupational Groups

While the bulk of the existing literature on empirical matching functions focuses on aggregate data and concentrates on the dynamic aspects of job creation (and destruction), little is known about cross-sectional patterns of matching. Recently, a few studies considered the regional dimension of labor markets and analyzed regional differences in matching.[31]

While all studies find effects of spatial and/or industrial heterogeneity on the matching outcomes, no study has used data disaggregated by occupations.[32] However, occupation might well be considered to define the relevant labor markets as faced by participants, that is job seekers and vacancy posting firms. Occupation captures skill requirements and characteristics, similarities of tasks etc. and therefore defines the potential labor market on which the participants search for a new match better than for example industry.[33] Moreover, occupations are explicitly defined in public data. This

[30] In order to check the robustness of the results, we estimated models (8) to (13) also using lagged stocks as well as adjusted stocks to alleviate problems of simultaneity and time aggregation, see Fahr and Sunde (2001b). Only the results for adjusted stocks only differ somewhat, confirming our confidence that the estimates obtained with unadjusted explanatory variables are robust and reasonably accurate.

[31] See Petrongolo and Pissarides (2001) for a survey of the - small - literature.

[32] Entorf (1998) uses information on occupations, but does not analyze the effects of occupations. He rather concentrates on the dynamic variation of matching functions and Beveridge Curves. Van Ours and Ridder (1995) and Berman (1997) distinguish between occupation-industry categories.

[33] Of course there might be correlations between certain industries and certain occupations.

is even more true for the German education system in which there are education and training certificates for nearly every occupation. Occupations are defined by the notification of the employer (or the last employer if the individual is currently not employed) in the individual's social security records. These notifications are not only reasonably precise but also keep track of the working history of an individual. In contrast to that, the definition of regional labor markets and travel-to-work areas arguably contains a bit more arbitrariness. Therefore, exploring labor market structure and matching functions for different occupations is an important issue which has been neglected so far. As described in the data section, we group occupations into nine broad occupational groups.[34] In order for matching functions to be estimated in a meaningful way, stocks and flows should correspond to the same labor market. For this to be true, the flows between occupations should not be too high in order for the distinction of labor markets by occupations to be relevant. In particular, large flows between groups would aggravate the measurement error problem mentioned above in section 3.2 even more. However, in the data at hand, on average less than two percent of the labor force change broad occupations in a given year. On average, less than a third of all new hirings involves a change of broad occupation. The bulk of these flows between occupations is concentrated among a few groups and the overall picture is that even fewer matches involve between group changes.[35] Therefore we emphasize occupations as defining labor markets as a feature of relevance which has been almost completely neglected so far. Figure A1 in the appendix illustrates the evolution of the sizes of the broad occupational groups over the entire observation period 1980 - 1995.

6.1 General Results Controlling for Occupation

Table 6-1 shows the results for estimations of the standard matching function of specification (3) when occupation dummies for nine broad occupational groups are added. As already discussed before, the dummy for a given occupational group acts as augmenting the total factor productivity of the matching function. Positive coefficients indicate an increase in the overall efficiency for the respective group, and negative coefficients a decrease. In order to obtain results that do not rely on which reference group is chosen, we transform the dummy coefficients following the approach suggested by Haisken-DeNew and Schmidt (1997). After regressing logged matches on a set of explanatory variables and dummies (excluding a reference group), the coefficients obtained for the dummies are renormalized as weighted devia-

[34] Berman (1997) constructs seven industry-occupation groups. We focus only on occupations and therefore construct groups which do not comprise any relation to industries classifications.

[35] For example, about 20 percent of the flows between broad occupational groups are flows between manufacturing occupations and crafts (groups 2 and 3).

tions from the weighted mean of all groups.[36] As a consequence, also the variance-covariance matrix and the standard errors of the coefficients are adjusted.

[INSERT TABLE 6-1 ABOUT HERE.]

In model (14), the dependent variable is all hirings, in model (15) only matches from non-employment are considered as dependent variable. The coefficients for the stocks of job seekers and vacancies for all hirings and hirings from non-employment are virtually identical to those already obtained before in Tables 4-2 and 4-3. The null-hypothesis of constant returns to scale can be rejected in all cases at any level of significance. The results for the dummies are striking. In models (14) and (15), except for group 9 (low skilled occupations), the dummy coefficients are qualitatively identical. Those which are highly significant are also quantitatively very similar for both concepts of hirings used, with the exception of the primary sector. The negative coefficient means that new matches are less likely for members of this group, than on average over all occupations. However, this disadvantage is a lot smaller for matches from non-employment than for all hirings. The groups for which the labor market seems to be particularly dynamic in the sense that compared to the average flows over all groups more matches occur *ceteris paribus*, are high skilled and health occupations (groups 6 and 8). The opposite is true for technical occupations (group 4) which create significantly fewer matches *ceteris paribus*. One reason for this might be that the market for these individuals is less dynamic, for example because the positions they usually fill have very particular skill requirements, high capital intensities or bear high responsibility. Another interpretation is that these jobs offer generous compensation packages and that thus turnover is low. Crafts (group 3) represent a relatively efficient labor market with more matches than the average. This could be explained by the fact that the German apprenticeship system provides a homogeneously high level of human capital which is highly transferable, so more employment relations can be created as desired skills are easily observed. For completeness, model (16) is added where the dependent variable is matches from unemployment. Constant returns can be rejected as in the two specifications described before. The other results are qualitatively the same with the exception of the occupation dummies for manufacturing (group 2) and health (group 8) occupations. The former exhibits a significant positive, the latter a significant negative effect, exactly opposite to what is found when regressing all hirings or hirings from non-employment.

Instead of pooling the data, we also estimated random effects and two-way fixed effects models. Hausman tests reveal that if broad occupations are

[36] The group sizes which are necessary as weights for the deviations as well as for the calculation of the mean are given by the number of occupations contained in the respective broad occupational group, because occupations are the unit of observation.

the relevant group variable, both approaches are absolutely identical with respect to their coefficient estimates for all models shown in Table 6-2. For all hirings as the dependent variable we get identical results as in the pooling case (with estimates of 0.459 for α and 0.369 for β regardless of the panel methodology used). For hirings from non-employment we get coefficients (with standard errors in parentheses) for the supply elasticity α of 0.408 (0.022) and for the demand elasticity β of 0.424 (0.018) which differ slightly from the pooling case.[37]

6.2 Results for Broad Occupational Groups

As seen before, if the matching function is imposed to have identical elasticities for stocks of job seekers and vacancies, as was done in the estimations with occupation dummies included, constant returns to scale are rejected in favor of decreasing returns to scale. In this subsection, matching functions of the same form (eq. 3) are estimated separately for each broad occupation group. The results are presented in Tables 6-2 and 6-3.

Model (17) takes all hires for the respective occupation group as dependent variable, while the dependent variable in model (18) is hirings from non-employment for the respective occupation group.[38] Bootstrapped panel tests for stationarity carried out for each broad occupational group and for each definition of flows (m_all, m_X, m_U) separately unambigously show that the null of stationarity cannot be rejected.[39] A first inspection shows that the results are very similar for both concepts of flows used. When total matches are looked at, the hypothesis of constant returns to scale can be rejected at any reasonable significance level.

[Insert Tables 6-2 and 6-3 about here.]

As before, one model takes all hirings as dependent variable (model 17), while another looks at hirings from non-employment only. With respect to the question of returns to scale, the results are almost identical. Constant returns to scale cannot be rejected for occupations in the primary sector (group 1), for service occupations (group 5) and for health occupations (group 8), and this is consistent regardless of which measure for matches is used as dependent variable. The only exception is low skilled (group 9), for which constant returns can be accepted for total hirings, but has to be rejected once hirings from non-employment are regressed. Interestingly, crafts and

[37]For hirings from unemployment, the respective estimates for α and β are 0.407 (0.031) and 0.449 (0.026).

[38]For matters of completeness and comparison, Table 6-3 contains estimates for all hirings (model 17) and hirings from unemployment (model 19).

[39]The null is that all time series in the respective panels are stationary processes. All tests control for serial dependence in the errors. Detailed results are available from the authors upon request.

technical occupations (groups 3 and 4) exhibit increasing returns to scale. The overall picture is therefore very mixed. The degree of homogeneity of the matching function very much depends on the occupation under consideration. Moreover, the argument of increasing returns on the micro level and constant returns on the macro level as suggested in the literature (see Diamond (1982) and Coles and Smith (1998) for a microfoundation based on overlap of adjacent labor markets) seems to be supported only to a limited extent by our data.

Furthermore, also with respect to the sizes of the coefficients for stocks of unemployed and vacancies, the results reveal a pronounced heterogeneity across groups. The overall picture is the same regardless of what measure for matches is taken as the dependent variable. Roughly speaking, there are three categories of occupations. For groups 1, 7 and 9 (manufacturing, social and low skilled occupations), new matches are more elastic with respect to the stock of vacancies than unemployment. The opposite is true for all other groups, and particularly strong in technical occupations. The only exception is manufacturing (group 2) where labor supply and demand have about equal weights in creating new hires. Also the time trends exhibit interesting differences across groups.[40] The coefficient for the time trend is negative and significant at the aggregate as well as for most groups when taken separately with a value of about three to five percent. The negative value is largest for primary sector occupations (group 1) with around 12 percent. On the other hand, the coefficient for the time trend is not significantly different from zero for technical, white collar, social and low skilled occupations (groups 4, 6, 7 and 9). This indicates that, when taken separately, overall matching efficiency did not decrease in all labor markets over the time period under consideration.[41]

7 Matching by Level of Educational Attainment

The discussion above indicated that there is considerable heterogeneity in the workings of labor markets for different occupations. This section takes a closer look at the matching functions for different educational cohorts. The *modus operandi* is the following: First, we look at matching functions estimated separately for different educational groups. Then, the interactions

[40]The detailed results are not contained in the table, but are available from the authors upon request.

[41]For better comparability among groups, Fahr and Sunde (2001b) contains estimates that result from imposing constant returns to scale for each occupational group. Panel estimates for the models shown are also available from the authors upon request. While again specification find no systematic differences between results from fixed and random effects models, the results differ in detail somewhat from the results in Tables 6-2 and 6-3. However, the main conclusions, in particular that of strong heterogeneity across groups, remain unaffected.

between education and occupation are examined more closely.[42]

7.1 Matching by Educational Groups

We divide individuals into three groups according to the education level they attained: We distinguish individuals with low educational backgrounds (group 1), defined as neither having successfully completed high-school (without *Abitur*), nor having completed an apprenticeship. Group 2 consists of individuals with an intermediate level of education, that is with either high-school diploma (*Abitur*), or completed apprenticeship, or both. Finally, we assort all individuals with a university degree or a degree from an applied university or polytechnic (*Fachhochschule*) into a high education group (group 3).

Table 7-1 presents summary statistics on group sizes and the flows from the different sources as obtained for the different educational groups under consideration. More than half of all hirings are made up for by the intermediate education group, while this group constitutes more than 60 percent of the population.[43] Less than 15 percent of total matches can be traced back to highly educated individuals who make up for less than 12 percent of all individuals. That is, with respect to their group sizes, highly educated and less educated are overrepresented in total matchings, the intermediate education group is underrepresented. This changes when one considers different measures of hirings. When looking at job-to-job matches, more than 65 percent can be attributed to individuals of intermediate education, whose share of employed is about 62 percent. About a fifth of hirings from employment go to individuals of low education, who represent around 27 percent of employed, and the rest of about 14 percent of job-to-job matches involve highly educated, while only 11 percent of employed are belonging to this group. Thus, groups 2 and 3 are overrepresented, group 1 forms fewer hirings from employment as would correspond to its size. This somewhat contradicts the conventional wisdom of the highly educated making up for the bulk of job-to-job changes.

Surprisingly, when it comes to hirings from non-employment, less educated form more matches (almost 41 percent) than would be expected from their share of non-employed (less than 34 percent). On the other hand, members of the intermediate and the high group form fewer (44 %) and much fewer (14.6 %) matches than would be expected from their shares of the pool of employed (46 and almost 20 percent, respectively). This indicates that

[42]Again, strictly speaking, the flows are not regressed on the correct corresponding stocks, because neither the stock of unemployed per year and occupation is available by education groups, nor - trivially - the stock of vacancies.

[43]Group sizes are relatively stable with the share of individuals with low education decreasing somewhat over time in favor of increases of the other two groups. Fahr and Sunde (2001b) contains a Figure displaying the development of the relative sizes of the educational groups over the entire observation period 1980 - 1995.

those with low education actually form more matches from non-employment than would be expected from their relative group size. This effect is even more pronounced when matches from out of the labor force are considered. The low education group, which is about 27 percent of the population (and less than 34 percent of non-employed), contributes almost 48 percent of the flows, while intermediate education only contributes to less than 37 percent but constitutes more than 60 percent of all individuals (and 46 percent of non-employed). Highly educated create relatively more matches from out of the labor force than their relative size (15.5 percent versus 11.84 percent), but less than their proportion of non-employment (19.8 percent). This suggests that individuals with low education exhibit the lowest attachment to the labor market, and that they revise their participation decision more often than more educated people.

[Insert Table 7-1 about here.]

Moreover, within the low education group, hirings from out of the labor force represent almost two thirds of all matches of this group, with matches from employment being a mere 18 percent. This means that the probability of having to go through an unemployment spell or even some period of non-participation is very high for members of this group. This pattern changes for intermediate education. More than 40 percent of all matches of this group are job-to-job matches, more than 27 percent are matches from unemployment. Surprisingly, the high education group has a distribution of all hirings that lies in between the two other groups. Particularly noteworthy is that a good 30 percent of their new hirings are matches from employment, while almost 50 percent are hirings from out of the labor force. Intuitively this means that highly educated either switch jobs directly or stop participating for some time (possibly living of their bonuses) before entering a new employment relation, while only a fifth goes through an unemployment spell before forming a new match.

The standard matching function is regressed for flows for three different educational groups. The results for matching functions for all flows m_{all} and hires from non-employment m_X are presented in Table 7-2.[44] Interestingly, the hypothesis that the matching function exhibits constant returns to scale can only be rejected in favor of decreasing returns to scale for the matching function estimated for the intermediate group using all hires. Regardless of the measure used as dependent variable, the group of low education produces the highest coefficient for the stock of unemployed and the lowest coefficient for vacancies. This indicates that an additional job seeker in this group creates a new match with a relatively higher probability than in the other groups. The opposite is true for a person in the intermediate group. On the

[44]Panel stationarity tests reveal that the null of stationarity of all hirings and hirings from non-employment cannot be rejected for any of the three educational groups.

other hand, an additional vacancy creates relatively fewer matches in the low group than in the high, and in particular the intermediate group.

[Insert Table 7-2 about here.]

The interpretation for this is straightforward. A given vacancy targets mostly intermediately and highly educated individuals, thus produces relatively more matches for members of these groups. On the other hand, variation in the stock of unemployment explains most matches for the low educated. This can either mean that this group is particularly affected by cyclical variations or that the labor market for this group is very dynamic in the sense that many matches split up after a short period of time, therefore creating a lot of variation in both matches and the stock of unemployed. Moreover, group 1 also exhibits a negative time trend that is about three times as large as the trend for the other groups, regardless of the dependent variable used, which could be taken as an indication for decreasing labor market efficiency for these people, or even a sign for skill biased technological change.

7.2 Educational Groups and Occupational Differences

In this subsection we repeat the estimations of empirical matching functions for different educational levels with controls for broad occupational groups. Table 7-3 summarizes the findings for regressing total hirings in the standard specification including occupational dummies. The coefficient estimates for stocks of job seekers and vacancies are similar to the ones obtained with the specification without occupations. The elasticity of matches with respect to vacancies is somewhat higher in this specification, about 10 percent higher in the group with low education attainment, about 7 percent higher in the high education group, and about 3 percent in the intermediate cohort. Again, the time trend is negative for low and intermediate education, but insignificantly positive for individuals with high education. The null of constant returns to scale of the matching function can be rejected for intermediate and high education groups in favor of decreasing returns.

[Insert Table 7-3 about here.]

As in the preceeding section, we repeated the estimation of the models in Table 7-2 using the method of seemingly unrelated regression. As for age groups, the null hypothesis of identical unemployment *and* vacancy elasticities for matches of either definition across education groups, respectively, can be rejected at any reasonable level.[45] Moreover, also the hypotheses of either identical vacancy elasticities for all groups or identical unemployment elasticities for all groups can as well be rejected.

[45]See Tables A7-2-1 and A7-2-2 for detailed results obtained with SURE.

Once again, the occupation effects are interesting. As compared to the average over all occupations, manufacturing occupations (group 2) exhibit relatively more matches for individuals with low education, but fewer for individuals with intermediate or high education attainment. Similarly, low and middle education groups create more matches in crafts (group 3) than the average, high education leads to fewer matches. Exactly the opposite is true for technical occupations (group 4), service and white collar occupations (groups 5 and 6), and social and health related occupations (groups 7 and 8). In these occupations, low education exhibits lower matching rates (in particular for technical and social occupations). On the other hand, intermediate and in particular high education leads to more matches than average, where the positive effects are strongest for technical, white collar and social occupations. Primary occupations (group 1) form fewer matches in general, regardless of the occupation, while the negative effect is insignificant for high and low education. Oddly, in low skilled occupations (group 9), high and low education lead to more matches than on average, while intermediate education exhibits fewer matches. Most of these occupation effects probably have to do with the fact that the structure of skill requirements differs systematically across occupations and thus determines the number of flows in different education groups as education is systematically related with the skills required.

Table 7-4 presents the results for matchings from non-employment as dependent variable.[46] The coefficient estimates are quite similar to those obtained before, with the exception that the unemployment elasticity for the low education group is about 7 percent and for the intermediate group about 4 percent higher than when regressed for all hirings. Constant returns to scale can be rejected for all three education groups. However, while this is done in favor of decreasing returns for intermediate and high education, the matching function for low education seems to exhibit increasing returns to scale. The pattern of interactions between education and occupation, captured by the dummy coefficient estimates is virtually identical apart from minor differences concerning significance of the coefficients.

[Insert Table 7-4 about here.]

Matching functions for the three education groups have been estimated as seemingly unrelated regressions (SURE) separately for all nine occupation groups.[47] The hypothesis that unemployment elasticities are the same for all three age groups cannot be rejected at the 5 percent level for primary (1), white collar (6), social (7) occupations regardless of the measure for

[46]Results for hirings from unemployment as dependent variable differ substantially from those obtained with all hirings and hirings from non-employment and are contained in Fahr and Sunde (2001b). But as before we forbear from a detailed interpretation.

[47]Detailed Results are available from the authors upon request.

matches. The null for identical vacancy elasticity for all education groups can be rejected at the 5 percent level for all occupations. The same is true for the null of both elasticities being the same across educational groups.

8 Conclusion

Summarizing the investigations of empirical matching functions disaggregated along different dimensions, the most striking, although not too surprising, finding is the apparent heterogeneity of the matching technology for different categories of jobs and workers. While the results presented for the aggregate matching function fit nicely into the findings reported in the literature, this aggregate matching function seems to be made up of very distinct matching functions on very distinct labor markets with very distinct characteristics of the matching process. The stylized results are the following: In general with only few exceptions, constant returns to scale of the matching function is rejected on the aggregate as well as on disaggregate levels. On the aggregate level, we find evidence for decreasing returns to scale.

On the occupational level, matching technologies for some occupations exhibit increasing returns to scale while at the same time those for other occupations exhibit decreasing returns. Some occupations exhibit a relatively high elasticity of matches with respect to the stock of job seekers, indicating that matches are relatively more supply determined, while this elasticity is relatively low for others. Examples for the former group are technical occupations, for the latter low skilled occupations. The same goes for the elasticity with respect to vacancies. Also differences in the absolute values of these elasticities are substantial across occupations.

Moreover, the patterns of this heterogeneity are modified once other concepts of flows like e.g. those from non-employment are looked at instead of all hirings within a profession. This sheds some more light on issues like whether the employment status matters for finding new employment and whether there is crowding out from employed job seekers.

Also, matching technologies are quite heterogenous for members of different education groups indicating that labor markets are heterogeneous also along this dimension. We find evidence that suggests that matches are more demand-determined for lower educational levels, indicating relative demand shortages for members of this group as compared to higher educated. However, the results differ across occupations, leading to the impression of different systematic education effects on matching for different types of skills.

The main implication from these findings is a caveat on the usefulness of an aggregate matching function to explain the working of the labor market as a whole since it discards a lot of potentially important information. In particular, for the conduct of economic policy it might be indispensable to look beyond the aggregate level in order to have a clear understanding of the structure of labor markets, and the pattern of frictions at work.

References

[1] ANBA (1997), *Amtliche Nachrichten der Bundesanstalt für Arbeit*, Arbeitsstatistik 1997, Jahreszahlen, 46. Jahrgang.

[2] Anderson, Patricia M. and Simon M. Burgess (2000): "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data", *Review of Economics and Statistics* 82(1), pp. 93-102.

[3] Baltagi, Badi (ed.) (2000): *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Advances in Econometrics Vol. 15, Elsevier, Amsterdam.

[4] Banerjee, Anindya (1999): "Panel Data Unit Roots and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics* Special Issue, pp. 607-629.

[5] Bender, Stefan, Anette Haas and Christoph Klose (2000): "IAB Employment Subsample 1975 - 1995. Opportunities for Analysis Provided by the Anonymised Subsample", *IZA Discussion Paper* 117, (February).

[6] Bender, Stefan, Juergen Hilzendegen, Goetz Rohwer and Helmut Rudolph (1996): "Die IAB-Beschaeftigtenstichprobe 1975 - 1990", *Beitraege zur Arbeitsmarkt- und Berufsforschung* 197.

[7] Berman, Eli, John Bound and Zvi Griliches (1994): "Changes in the Demand for Skilled Labor Within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures", *Quarterly Journal of Economics* 109(2) (May), pp. 367-397.

[8] Blanchard, Olivier J. and Peter Diamond (1989): "The Beveridge Curve", *Brookings Papers on Economic Activity* 1:1989, pp. 1-76.

[9] Blanchard, Olivier J. and P. Diamond (1990): "The Aggregate Matching Function", in: P. Diamond (ed.): *Growth, Productivity, Unemployment*, Cambridge Massachusetts, S. 159-201.

[10] Broersma, Lourens and Jan C. Van Ours (1999): "Job Searchers, Job Matches and the Elasticity of Matching", *Labour Economics* 6, pp. 77-93.

[11] Burda, Michael and Stefan Profit (1996): "Matching Across Space: Evidence on Mobility in the Czech Republic", *CEPR Discussion Paper* 1364 (March).

[12] Burda, Michael and Charles Wyplosz (1994): "Gross Worker and Job Flows in Europe", *European Economic Review* 38, pp. 1287-1315.

[13] Burgess, Simon M. (1993): "A Model of Competition Between Unemployed and Employed Job Searchers", *Economic Journal* 103, pp. 1190-1204.

[14] Burgess, Simon and Stefan Profit (2001): "Externalities in the Matching of Workers and Firms in Britain", *mimeo*, University of Bristol.

[15] Buttler, Friedrich and Ulrich Cramer (1991): "Entwicklung und Ursachen von mis-match-Arbeitslosigkeit in Westdeutschland", *Mitteilungen aus Arbeits- und Berufsforschung*, 24(3), pp. 483-500.

[16] Chang, Yoosoon (2000): "Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency", *mimeo*, Rice University.

[17] Coles, Melvyn G. and Eric Smith (1996): "Cross Section Estimation of the Matching Function: Evidence from England and Wales", *Economica* 63(252), pp. 589-597.

[18] Coles, Melvyn G. and Eric Smith (1998): "Marketplaces and Matching", *International Economic Review* 39(1) (February), pp. 239-254.

[19] Diamond, Peter (1982): "Wage Determination and Efficiency in Search Equilibrium", *Review of Economic Studies* 49(2) (April), pp. 217-227.

[20] Entorf, Horst (1998): *Mismatch Explanations of European Unemployment, A Critical Evaluation*, Springer, Berlin.

[21] Fahr, René and Uwe Sunde (2001a): "Endogenous Hiring Behavior in Empirical Matching Functions", *IZA Discussion Paper* 320 (June).

[22] Fahr, René and Uwe Sunde (2001b): "Disaggregated Matching Functions", *IZA Discussion Paper* 335 (July).

[23] Gregg, Paul and Barbara Petrongolo (1997): "Random or Non-Random Matching?", *Centre for Economic Performance* Discussion Paper 348 (June).

[24] Gross, Dominique M. (1997): "Aggregate Job Matching and Returns to Scale in Germany", *Economics Letters* 56, pp. 243-248.

[25] Hadri, Kaddour (2000): "Testing for stationarity in heterogeneous panel data", *Econometrics Journal* 3, pp. 148-161.

[26] Haisken-DeNew, John P. and Christoph M. Schmidt (1997): "Interindustry and Interregion Differentials: Mechanics and Interpretation", *Review of Economics and Statistics* 79(3) (August), pp. 516-521.

[27] Hall, Robert (1989): "Comment on Blanchard and Diamond", *Brookings Papers on Economic Activity* 1:1989, pp. 61-71.

[28] Im, K.S., M.H. Pesaran and Y. Shin (1997): "Testing for Unit Roots in Heterogeneous Panels", *mimeo*, Dept. of Applied Economics, University of Cambridge, United Kingdom.

[29] Jackman, Richard, Richard Layard and S. Savouri (1991): "Mismatch: A Framework for Thought", in: Padoa-Schioppa, Fiorella (ed.) *Mismatch and Labor Mobility*, Cambridge University Press, Cambridge U.K., ch. 2.

[30] Levin, Andrew and Chien-Fu Lin (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", *University of California, San Diego* Discussion Paper 92-23 (May).

[31] Li, Hongyi and G.S. Maddala (1996): "Bootstrapping Time Series Models", *Econometric Reviews* 15(2), pp. 115-158.

[32] Maddala, G.S. and Shaowen Wu (1999): "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", *Oxford Bulletin of Economics and Statistics* 61(0) Special Issue (November), pp. 631-652.

[33] Moulton, Brent R. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro-Units", *Review of Economics and Statistics* 72(2), pp. 334-338.

[34] Mumford, Karen and Peter N. Smith (1999): "The Hiring Function Reconsidered: On Closing the Cycle", *Oxford Bulletin of Economics and Statistics* 61(3), 343-364.

[35] Ours, Jan van (1991): "The Efficiency of the Dutch Labour Market in Matching Unemployment and Vacancies", *De Economist* 139, pp. 358-378.

[36] Ours, Jan van, and Geerd Ridder (1995): "Job Matching and Job Competition: Are Lower Educated Workers at the Back of Job Queues?", *European Economic Review* 39 (December), pp. 1717-31.

[37] Petrongolo, Barbara and Christopher Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, forthcoming.

[38] Phillips, Peter C.B. and Hyunsgsik R. Moon (2000): "Nonstationary Panel Data Analysis: An Overview of Some Recent Developments", *Econometric Reviews* 19(3), pp. 263-286.

[39] Pissarides, Christopher (1979): "Job Matchings with State Employment Agencies and Random Search", *Economic Journal* 89 (December), pp. 818-833.

[40] Pissarides, Christopher (1986): "Unemployment and Vacancies in Britain", *Economic Policy* 3, pp. 499-559.

[41] Pissarides, Chistopher (2000): *Equilibrium Unemployment Theory*, 2nd ed., MIT Press, Cambridge Massachusetts.

[42] Storer, P. (1994): "Unemployment Dynamics and Labour Market Tightness", *Journal of Applied Econometrics* 9, pp. 389-419.

[43] Warren, Ronald S. (1996): "Returns to Scale in a Matching Model of the Labor Market", *Economics Letters* 50, pp. 135-142.

[44] Yashiv, Eran (2000): "The Determinants of Equilibrium Unemployment", *American Economic Review* 90(5) (December), pp. 1297-1322.

[45] Yin, Yong and Shaowen Wu (2000): "Stationarity Tests in Heterogeneous Panels", in: B. Baltagi (ed.): *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Advances in Econometrics Vol. 15, Elsevier, Amsterdam, pp. 275-296.

A Data Appendix

The data used for the empirical analysis are yearly data for Western Germany disaggregated into 83 occupational groups. They are reported in official labor statistics as published in the *Amtliche Nachrichten der Bundesanstalt fuer Arbeit*. For all estimations the occupational group collecting all unspecified occupations (other occupations), trainees and apprentices (group 98) is dropped, which leaves us with 82 occupational groups throughout the empirical analysis. The occupational groups were further clustered in nine broad occupational groups (see Table A1 for details). The data include information on unemployment, vacancies, employment levels as well as flows from registered vacancies to employment. The stock data are reported as measured on the 30th of September of each year as reference date. The flows from registered vacancies to employment are reported as the flows aggregated over one year. As an further attempt to reproduce the relevant stocks for job seekers and vacancies for the entire period and check the robustness of our results, the stocks of occupational measures as reported on the reference date of each year are augmented by a correction factor. This factor is obtained by dividing the aggregate flows aggregated over an entire period by the stock of the aggregate measure at the reference date (30.09.). This practice is necessary due to the lack of detailed vacancy flows data on the disaggregate (occupational) level.

The hirings were constructed using an anonymized representative 1% sample of Western German social security records from the German Institute for Employment Research (IAB). The basis of the IAB employment subsample 1975-1995 is the integrated notifying procedure for health insurance, statutory pension scheme and unemployment insurance which is regulated through German legislation. The employment statistics include all employees obliged to pay social insurance contributions and covers about 80% of all employed persons in Western Germany. In total this data set includes 6,711,153 notifications of 483,327 Western Germans (calculated on the basis of final notifications) (cf. Bender *et al.*, 2000). The data contain information on individual characteristics, as well as a firm identifier. They are supplemented by person-related information on periods in which the Federal Employment Service paid benefits from the benefits recipients file. With this information the hirings from different sources could be identified. A flow from employment to employment is identified by observing a change in the firm identifier while an individual is employed (pays social security contributions) from on year to another. There is a negligible measurement error, because a change in the firm identifier could result from a merger or a move between different plants of the same firm. A hiring from unemployment is identified by observing in one year the notification of an individual as stemming from the benefits recipients file (characterizing the individual as an unemployment benefits receiver) while observing the information for the individual in the next year in the employed sample. Accordingly a flow from out of the labor force to employment is identified by missing information on the status in one year (neither employed nor unemployed) while finding this individual in the employed (social security contribution paying) sample in the next year. Because the employment status is identified by social security payments, self-employed and individuals in low-paid jobs which are exempt from social security contributions show up as being out of the labor force. Thus, some flows measured as hirings from out of the labor force are in fact hirings from employment.

In addition, there are some sources of measurement error in distinguishing flows from out of the labor into employment from flows from unemployment. This is due to the fact that there are three sources of mistaking an unemployed as an individual from out of the labor force. Firstly, information from the benefits recipients file could only be matched to the data from the employment subsample when the recipient had a social security number. But between 1.4% and 8% of all notifications in the benefits recipients file are reported without a social security number (see Bender *et al.* 1996). The hiring from an unemployed who received unemployment benefits without having a social security number would be mistaken as a hiring from out of the labor force. Secondly, certain preconditions have to be fulfilled to be entitled to receive unemployment benefits. This means that some people are unemployed without receiving unemployment benefits and could therefore not be identified as unemployed in the data set. Finally the benefit recipients file does not record all benefits paid by the Federal Employment Service. Some payments related to measure of active labor market policies did not show up in the benefits recipients file (for a detailed list see Bender *et al.* 1996). All three source of measurement error lead to a considerable underreporting of hirings from unemployment while measuring to many hirings from out to the labor force.

Because the anonymization procedure leads to missings in the codings of occupations which are not found in the official statistics, only the relative hirings in the sub-sample are regarded as representative. In order to obtain absolute values the relative numbers of hirings for each occupations were multiplied by the respective occupational employment levels from aggregate labor statistics.

To keep the data comparable to the aggregated data from labor statistics and due to some measurement error in the data in the years before 1980 we retain all observations from 1980 to 1995, with the exception of notifications for a second job, for the construction of total hirings per year. The hirings for a specific occupation for a specific year were calculated by comparing all employees at the 30th of September of each year in a specific occupation and with a specific firm identifier to the values of these variables and the variable denoting the employment status (stating whether the observation for an individual is taken from the employment statistics or from the benefits recipients file) at the previous reference date. A problem of this procedure is that one misses short employment spells which take place within the year. The 30th of September was chosen to make the hirings information comparable to the data on the occupational level.

B A Bootstrapping Procedure to Test Panels for Stationarity

Recently, there has been a vivid interest in testing panel data for unit roots or stationarity. As in short time series, a small number of usable observations poses an additional problem for inference. A commonly accepted test strategy for panels is to test individual time series of the single cross-sectional units and then to construct a common test statistic for the entire panel. The challenge is to infer as much as possible about the intertemporal structure of the data without destroying (systematic) cross-sectional dependencies and heterogeneity. This appendix proposes a version of the bootstrap methodology introduced by Li and Maddala (1996) and Maddala and Wu (1999) modified for the use for panel data stationarity tests.

B.1 Issues in Resampling/Creation of the Bootstrap

The conventional bootstrap methodology presented must be modified in order to account for panel structures in form of cross-sectional dependencies. The strategy presented in this section roughly follows Maddala and Wu (1999), Chang (2000) and Yin and Wu (2000), and proposes a way for obtaining bootstrap results for null hypotheses of unit roots or stationarity for panels.[48]

Following Hadri (2000), consider a panel sample of random variables y_{it}, where $i = 1, 2, ..., N$ denotes the cross-sectional dimension and $t = 1, 2, ..., T$ denotes the time dimension. The data are generated by the model:

$$y_{it} = r_{it}(+\gamma_i t) + \epsilon_{it}. \tag{4}$$

We assume that r_{it} is a random walk:

$$r_{it} = r_{it-1} + u_{it}. \tag{5}$$

The ϵ_{it} and u_{it} are mutually independent normally distributed and *i.i.d.* across i and over t.[49] Moreover, $E[\epsilon_{it}] = 0$, $E[\epsilon_{it}^2] = \sigma_\epsilon^2 > 0$, $E[u_{it}] = 0$ and $E[u_{it}^2] = \sigma_u^2 \geq 0$. We want to test stationarity of the series for all i. This is done by testing the hypothesis that $\sigma_u^2 = 0$.[50]

Substituting-in and solving backwards, one can write:

$$y_{it} = r_{i0}(+\gamma_i t) + \sum_{\tau=1}^{t} u_{i\tau} + \epsilon_{it} = r_{i0}(+\gamma_i t) + e_{it}. \tag{6}$$

with $e_{it} = \sum_{\tau=1}^{t} u_{i\tau} + \epsilon_{i\tau}$. Thus, if $\sigma_u^2 = 0$, e_{it} is reduced to ϵ_{it} and thus stationary.

The test is then formulated as:

$$H_0 : Z = 0 \text{ against } H_1 : Z > 0. \tag{7}$$

where $Z = \frac{\sigma_u^2}{\sigma_\epsilon^2}$.[51]

The following steps describe the procedure to obtain bootstrapped test statistics for testing the null.

STEP 1: Create residuals

The sample of residuals to be bootstrapped is created under the null to be tested. Let $\varepsilon_{i,t}$ be the relevant sample residuals for unit i at time t:

$$\forall i, t : \varepsilon_{i,t}^0 = y_{i,t} - r_{i0}(-\gamma_0 t) \longrightarrow E^0 = \begin{bmatrix} \varepsilon_{1,1}^0 & \varepsilon_{2,1}^0 & \cdots & \varepsilon_{N,1}^0 \\ \varepsilon_{1,2}^0 & \varepsilon_{2,2}^0 & \cdots & \varepsilon_{N,2}^0 \\ \cdots & \cdots & \cdots & \cdots \\ \varepsilon_{1,T}^0 & \varepsilon_{2,T}^0 & \cdots & \varepsilon_{N,T}^0 \end{bmatrix}. \tag{8}$$

[48]Strictly speaking, the contribution goes a bit further, since Maddala and Wu (1999) only test for unit roots, while Yin and Wu (2000) perform a Monte Carlo study on the behavior of stationarity tests.

[49]The ϵ_{it} could also be serially correlated.

[50]Under the null, y_{it} is stationary around a level if no time trend is included in model (4), and stationary around a trend if the trend is included.

[51]Note that Z is a pivotal statistic, that is a statistic whose distribution is independent of the true parameter of the model.

In practice, we regress y_{it} on a constant (and where appropriate a time trend) only, and use the coefficient estimates $\hat{r_{i0}}$ and $\hat{\gamma}$ to calculate the residuals ε_{it}^0 under the null. For each individual cross-sectional unit, there is a time series of residuals under the null, and the time series for all units have the same dimensions. The bootstrap proposed below relies on the assumption of a balanced panel, that is that $T_i = T \; \forall i = 1, 2, ..., N$. If the panel to be investigated is not balanced, the creation of bootstrap samples has to be modified.

STEP 2: Create Sample Statistics $\hat{Z}$

In order to create the sample statistic $\hat{Z}$, we use again the residuals ε_{it}^0 created under the null (using estimated coefficients for constant and trend if appropriate) from STEP 1. The Hadri (2000) statistic is then calculated to be:

$$\hat{Z} = \frac{\frac{1}{N}\sum_i^N \frac{1}{T^2}\sum_{t=1}^T S_{it}^2}{\sigma_{\varepsilon^0}^2}, \tag{9}$$

where $S_{it} = \sum_{\tau=1}^t \varepsilon_{i\tau}^0$ is the partial sum of residuals, and $\sigma_{\varepsilon^0}^2$ is a consistent estimator of σ_ε^2 under the null. In order to capture the underlying panel structure as good as possible, one can think of estimating the parameters needed (the constant and γ) not only by using e.g. OLS equation by equation. Panel estimation methods, like Fixed Effects models could be estimated as well as SURE models.

Intuitively, in order to reject the null for a given unit i, this statistic should have an absolute value as large as possible.

STEP 3: Create Bootstrap Sample

Note that now $B \times N$ bootstrap samples have to be generated. The core of the sampling strategy is to permute the time dimension of the panel in order to create bootstrap samples E^* from E^0 while losing as little of the cross-sectional correlation as possible. To preserve the cross-correlation structure of the error term, draw (with replacement) B samples of size T from the sample residual matrix E obtained in STEP 1 *by only permuting entire rows*, that is keeping the cross-section index fixed. Denote the rows of E^0 as $E_t^0 = (\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{N,t}),\ t \in [1, T]$. Then, similar to the simple time series case, the matrix of bootstrap residuals can be generated by permuting $E^0 = (E_1^0, E_2^0, ..., E_T^0)\prime$ B-times:

$$\longrightarrow E^* = (E_1^*, E_2^*, ..., E_T^*)\prime = \begin{bmatrix} \varepsilon_{1,1}^* & \varepsilon_{2,1}^* & \cdots & \varepsilon_{N,1}^* \\ \varepsilon_{1,2}^* & \varepsilon_{2,2}^* & \cdots & \varepsilon_{N,2}^* \\ \cdots & \cdots & \cdots & \cdots \\ \varepsilon_{1,T}^* & \varepsilon_{2,T}^* & \cdots & \varepsilon_{N,T}^* \end{bmatrix}. \tag{10}$$

STEP 4: Create Bootstrap Panel Data

The data necessary to construct the coefficients and statistics for the statistical inference have to be created using the bootstrap residuals and obeying the null. As in the time series case, the bootstrap data for bootstrap samples $b = 1, ..., B$ (using all B bootstrap matrices E^*) are constructed:

$$\forall i, t : y_{i,t}^*(b) = \hat{r_{i0}}(+\hat{\gamma}t) + \varepsilon_{i,t}^*(b),\ i = 1, ..., N,\ t = 1, ...T. \tag{11}$$

STEP 5: Create Bootstrap Statistics

STEP 2 has to be replicated (panel regression of y_{it}^* on a constant and possibly a trend) to obtain new residuals that allow to construct the bootstrap statistics $\hat{Z}^*$. Note that, depending on the estimation technique used in STEP 2 for the sample estimates, this might again call for panel estimation techniques. Of course for obtaining an empirical distribution for each of the $\hat{Z}^*$, this is done B times using all the bootstrap samples previously created. The result of this step should be a set of B statistics $\hat{Z}^*$. The construction of an empirical bootstrap distribution of test statistics for all panels is straightforward:

$$\longrightarrow \hat{F}^* := CDF_B(\hat{Z}^*) \tag{12}$$

STEP 6: Statistical Inference for Cross-sectional Units using Bootstrap Statistics

The entire enterprise has the purpose to infer whether the original sample statistic is such that the null can be rejected. In order to make this inference, one has to find out which quantile of the distribution $\hat{F}^*$ is defined by the original sample statistic $\hat{Z}$, that is which is the empirical (bootstrap) p-value of $\hat{Z}$. The interpretation of this p-value has to take into account under which null it was obtained. The hypothesis to test was whether the null holds for the entire sample rather than only for the series at hand for which the p-value was constructed. Still, the information obtained by this investigation uses (or imposes) more structure than if only every of the i time series would have been investigated separately.

TABLE A1: OCCUPATIONAL GROUP DEFINITIONS

Occupation	Occupational codes (2-digit, from official statistics)	Broad occupational groups*
farmer, fisher	01	1
agricultural administrator	03	1
helper in the agricultural sector, agricultural workers, stockbreeding professions	04	1
gardener, florist	05	1
forester and huntsman	06	1
miner and related professions	07	2
exhauster of mineral resources	08	2
mineral rehasher, mineral burner	09	2
stone processor	10	2
,producer of building materials	11	2
ceramicist, glazier	12	2
glazier, glass processor, glass refiner	13	3
chemical worker	14	2
polymer processor	15	2
paper producer	16	2
printer	17	2
woodworker, wood processor	18	3
metal worker	19	2
moulder, caster, semi-metal cleaner	20	2
metal press workers, metal formers	21	2
turner, cutter, driller, metal polisher	22	2
metal burnisher, galvanizer, enameler	23	2
welder, solderer, riveter, metal gluter	24	2
steel smith, copper smith	25	2
plumber, plant locksmith	26	3
locksmith, fitter	27	3
mechanic	28	3
toolmaker	29	2
metal precision-workers, orthodontists, opticians	30	3
electricians	31	3
assemblers and metal related professions	32	2
spinner, ropemaker	33	2
weaver, other textile producer	34	2
tailor, sewer	35	2
textile dyer	36	2
leather and fur manufacturers, shoemaker	37	2
baker, confectioner	39	3
butcher, fishworkmansip and related	40	3
cooks, convenience food preparatory	41	3
brewer, manufacturer for tobacco products	42	2
milk/fat processor, nutriments producer	43	2
bricklayer, concrete builder	44	3
carpenter, roofer, spiderman	45	3
road/track constructors, demolisher, culture structurer	46	3
helper in the construction sector	47	3
plasterer, tiler, glazier, screed layer	48	3
interior designer, furniture supplier	49	3
joiner, modeler, cartwright	50	3
painter, varnisher and related professions	51	3
goods tester, consignment professions	52	2
unskilled worker	53	9
machinist and related professions	54	2

Table A1 continued

engineer, architect	60	4
chemist, physicist	61	4
technician	62	4
technical specialist	63	4
merchandise manager	68	5
banking professional, insurance merchant	69	6
merchant/ specialist in conveyance, tourism, other services	70	6
conductor, driver, motorist	71	5
navigator, ship engineer, water/air traffic professions	72	5
mail distributer	73	5
storekeeper, worker in storage and transport	74	9
manager, consultant, accountant.	75	6
member of parliament, association manager	76	6
accounting clerk, cashier, data processing expert	77	6
clerk, typist, secretary	78	6
plant security, guard, gate keeper, servant	79	5
other security related professions, health caring professions	80	5
law related professions	81	5
publicist, translator, librarian	82	7
artist and related professions	83	7
physician, dentist, apothecaries	84	8
nurse, helper in nursing, receptionist and related	85	8
social worker, care taker	86	7
professor, teacher	87	7
scientist	88	7
helper for cure of souls and cult	89	7
beauty culture	90	8
guest assistant, steward, barkeeper	91	5
domestic economy, housekeeping	92	5
cleaning industry related professions	93	5
trainee, apprentice	98	**

*The occupations are merged into the following broad occupational groups:

(1) Primary sector
(2) Industry and manufacturing
(3) Crafts
(4) Technical
(5) Service
(6) White collar/ clerical
(7) Social and cultural
(8) Health
(9) Low skilled

** Dropped in empirical analysis.

TABLE 5-1: DESCRIPTIVE STATISTICS (AGGREGATE SAMPLE)

		Average 1980-1995	Share
Labor force		23,363,770	100 %
Employement levels		21,500,186	
Unemployed		1,863,583	7.97 %
Vacancies		211,377	0.90 %
Hirings			
total:	m_{all}	3,475,697	100 %
from nonemployment:	m_X	2,384,188	68.60 %
from employment:	m_E	1,091,509	31.40 %
from unemployment:	m_U	800,725	23.04 %
from out of labor force:	m_{OL}	1,583,460	45.56 %
from registered vacancies:	m_R	2,038,274	58.64 %

Note: All data are aggregated over all 82 occupations and averages over the period 1980-1995.

TABLE 5-2: EMPIRICAL AGGREGATE MATCHING FUNCTIONS

Dependent Variable: Total Hirings (per occupation and year) m_{all}

	(1)*	(2)**	(3)	(4)	(5)	(6)	(7)
Log of registered unemployed in the occupation: **ln*U***	0.447 (0.018)	0.446 (0.169)	0.446 (0.025)	0.460 (0.025)	0.460 (0.021)	0.442 (0.019)	0.456 (0.022)
Log of registered vacancies in the occupation: **ln*V***	0.409 (0.016)	0.379 (0.017)	0.411 (0.022)	0.367 (0.024)	0.399 (0.019)	0.415 (0.017)	0.404 (0.020)
Linear time trend	-0.032 (0.003)	-0.031 (0.003)			-0.054 (0.016)	-0.032 (0.004)	-0.055 (0.016)
Time dummies	No	No	Yes***	Yes***	No	No	No
Dummies for 9 broad occupational groups	No	Yes	No	Yes	No	No	No
Log GDP					0.922 (0.635)		0.922 (0.641)
Log GDP-growth						0.011 (0.016)	0.010 (0.016)
R^2	0.858	0.878	0.860	0.879	0.858	0.863	0.864
Observations	1311	1311	1311	1311	1311	1147	1147

Note: Robust standard errors are in parentheses. Coefficient estimates for dummies are available from the authors upon request.

* F-statistic for H_0: constant returns to scale of the matching function with respect to unemployment and vacancies ($\alpha+\beta=1$): F(1, 1307) = 128.72.

** F-statistic for constant returns to scale: F(1, 1299) = 186.23.

*** All dummies are negative. Dummies for 1986 and all years after are significantly negative at the 1%-level.

TABLE 5-3: EMPIRICAL AGGREGATE MATCHING FUNCTIONS BY SOURCES OF FLOWS

	Dependent Variable: Hirings (per occupation and year) by sources					
	(8) total	(9) from non-employment	(10) from employment	(11) from un-employment	(12) from out of labor force	(13) from registered vacancies
	m_{all}	m_X	m_E	m_U	m_{OL}	m_R*
Log unemployed: $\ln U$	0.447	0.450	0.460	0.520	0.441	0.614
	(0.018)	(0.018)	(0.025)	(0.030)	(0.017)	(0.040)
Log registered vacancies: $\ln V$	0.409	0.413	0.417	0.346	0.451	0.320
	(0.016)	(0.015)	(0.022)	(0.023)	(0.015)	(0.035)
Linear time trend	-0.032	-0.036	-0.022	-0.031	-0.033	-0.011
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.007)
R^2	0.858	0.839	0.810	0.710	0.836	0.616
Observations	1311	1310	1308	1298	1308	640
H_0: constant returns to scale	F(1,1307)	F(1,1306)	F(1,1304)	F(1,1294)	(1,1304)	F(1,636)
	128.72	98.38	67.89	54.66	53.83	6.29

Note: Robust standard errors are in parentheses.

* In contrast to flows from pools of job seekers, data of flows from registered vacancies are only available for a broader classification of 40 occupational groups.

TABLE 6-1: EMPIRICAL AGGREGATE MATCHING FUNCTIONS BY SOURCES OF FLOWS AND BROAD OCCUPATIONAL GROUPS

Dependent Variable: Hirings (per occupation and year) by sources						
	(14) total hirings m_{all}		(15) hirings from non-employment m_X		(16) hirings from unemployment m_U	
Log unemployed: $\ln U$	0.446*	(0.018)	0.456*	(0.017)	0.538*	(0.029)
Log registered vacancies: $\ln V$	0.379*	(0.017)	0.379*	(0.016)	0.337*	(0.024)
Linear time trend	-0.031*	(0.003)	-0.034*	(0.003)	-0.031*	(0.004)
Group 1 (primary/agricultural)	-0.152	(0.082)	-0.046	(0.099)	0.044	(0.149)
Group 2 (industry/manufacturing)	-0.126*	(0.021)	-0.127*	(0.023)	0.091*	(0.029)
Group 3 (crafts)	0.099*	(0.019)	0.160*	(0.020)	0.310*	(0.032)
Group 4 (technical occupations)	-0.204*	(0.055)	-0.378*	(0.057)	-0.635*	(0.082)
Group 5 (services)	-0.040	(0.038)	-0.041	(0.037)	-0.192*	(0.050)
Group 6 (white collar/clerical)	0.526*	(0.041)	0.326*	(0.047)	-0.024	(0.046)
Group 7 (social and cultural)	-0.067	(0.057)	-0.009	(0.061)	-0.506*	(0.061)
Group 8 (health)	0.332*	(0.042)	0.306*	(0.038)	-0.393*	(0.082)
Group 9 (low skilled)	-0.014	(0.060)	0.008	(0.062)	0.237*	(0.069)
R^2	0.878		0.855		0.747	
Observations	1311		1310		1298	
H_0: constant returns to scale	$F(1,1299) = 186.23$		$F(1,1298) = 134.91$		$F(1,1286) = 52.41$	

Note: Robust standard errors are in parentheses. For a detailed description of the definition of occupational groups, refer to the data appendix. Coefficients and standard errors are obtained using a 2-step procedure following Haisken-DeNew and Schmidt (1997). The first step provides robust OLS-results with group 1 as reference group. On the second step, dummy coefficients are renormalized as deviations weighted by group-size from a group-size weighted mean (where group size refers to the number of occupations contained in a group) and standard errors are adjusted accordingly. The procedure therefore renders the results independent from the choice of the reference group.

* Coefficient significant at the 1%-level.

TABLE 6-2: MATCHING FUNCTIONS DISAGGREGATED BY BROAD OCCUPATIONAL GROUP

	Dependent Variable:									
	(17) total hirings m_{all}					(18) hirings from non-employment m_X				
	α	β	CRS	Obs.	R^2	α	β	CRS	Obs.	R^2
Group 1	0.401	0.533	F(1,76)	80	0.728	0.388	0.614	F(1,76)	80	0.675
(primary sector)	(0.091)	(0.072)	1.18			(0.118)	(0.087)	0.00		
Group 2	0.372	0.359	F(1,443)	447	0.814	0.345	0.379	F(1,442)	446	0.776
(industry/manufacturing)	(0.025)	(0.027)	209.38**			(0.023)	(0.022)	166.91**		
Group 3	0.594	0.442	F(1,284)	288	0.948	0.675	0.402	F(1,284)	288	0.937
(crafts)	(0.018)	(0.015)	4.44*			(0.020)	(0.017)	14.73**		
Group 4	0.781	0.425	F(1,60)	64	0.930	0.775	0.382	F(1,60)	64	0.901
(technical occupations)	(0.046)	(0.042)	29.64**			(0.069)	(0.055)	11.90**		
Group 5	0.608	0.416	F(1,156)	160	0.929	0.654	0.382	F(1,156)	160	0.934
(services)	(0.045)	(0.042)	0.96			(0.047)	(0.043)	2.47		
Group 6	0.420	0.336	F(1,92)	96	0.920	0.534	0.270	F(1,92)	96	0.898
(white collar/clerical)	(0.060)	(0.059)	90.00**			(0.071)	(0.068)	46.85**		
Group 7	*0.211*	0.412	F(1,92)	96	0.845	0.319	0.331	F(1,92)	96	0.830
(social and cultural)	(0.134)	(0.099)	65.60**			(0.142)	(0.106)	54.33**		
Group 8	0.593	0.343	F(1,44)	48	0.912	0.640	0.326	F(1,44)	48	0.945
(health)	(0.050)	(0.051)	2.84			(0.034)	(0.039)	1.11		
Group 9	0.425	0.503	F(1,28)	32	0.865	0.382	0.399	F(1,28)	32	0.838
(low skilled)	(0.058)	(0.048)	1.34			(0.058)	(0.047)	12.99**		

Note: α is the estimated coefficient for $\ln U$, β is the estimated coefficient for $\ln V$. All models are estimated including a constant and a linear time trend. Robust standard errors are in parentheses. Insignificant coefficients are set in italics. CRS: Contains the F-statistics for H_0: The matching function exhibits constant returns to scale. For a detailed description of the definition of occupational groups, refer to the data appendix.

* H_0: Constant returns to scale can be rejected on the 5%-level.

** H_0: Constant returns to scale can be rejected on the 1%-level.

TABLE 6-3: MATCHING FUNCTIONS DISAGGREGATED BY BROAD OCCUPATIONAL GROUP

	Dependent Variable: (17) total hirings m_{all}					(19) hirings from unemployment m_U				
	α	β	CRS	Obs.	R^2	α	β	CRS	Obs.	R^2
Group 1 (primary sector)	0.401 (0.091)	0.533 (0.072)	F(1,76) 1.18	80	0.728	0.605 (0.219)	0.542 (0.136)	F(1,73) 1.54	77	0.483
Group 2 (industry/manufacturing)	0.372 (0.025)	0.359 (0.027)	F(1,443) 209.38**	447	0.814	0.285 (0.039)	0.374 (0.036)	F(1,437) 141.40**	441	0.593
Group 3 (crafts)	0.594 (0.018)	0.442 (0.015)	F(1,284) 4.44*	288	0.948	0.828 (0.042)	0.211 (0.031)	F(1,284) 1.52	288	0.746
Group 4 (technical occupations)	0.781 (0.046)	0.425 (0.042)	F(1,60) 29.64**	64	0.930	1.257 (0.090)	0.161 (0.067)	F(1,59) 27.46**	63	0.846
Group 5 (services)	0.608 (0.045)	0.416 (0.042)	F(1,156) 0.96	160	0.929	0.781 (0.062)	0.302 (0.053)	F(1,154) 5.31*	158	0.876
Group 6 (white collar/clerical)	0.420 (0.060)	0.336 (0.059)	F(1,92) 90.00**	96	0.920	0.678 (0.077)	*0.154* (0.087)	F(1,91) 20.20**	95	0.923
Group 7 (social and cultural)	*0.211* (0.134)	0.412 (0.099)	F(1,92) 65.60**	96	0.845	0.434 (0.138)	0.423 (0.106)	F(1,92) 9.78**	96	0.874
Group 8 (health)	0.593 (0.050)	0.343 (0.051)	F(1,44) 2.84	48	0.912	1.023 (0.037)	0.393 (0.060)	F(1,44) 85.32**	48	0.958
Group 9 (low skilled)	0.425 (0.058)	0.503 (0.048)	F(1,28) 1.34	32	0.865	0.472 (0.068)	0.215 (0.058)	F(1,28) 23.14**	32	0.780

Note: α is the estimated coefficient for $\ln U$, β is the estimated coefficient for $\ln V$. All models are estimated including a constant and a linear time trend. Robust standard errors are in parentheses. Insignificant coefficients are set in italics. CRS: Contains the F-statistics for H_0: The matching function exhibits constant returns to scale. For a detailed description of the definition of occupational groups, refer to the data appendix.

* H_0: Constant returns to scale can be rejected on the 5%-level.

** H_0: Constant returns to scale can be rejected on the 1%-level.

TABLE 7-1: DESCRIPTIVE STATISTICS (EDUCATION GROUPS, AGGREGATE)

		Total Shares (Averages 1980-1995)	Education Group 1 (low)	Education Group 2 (middle)	Education Group 3 (high)
Relative group sizes:					
sample population		100 %	27.40 %	60.76 %	11.84 %
employed		100 %	26.98 %	61.66 %	11.36 %
non-employed		100 %	33.81 %	46.37 %	19.82 %
Hirings:					
total matchings:	$\boldsymbol{m_{all}}$	100 %	34.40 % (100 %)	51.01 % (100 %)	14.51 % (100 %)
from employment:	$\boldsymbol{m_E}$	31.40 %	20.55 % (18.76 %)	65.24 % (40.11 %)	14.21 % (30.74 %)
from nonemployment:	$\boldsymbol{m_X}$	68.60 %	40.74 % (81.24 %)	44.60 % (59.89 %)	14.65 % (69.26 %)
from unemployment:	$\boldsymbol{m_U}$	23.04 %	26.43 % (17.70 %)	60.59 % (27.32 %)	12.98 % (20.61 %)
from out of labor force:	$\boldsymbol{m_{OL}}$	45.56 %	47.98 % (63.54 %)	36.52 % (32.57 %)	15.50 % (48.65 %)

Note: All data are aggregated over all 82 occupations and averages over the period 1980-1995. Table entries are shares of the education group characteristics with respect to total shares (that is they add up to 100 % horizontally). Entries in parentheses are shares of the respective flows with respect to the respective educational group (that is they add up to 100 % vertically). Therefore, the share of a given flow of a given educational group with respect to total hirings can be calculated by multiplying the entry with the share of that flow with respect to total hires (that is the first column value).

TABLE 7-2: EMPIRICAL AGGREGATE MATCHING FUNCTIONS BY EDUCATION GROUPS AND SOURCES

Dependent Variable: All Hirings m_{all} and Hirings from Non-Employment m_X by Education Groups and Sources

	(35)	(36)	(37)	(38)	(39)	(40)
	m_{all} Group 1 (low)	m_{all} Group 2 (middle)	m_{all} Group 3 (high)	m_X Group 1 (low)	m_X Group 2 (middle)	m_X Group 3 (high)
Log unemployed: **ln*U***	0.607	0.452	0.512	0.652	0.484	0.520
	(0.048)	(0.030)	(0.043)	(0.053)	(0.032)	(0.050)
Log registered	0.362	0.465	0.482	0.382	0.475	0.521
vacancies: **ln*V***	(0.038)	(0.023)	(0.036)	(0.039)	(0.024)	(0.044)
Linear time trend	-0.059	-0.022	-0.002	-0.062	-0.027	-0.002
	(0.010)	(0.005)	(0.008)	(0.011)	(0.006)	(0.009)
R^2	0.405	0.783	0.528	0.409	0.701	0.499
Observations	1311	1311	1311	1311	1311	1311
H_0: constant returns to	F(1,1307)	F(1,1307)	F(1,1307)	F(1,1307)	F(1,1307)	F(1,1307)
scale	0.76	17.49*	0.02	0.69	2.64	1.07

Note: Robust standard errors are in parentheses. Group definitions are as follows: Members of Group 1(low education) have neither finished high school (*Abitur*) nor an apprenticeship successfully. Group 2 (intermediate education) members have either finished high school (*Abitur*) or an apprenticeship or both successfully. Members of Group 3 (high education) hold a degree from a university or an applied university (*Fachhochschule*).

* H_0: Constant returns to scale can be rejected on the 1% level.

TABLE 7-3: EMPIRICAL AGGREGATE MATCHING FUNCTIONS BY EDUCATION GROUP

Dependent Variable: All hirings m_{all} (per occupation and year) by education

	(41) Group 1 (low) m_{all}		(42) Group 2 (middle) m_{all}		(43) Group 3 (high) m_{all}	
Log unemployed: **ln*U***	0.612*	(0.044)	0.450*	(0.030)	0.452*	(0.043)
Log registered vacancies: **ln*V***	0.468*	(0.042)	0.431*	(0.024)	0.414*	(0.037)
Linear time trend	- 0.063*	(0.009)	- 0.021*	(0.004)	*0.003*	(0.007)
Group 1 (primary/agricultural)	*- 0.115*	(0.270)	- 0.270*	(0.102)	*- 0.117*	(0.125)
Group 2 (industry/manufacturing)	0.746*	(0.067)	- 0.138*	(0.030)	- 0.729*	(0.051)
Group 3 (crafts)	0.441*	(0.038)	0.212*	(0.028)	- 0.259*	(0.053)
Group 4 (technical occupations)	- 2.837*	(0.317)	- 0.525*	(0.152)	0.856*	(0.111)
Group 5 (services)	- 0.426*	(0.121)	*- 0.032*	(0.058)	0.501*	(0.079)
Group 6 (white collar/clerical)	- 0.707*	(0.156)	0.907*	(0.043)	1.159*	(0.087)
Group 7 (social and cultural)	- 1.280*	(0.214)	*- 0.178*	(0.096)	1.371*	(0.072)
Group 8 (health)	- 0.488	(0.236)	*0.008*	(0.112)	0.505	(0.214)
Group 9 (low skilled)	0.369	(0.101)	- 0.293*	(0.085)	0.258*	(0.083)
R^2	0.581		0.784		0.653	
Observations	1311		1311		1311	
H_0: constant returns to scale	F(1,1299) = *3.55*		F(1,1299) = 35.70**		F(1,1299) = 16.01**	

Note: Robust standard errors are in parentheses. Coefficients without asterisk are significant at the 5%-level. Italics indicate that the null that the coefficient is zero cannot be rejected at the 5%-level. For a detailed description of the definition of occupational groups, refer to the data appendix. Coefficients and standard errors are obtained using a 2-step procedure following Haisken-DeNew and Schmidt (1997). The first step provides robust OLS-results with group 1 as reference group. On the second step, dummy coefficients are renormalized as deviations weighted by group-size from a group-size weighted mean (where group size refers to the number of occupations contained in a group) and standard errors are adjusted accordingly. The procedure therefore renders the results independent from the choice of the reference group.

* Significant at the 1%-level.

** H_0 can be rejected at the 1%-level.

TABLE 7-4: EMPIRICAL AGGREGATE MATCHING FUNCTIONS BY EDUCATION GROUP

Dependent Variable: Hirings from non-employment m_X (per occupation and year) by education

	(44) Group 1 (low) m_X		(45) Group 2 (middle) m_X		(46) Group 3 (high) m_X	
Log unemployed: **ln*U***	0.671*	(0.048)	0.490*	(0.033)	0.456*	(0.050)
Log registered vacancies: **ln*V***	0.483*	(0.042)	0.440*	(0.027)	0.454*	(0.044)
Linear time trend	- 0.066*	(0.009)	- 0.026*	(0.006)	*0.003*	(0.008)
Group 1 (primary/agricultural)	*- 0.043*	(0.280)	*- 0.153*	(0.158)	*- 0.055*	(0.161)
Group 2 (industry/manufacturing)	0.746*	(0.070)	- 0.136*	(0.036)	- 0.758*	(0.059)
Group 3 (crafts)	0.510*	(0.040)	0.258*	(0.030)	- 0.251*	(0.064)
Group 4 (technical occupations)	- 2.879*	(0.330)	- 0.717*	(0.173)	0.731*	(0.106)
Group 5 (services)	- 0.505*	(0.128)	*- 0.027*	(0.056)	0.563*	(0.086)
Group 6 (white collar/clerical)	- 0.954*	(0.187)	0.705*	(0.050)	1.018*	(0.093)
Group 7 (social and cultural)	- 1.137*	(0.217)	*- 0.094*	(0.104)	1.522*	(0.083)
Group 8 (health)	*- 0.417*	(0.231)	*0.006*	(0.099)	0.495	(0.208)
Group 9 (low skilled)	0.253	(0.119)	- 0.311*	(0.082)	0.366*	(0.082)
R^2	0.573		0.732		0.610	
Observations	1311		1311		1311	
H_0: constant returns to scale	F(1,1299) = 10.72**		F(1,1299) = 7.60**		F(1,1299) = 6.46**	

Note: Robust standard errors are in parentheses. Coefficients without asterisk are significant at the 5%-level. Italics indicate that the null that the coefficient is zero cannot be rejected at the 5%-level. For a detailed description of the definition of occupational groups, refer to the data appendix. Coefficients and standard errors are obtained using a 2-step procedure following Haisken-DeNew and Schmidt (1997). The first step provides robust OLS-results with group 1 as reference group. On the second step, dummy coefficients are renormalized as deviations weighted by group-size from a group-size weighted mean (where group size refers to the number of occupations contained in a group) and standard errors are adjusted accordingly. The procedure therefore renders the results independent from the choice of the reference group.

* Significant at the 1%-level.

** H_0 can be rejected at the 1%-level.

Figure A1: Employment Levels of Broad Occupational Groups in Data Set