MSc Economics

Labor market polarization in the U.S. - An empirical approach

A Master's Thesis submitted for the degree of
"Master of Science"

supervised by

Tamás K. Papp

Fruzsina Mayer

1226492

Vienna, 10. June 2014

MSc Economics

Affidavit

I, Fruzsina Mayer

hereby declare

that I am the sole author of the present Master's Thesis,

Labor market polarization in the U.S. - An empirical approach

46 pages, bound, and that I have not used any source or tool other than those referenced or any other illicit aid or tool, and that I have not prior to this date submitted this Master's Thesis as an examination paper in any form in Austria or abroad.

Vienna, 10. June 2014

Signature

Contents

List of Figures

List of Tables

Abstract

Labor market polarization is an ongoing structural change of the U.S. labor market which consists of the "hollowing out" of the wage and skill distribution. This paper examines the effects of this process on individual workers, focusing on the feedback effect from the changes in the labor market aggregates represented by employment-to-population ratios and by the employment shares of the different occupational groups on the transition between unemployment and the three main occupational groups. The main occupational groups defined by skill and wage levels are the routine, non-routine manual and non-routine cognitive occupations. In order to estimate the effects of the aggregates on the individual level, transition probabilities are estimated with logistic regressions. The 1991 and the 2001 panels of the Survey of Income and Program Participation are used as data source for the analysis, in contrast to most of the literature which uses CPS data. The findings show that polarization was present in the period of the 2001 panel, while it cannot be detected in the 1991 panel. The results suggest that, as expected, the aggregate rates had an effect on the transition toward unemployment in both panels, but they did not influence the movement between the routine, non-routine cognitive and non-routine occupational groups. The results also imply that in the 2001 panel the probability of acquiring a routine job after job change is smaller for those who were employed in any occupational category than for those who were unemployed before the change. This indicates that on the level of individuals the polarization made the direct transition from any job towards the routine occupations more difficult.

1 Introduction

Movements in the labor market are closely related to the changes in the aggregate output, as labor is one of the main factors of production. Looking at the deviations of aggregate production and the employment rate from their long run trend one can observe this co-movement in U.S. data. Therefore, when a rather large and permanent discrepancy appeared between the cyclical behavior of the output and the employment rate, namely that the latter caught up after economic downturns slower compared to other economic indicators, among them the aggregate production, it became a topic of economic research. The phenomenon I am referring to is the so-called jobless recovery, which is observed in the U.S. economy since the beginning of the 1990s (Jaimovich and Siu, 2013). So far there is no agreement in the literature about the cause of the slow increase in the employment after the last three recessions. One possible explanation of the issue offered by Jaimovich and Siu (2013) is that job market polarization was faster in the periods of the most recent recessions.

Job polarization is a structural change of the labor market which has been present in the U.S. for the past thirty years and it is also prevalent in the economies of Europe (Autor, 2010; Goos, Manning, and Salomons, 2010). Polarization is defined shortly as the "hollowing out" of the wage distribution. This definition captures the essence of the polarization; however, it does so without expressing the complexity of the process. Polarization did not affect the lower and the upper ends of the wage distribution equally over time and it was also influenced by the business cycles.[1] In the 1990s polarization was most favorable for the workers on the higher end of the wage distribution while in the 2000s the direction of the process changed and the lower end of the distribution gained relatively more both in terms of employment share and wages (Autor, 2010). Moreover, polarization was faster during recessionary periods thus contributing to the above mentioned jobless recoveries in the last two decades (Jaimovich and Siu, 2013; Tüzemen and Willis, 2013).

Labor market polarization is also closely related to the skill-demand of the different occupations, as there is a close correspondence between the skill and the wage level of the occupations. Low skill occupations are the ones which are positioned on the lower end of the wage distribution, while the jobs which require more skilled employees are also the ones with higher wage rates. This link between the wages and the skill requirement of jobs leads to the result that polarization can be described as the disappearance of middle skill jobs. Autor

[1]The nature of the relationship between the economic cycles and the job polarization is a topic of discussion, for further details see Section 2.

(2010) shows that while in the 1980s the occupations at the higher percentiles of the skill distributions gained and the ones at the lower tail lost in terms of employment share, so that a positive relationship was observable between the skill requirements and the employment share of the occupations, this was not true anymore in the 1990s. In this period the occupations at the lower and higher end witnessed an increase in their employment share, which was relatively smaller for the low-skilled occupations, while the professions in the middle of the skill distribution showed a decline. Similarly, in the first decade of the 21^{st} century the middle skill occupations faced further decline whereas the employment share of the lower tail professions increased far more than previously (Autor, 2010; Autor and Dorn (2013)).

Ordering the occupations based on their skill content and exploring polarization from this point of view makes it possible to apply another type of classification of the jobs, which is generally used by the literature. Occupations can be categorized by the types of labor tasks they include. For example Autor et al. (2003) differentiate between routine and non-routine occupations based on the ratio of tasks in a given job which can be computerized. Within these groups they also make a distinction between manual and cognitive occupations.[2] Matching this categorization with the skill distribution shows that non-routine manual tasks are characteristic of the lower tail of the skill distribution and non-routine cognitive activities are attributed to high skill occupations. Routine manual and routine cognitive tasks are prevalent in the middle skill professions (Jaimovich and Siu, 2013). However, this means that due to polarization exactly the jobs belonging to the latter two groups are the ones which vanish (Autor et al., 2003; Autor and Dorn (2013)). The disappearance of the routine occupations causes an increase in wage inequality (Autor et al., 2008) and also as routine jobs have the highest employment share, their vanishing can lead to higher unemployment and lower employment-to-population ratios (Jaimovich and Siu, 2013; Autor, 2010). Profound investigation of polarization is therefore useful for formulating economic policy in response to these structural changes of the labor market.

A significant part of the literature focusing on polarization generally examines the movements between the three or four occupational categories defined by task content without considering the flows between employment and unemployment (e.g. Autor et al., 2003; Autor and Dorn (2013)). On the one hand, excluding the unemployment state does not affect the results when the focus is on the effects of polarization on the structure of employment in the long run or the causes leading to polarization. On the other hand including the flows between employment and

[2]For further details see Section 2.

unemployment in the investigation can result in a more precise understanding of the underlying process of polarization. Expanding the scope of the examination in this direction allows for considering the effects of the polarization on the individual workers as well. Even though polarization is observable on the level of the aggregates, without including unemployment little can be said about how aggregate changes relate to the changes on the level of individuals. For example, an increase in the employment share of non-routine occupations can be caused by workers moving away from the routine professions towards the non-routine ones or it can be the result of new labor market entrants being employed in these occupations while the middle-skill workers who left or lost their jobs became unemployed or left the labor force altogether. In other words, looking only at the aggregates one can miss important implications of the polarization which could be useful when it comes to drawing up policy advice on this topic. The more recent works in the topic of polarization are closer to addressing this issue as they explore the relationship between polarization and the business cycle, which requires the inclusion of the flows between employment and unemployment in the analysis (Jaimovich and Siu, 2013; Foote and Ryan, 2012).

My goal is to assess the effect of polarization on the level of individuals, therefore I include the unemployment state into my examinations. I am interested mostly in determining how the presence of polarization in the U.S. economy affects the individual's labor decisions, and whether there is a significant difference in this effect for the various skill-groups and occupational categories. I use two panels of the Survey of Income and Program Participation (SIPP) as my main data source, which allows me to look at employment history changes of the sample members. My intention is to estimate transition probabilities between unemployment and employment and between the different occupational categories using logistic regressions to examine whether there is a feedback effect from the aggregate level on these probabilities. I execute this estimation on the 1991 and on the 2001 SIPP panels in order to compare the results for the same model in different time periods. The most important variables for my analysis are the ones containing the change in the employment-to-population ratios for the three main occupational categories (routine, non-routine manual and non-routine cognitive) in the period when individuals changed their occupation or job.

In the case of both panels the changes in the employment-to-population ratios of the non-routine occupations have a significant effect on the probability of moving to unemployment in case of a job change. The coefficient of the change in the non-routine manual ratio has a negative sign, implying that in periods when this ratio increased, the probability of becoming unemployed after the job change was smaller in the sample. The change in the employment-to-population ratio for the non-routine cognitive occupations has an opposing sign, which can happen because in the sample periods of the two panels the ratio of this category moved in the opposite direction as the rate for the non-routine manual occupations. This result remains significant even after controlling for the occupational category to which the individuals belonged to before the job change. However, the job loss probability is not affected significantly by the occupational category before the change.

The logistic regressions also show that the employment-to-population ratios have no significant effect on the probability of acquiring a high-, middle- or low-skilled job after the job change in the 1991 an 2001 panels. In the regressions for the 1991 panel the effect of the occupation before the change is significant. The results show here that workers moved with higher probability to the same type of job as the one they had before the change, therefore no polarizing pattern is observable in the movements. The same regressions for the 2001 panel show that the probability of being employed in a routine occupation after the change is smaller in this panel even for those who worked in routine occupations before the change compared to the ones who were unemployed beforehand. In the other two categories those who moved to the same occupation as the one they had before the change still had a higher probability to do so than those who were employed in other categories or were unemployed before the change.

These results all in all show that polarization is present more clearly in the 2001 panel, where the transition probabilities seem to imply that individual workers do not move to routine jobs straight after their previous employment, but rather spend some time in unemployment before they do so. Comparing average predictive differences between the two sets of regressions for the two panels also reveals, that those, who had a routine job before the job change faced smaller disadvantage in terms of probability in the 2001 panel when it comes to moving towards non-routine occupations than in the 1991 panel.

After this short introduction of the issue at hand the next section contains a review of the related literature. In Section 3 I introduce the SIPP data and describe the variables used in the analysis. Section 4 then summarizes the methodology I apply with focus on multilevel logit models. In Section 5 I give a detailed description of the results and finally Section 6 contains the conclusion and further directions in which this research could be expanded.

2 Literature overview

The papers addressing the polarization approach different aspects of the process and a certain change can be observed in the general topic of these papers in th last few years. A large part of the polarization literature aims at explaining the process and its causes by testing the predictions of various models against the data (e.g. Acemoglu, 1999; Autor et al., 2003). More recent papers, however, examine the relationship between polarization and the business cycles (e.g. Jaimovich and Siu, 2013; Foote and Ryan, 2012). Papers which belong to this latter category are motivated by the appearance of jobless recoveries mentioned in Section 1. Apart from the studies belonging to these two categories there are also articles which focus on the effects of the process (e.g. Autor, 2010) and provide policy advice. In the following I will briefly introduce the most important pieces of the polarization literature in order to show what are the similarities and differences between these papers and my research.

The most widely held view is that polarization can be best explained with the theory of skill-biased technological change (SBTC) (Autor, 2010). The SBTC is such an improvement in technology which increases the relative demand for skilled labor and therefore it leads to an increase in the wage inequality between skilled and unskilled workers. This feature of the SBTC can explain the upsurge in the wages and employment share of non-routine cognitive or high skilled workers in the period of polarization and even before.

Acemoglu (1999) uses a search and matching framework to show that the skill-biased technological change can affect the job composition of the labor market by moving it from a pooling to a separating equilibrium in which the demand for high-skilled workers increases and job opportunities for the less skilled are scarce. Acemoglu (1999) also connects the skill with educational attainment and by doing so he gives an explanation to the increasing returns of higher education. This latter phenomenon accompanies polarization over its whole presence in the data (Autor, 2010).

Similarly to Acemoglu (1999), Autor et al. (2003) are also focusing on the skill-biased technological change as an explanation of the higher relative demand for workers with college education starting from the 1970s; however, they build a general equilibrium model with multiple industries to do so where SBTC is identified with computerization. In their model economy computers are substitutes of human work force in job tasks which are routine in the sense that they can be described by a set of well-defined rules, while they are complements of the workers in job tasks which do not belong to the previous category. Tasks belonging to the latter group are called non-routine. Within each group they also differentiate between cognitive (analytic or interactive) and manual tasks. In the model economy the constant decrease in the price of computer capital results in a decline in the relative wages of routine workers and an increase in computerization in those industries where the share of routine workers was originally high. At the same time the demand for non-routine cognitive job tasks increases resulting in higher relative wages for non-routine cognitive workers.

Autor et al. (2003) use data up until 1998 and they focus on the effect of polarization on the higher end of the wage distribution. They claim that it is not clear that the non-routine manual tasks are affected by computerization therefore they do not explain the increase of wages and employment share of the occupations containing mostly these types of tasks in the 1990s. Autor et al. (2003) use the Dictionary of Occupational Titles (DOT) and the O*NET database to create an index which shows the ratio of routine tasks in each occupation. Based on this index they categorize the occupations in the Current Population Survey (CPS) data which they use in the analysis. Testing the predictions of their model on the data Autor et al. (2003) find that indeed computerization was faster in those industries where the share of routine employees was higher previously and that in these sectors of the economy the relative demand for high-skill (or more educated) workers increased. They also observe in the data that the task composition of occupations changed a lot in industries were the computerization was faster.

The model of Autor et al. (2003) and its predictions was challenged by Autor and Dorn (2013) as the latter authors found that it is a major problem of the model that it does not explain the increase in the employment share and wages of the lower tail of the skill distribution, especially in the service occupations. Autor and Dorn (2013) argue that one has to take into consideration the effect of consumer preferences on the skill demand just as much as the effect of the skill-biased technological change. Similarly to Autor et al. (2003) they use the DOT to classify the occupations based on their routine skill content and test the predictions of their spatial equilibrium model on CPS data. Their findings are similar to the ones of Autor et al. (2003), and their results also show that low-skill workers moved towards non-routine manual – mainly service – occupations in those regions where the share of routine employment was high previously. Autor and Dorn (2013) argue that this happens because in the non-routine manual occupations computers cannot substitute out human labor force, but computerization can have a positive effect on their wages if it raises the demand for the services offered by them.

Even though skill-biased technological change is a good explanation for the job polarization, there are other factors which can contribute to this phenomenon. The declining real wage and the decreasing union coverage were both considered as possible reasons of the increasing wage inequality in the 1980s. Autor et al. (2008) examine whether the increase in the wage inequality in this period was caused by the changes in the real minimum wage, but they find that the SBTC does far better at explaining the increasing inequalities. As for the deunionization, Autor (2010) finds, that the decline in the union participation does not have a large enough effect to explain the permanent increase in the wage inequality. Similarly, outsourcing of jobs – mostly routine ones – to other countries was judged to be only marginally contributing to polarization as Autor (2010) concludes.

In the aftermath of the Great Recession the main question in connection with the polarization became that how it is related with the business cycle. One of the first papers tackling these questions was the one by Jaimovich and Siu (2013). Jaimovich and Siu (2013) aim at explaining the jobless recoveries after the recessions in the last two and a half decades with the help of the polarization. They use a search and matching model with labor market frictions to this end, where originally high- and low-skilled (respectively non-routine cognitive and routine) agents work in separate markets. Adding routine-biased technological change and recession to the model results in the movement of low-skill (routine) workers' towards the high-skilled market through the switching market. However, in the model it is more difficult for the ex-routine workers to find a job on the switching market which leads to higher unemployment rates over a longer period of time. In a counterfactual experiment Jaimovich and Siu (2013) also show that if the routine-biased technological change did not take place then the recoveries would be faster in the last recessions than they actually were.

Jaimovich and Siu (2013) also estimates transition rates between the different occupational categories in the recessionary periods both before polarization started and when it was present in the U.S. labor market. As a result they find that the probability of job finding in the routine category became significantly smaller for all workers in the recessions which happened during the period of polarization compared to the earlier ones.

The results of Jaimovich and Siu (2013) are very suggestive and they motivated others to look into the relationship of the business cycles and the polarization more deeply. Foote and Ryan (2012), for example, use CPS data for the period between 1979 and 2012 in order to re-evaluate the findings of Jaimovich and Siu (2013).[3] They examine the job finding and separation rates for the non-routine manual (low-skill), the non-routine cognitive (high-skill) and the routine (middle-skill) occupations in recessionary periods. However, they find that routine workers do not have higher separation rates in the recessions as expected compared to the ones for the other groups. They also find that routine workers who became unemployed are most likely to find a routine job later rather than to switch to occupations in the other two skill categories or to leave the labor force. The main difference Foote and Ryan (2012) finds between the more recent recessions (in 2001 and 2007) and the earlier ones is that lately even the non-routine cognitive group faces a large increase in their separation rate. Therefore they conclude that polarization cannot be the only reason behind the jobless recoveries. Tüzemen and Willis (2013) also support this conclusion of Foote and Ryan (2012) as examining CPS data they find that even though more polarization happened during recessionary periods, the larger part of the decline in the employment share of routine occupations took place in non-recessionary periods.

Even though the link between polarization and the business cycle is still debated, the effects of polarization on the different population categories are far less controversial. Autor (2010) gives a very accurate summary of how the different gender and educational groups are affected by this process. Similarly to the previously quoted literature, Autor (2010) uses CPS data in his research and finds that there is a significant difference in how males and females reacted to the polarization. While the employment share of males was increasing in the non-routine cognitive and non-routine manual occupations with almost the same rate and declined in routine occupations since the 1990s, for females the employment share increased more in the high-skilled category in the same period. Autor (2010) also finds that the return on educational attainment is increasing in the level of education over time and that polarization hurt those the most who did not even receive a high school diploma. Tüzemen and Willis (2013) have similar results for the different gender and educational groups; however, they also examine the reaction of different age categories. Their results show that elderly workers are more likely to move towards non-routine occupations while employees younger than 24 years are more probable to work in some low-skill routine occupation.

An interesting common point of the above introduced literature is that most of them are using the CPS as their main data source, which is the primary source of employment related data in the U.S. The CPS is collected on monthly basis and it contains information about several aspects of the labor market properties

[3] Foote and Ryan (2012) are reacting to an earlier version of Jaimovich and Siu (2013) from (2012).

of its members. The CPS is an address-based survey which means that if the people who were interviewed in the beginning of the survey move away, then they are not followed, but the new residents on the same address will be interviewed. Consequently there are inconsistencies in the data caused by the movers.

Before moving on to the next section, it is also worth to note that there are differences in the way the various papers define the low- middle- and high-skill occupational categories. Autor et al. (2003) created an index which determines the share of routine tasks in the given occupations and they based their categorization on it. Autor and Dorn (2013) follow a similar strategy as they order occupations into groups based on the ratio of abstract, manual and routine tasks in them. Autor (2010) on the other hand aggregates the small occupational categories given in the CPS into ten large groups from which three-three are non-routine cognitive and manual respectively and four are considered to be routine. Foote and Ryan (2012) use this categorization but they aggregate even it further making a distinction between high-skill (non-routine cognitive), low-skill (non-routine manual), construction, manufacturing and other middle-skill (other routine) occupations. The classification applied by Tüzemen and Willis (2013) is somewhat similar, as they divide the economy into four sectors: manufacturing, construction, education and health and "other". They justify this set up with the fact the routine occupations are in surplus in the first two sectors, while the non-routine cognitive occupations have a high share in the education and health sector. Jaimovich and Siu (2013) also use their own classification and they claim that their results are robust to changes in the categorization of occupations. All of these classifications are valid, they mostly differ in the complexity of the method used in generating them. In this essay I apply the categorization used by Jaimovich and Siu (2013), which I introduce in Section 3.

All in all the literature of labor market polarization focused on two main questions, namely on the causes of the process and on its relationship with the business cycle. From the point of view of my research the latter part of the literature is more important as these works examine the movement not only between the three occupational categories but they include the flows between employment and unemployment as well. The findings of these types of papers are somewhat controversial which justifies my intention to examine the movements of the individuals between the three occupational categories and the unemployment. The novelty of my research is that I use SIPP data to estimate transition probabilities between the occupational groups and the unemployment and that I focus on the effect of the aggregate level variables on the individual movements. As the SIPP collects data about its sample members over a longer period of time (usually more than 2 years), it contains more information about the employment history of the individual sample members. Therefore in the SIPP data the effects of the polarization on the individual level can be more closely examined. Further details about the SIPP are included in the following section.

3 Data description

It is clear from Section 2 that labor market polarization is a topic which has received a lot of attention in the past couple of decades. As I pointed it out previously, most of the above introduced literature used the CPS data in their analyses, which collects information about the employment characteristics of a very large sample of the population over half a century now. My research deviates from this practice because I use the SIPP as data source. The SIPP and the CPS have many common points, however, they differ in two important features. Firstly, while the CPS is address based, the observation units of the SIPP are the households regardless of their locations. Secondly, the SIPP is structured into panels which contain information about a set of households over a longer period of time, while the CPS has also panel dimension, it is more suitable for the analysis of cross-sectional and short-run movements in the aggregates. Because of these properties the SIPP is more suitable for my research, as I focus on the effects of polarization on the changes in the employment characteristics of individuals. The panel structure of the SIPP and the fact that it follows the households throughout its sample period makes it possible to concentrate more on the individual workers while the CPS is more appropriate for studies exploring the effects of polarization on the level of aggregates. In this section first I will introduce the SIPP by comparing it to the CPS then I will describe more precisely the parts of this data source which I included in the analysis. Finally I will show how the polarization can be captured in this dataset.

3.1 The Survey of Income and Program Participation

The SIPP was originally designed as a longitudinal survey that could augment the CPS in fields on which the latter did not lay much emphasis (Westat, 2001). These areas are the income and program participation properties of households with special focus on the asset holdings. The data provided by the SIPP is therefore particularly useful for policy analysis, which was based beforehand on the appropriate parts of the CPS. Apart from these specifics the SIPP also covers topics which are parts of the CPS as well, such as employment properties of the sample members. From the point of view of the welfare and policy analysis these data are also important as the eligibility for certain social programs depends on the employment conditions of the applicants (e.g. unemployment benefits).

The method of data collection applied in the SIPP differs from the one used for the CPS. As the SIPP is a longitudinal survey it is organized into panels, the first of which was executed in 1984. The panels are broken up into waves, where one wave covers a four month period. Originally each panel was planned to be eight waves long, however, not all of them were carried out completely. There was also a restructuring in 1996 after which the number of waves was officially increased to twelve, so that each panel supposed to cover four years instead of the 32 months period before the redesign. One result of the restructuring was that the overlapping structure of the panels was terminated; this was supposed to be compensated with larger sample size, however, it also meant that in case a panel was not fully carried out, there appeared gaps between the panels what for example makes data aggregation more difficult.[4] A result of the panel structure is that the sample members were surveyed over a longer period of time, which results in more data on the individual level. This feature of the SIPP is one of its clear advantages over the CPS, where households are included in the sample for two four-month periods and are dropped from it in between for eight months. Moreover the SIPP is collected on a four month basis which can hopefully mean fewer inconsistencies in the data which can occur if the questions are asked respectively about a longer period of time.

Similarly to the CPS the households are the basic units of the survey in the SIPP, which collects data both about the household as a whole and about each member of it. However, unlike the CPS the SIPP follows movers, which ensures that the data about the households is consistent. What more, if some sample members move out from the original address, because, for example they start a new family, then these individuals are not excluded from the survey. This approach ensures that the core of the sample remains mostly intact over the whole sample period. The survey also takes up those who enter a household in the sample during the period of the survey. This process increases the sample size and it seems to be a natural step because the presence of new members obviously affects the income and other properties of the households. For the same reason if any of these additional sample members leaves the household he or she is dropped from the survey. Hence there are several incomplete data series in the whole sample.

[4] Fujita et al. (2007) for example use standard maximum likelihood interpolation to fill in the gaps.

The panel structure of the SIPP allows for having information about each sample member over a longer period of time than it is possible with the CPS. This feature of the survey is useful for my research, because the longer each person is observed the more chance there is that the effects of polarization will be observable on the individual's employment history.

3.2 The 1991 and 2001 SIPP panels

After this general introduction of the SIPP I discuss the two panels which are used in the following analysis in more detail. I choose the 1991 and 2001 panels because both of them were executed in the periods that are usually considered to exhibit polarization, but as there is a decade difference between them, they depict different stages of the process.

The 1991 panel was one such panel before the 1996 redesign which was completed as planned and so it consists of eight waves. The 2001 panel on the other hand has only nine waves out of the originally intended twelve. Even so the 2001 panel has more observations than the other one, because the number of households in the sample was increased for every survey after 1996. In both panels the sample was divided into four rotation groups and in each month of the survey one of these groups was interviewed about its previous four month period as it is a general practice of the SIPP. This means that for the first and last three months the sample size is smaller than in the other periods of the survey. Similarly to Fujita et al. (2007) I left these six months out of the analysis. As a result I use data of the 1991 panel between January of 1991 and May of 1993, while the relevant period for the 2001 panel spans from January of 2001 until August of 2003.

Apart from the ones deriving from the redesign in 1996 there are other differences between the two panels which are related to the questions the surveys contained in the two periods. Most variables are essentially the same in both panels, however, in some cases the content of the questions changed between 1991 and 2001. There are also variables which are present in one panel but not in the other. In such cases the best way to solve the problem is to look for variables with similar meaning to the missing one and use those in the analysis. This method is the one which I apply when I determine whether somebody was self-employed in 2001 panel. While in the 1991 survey there is a variable coding self-employment, in the 2001 panel one can only indirectly find out who belongs to this category by using a variable which contains information about the number of businesses a sample member held in the given month. My assumption is here that if somebody owns a business, than he or she is self-employed.

A more severe problem is that in the 1991 panel in some cases the same value denoted if the question remained unanswered ("Not in the universe.") and when the value of the variable is actually zero. Such variable was for example the usual weekly hours worked. To be consistent with the notation between the two panels I recode such variables in the 2001 panel such that the "not in the universe" case takes up zero as well. This exercise is mostly harmless from the point of view of the data, because those for whom the weekly hours worked is zero were generally people who were not employed.[5]

Another difference was the coding of the occupations between the two panels. The occupational codes applied in the SIPP correspond to the ones defined by the IPUMS.[6] However, this coding system also changed over time. While the 1991 panel uses the 1980 IPUMS occupational codes the second panel is coded with the 1990 list.

From the point of view of this research those variables of the SIPP are important which contained information about the employment and personal characteristics of the sample members, therefore variables recording the various program participation properties of the sample members were not used in the analysis. I also restrict my sample to people between the ages of 16 and 64, as it is a general practice in this line of literature to focus on the working age population. This restriction is also necessary because people older than 64 are tendentiously moving out of employment (and the labor force as well) and children younger than 16 are also mostly not part of the labor force, therefore their presence causes a bias or has no effect at all when it comes to examining the properties of job changes in the sample. One could argue for the exclusion of women from the sample as well, as their labor market attachment is smaller, however, however as the literature I introduced beforehand does not follow this approach, I include both females and males in the analysis.

Similarly to the general approach followed by the literature, self-employed individuals are also not included in the analysis. I use the Core Wave files of the SIPP to create dataset for the analysis, each of which files contains the data of one wave. This makes it possible to look for self-employed people in every wave separately and so if somebody turns out to be self-employed in some of the waves then he or she is dropped from the sample in those waves only. On the one hand this process leaves me with more observations than if all observation of the self-employed were deleted; on the other hand, however, it also increases the number of incomplete individual datasets.

[5]In the end I recode with this method two variables only: the one which stands for the occupations and one which measured the number of weekly hours worked.

[6]Integrated Public Use Microdata Series, 1980 and 1990 Occupation Codes

Below I display and briefly discuss some summary statistics of the above described dataset. Tables 3.1 and 3.2 contains the main statistics for the two panels. Looking at these statistics what one can notice at first sight is that the mean values are almost without exception higher in the 2001 panel than in the other one, resulting in higher average ages and real wages[7], while the number of weekly hours worked is slightly smaller than in the 1991 sample. Wages show an increase for both genders, even though the labor force participation for males is lower and for females higher in the 2001 panel. Educational attainment is not contained in the table as it is measured differently in the two panels. 53% of the 1991 panel's population has finished maximum high school while this number drops to 47% in the 2001 panel, where the categorization of educational attainment is much finer. All in all people seem to have higher educational attainment and higher real wages in the second panel than in the 1991 dataset.

Table 3.1: Summary statistics for the 1991 and 2001 panels, continuous variables

	1991		2001	
	Mean	*St.Dev.*	Mean	*St.Dev.*
Age	36.9	12.61	38.18	12.87
Monthly real earnings	976.7	1098.29	1081.5	1481.89
… for males	1303	1264.02	1402	1784.52
… for females	692.8	832.22	802	1079.71
Weekly hours worked	28.85	19.98	28.71	19.44

Source: SIPP, author's own calculation

Table 3.2: Summary statistics for the 1991 and 2001 panels, binary variables

	1991	2001
Number of observations	561 032	1 253 444
Ratio of males	47%	47%
Labor force participation	77%	77%
…for males	86%	83%
… for females	70%	71%

Source: SIPP, author's own calculation

[7]The SIPP collects data about the nominal wages only, therefore to calculate the real wages I deflate the nominal wages with the monthly CPI data available in the FRED database.

3.3 Polarization in the two panels

The occupational codes in the SIPP panels range between 0 and 909; however, there are values to which no occupation is ordered. Still, there are around 500 categories on the list for each panel. As polarization is defined on the level of larger skill groups, some level of aggregation has to be applied on the SIPP occupations before one can search for the signs of polarization in the panels. The aggregation is also necessary because without it in each category there would be only a few observations which leads to noisy data.

There are several ways one can divide up the jobs between the low-skill, middle-skill and high-skill categories, some of which I have introduced in Section 2. The classification I used is the one described by Jaimovich and Siu (2013) on the example of the data collected by the Bureau of Labor Statistics (BLS) available in the FRED database, which is their data source for the period between 1983 and 2011. The employment data in question is collected on a monthly basis from a large sample of firms in the U.S. and it already contains some level of aggregation, however, its categories mostly correspond to the larger occupational groups defined on the IPUMS code list. Table 3.3 shows how the occupational groups in the SIPP were matched to the ones defined by Jaimovich and Siu (2013).[8]

I compare the two categorizations in order to show how well the aggregation applied in this research matches with one used in the polarization literature. This step is necessary as it demonstrates that the data in question does not differ much from the one provided by the BLS and therefore it is justified to use it in the analysis of the polarization.

Based on the above classification one can compare the summary statistics of the three main occupational categories between the SIPP and the FRED data. I consider here the employment-to-population ratio which is defined as the number of employed people divided by the whole sample size in the given period, and the employment share of an occupational group which is the size of the employment in the given category divided by the number of all employed sample members in the same period. In each case longitudinal panel weights were used in the calculations which supposed to compensate for the over and under-sampling of the various subpopulations (Westat, 2001:8-2). I calculated the same rates for each month of the two sample periods for the SIPP samples and for the FRED data as well. Because individuals who are younger than 16 or older than 64 are excluded from the SIPP sample, therefore I use the Working Age Population data from the FRED database from which I subtract self-employment before calculating the employment-to-population ratios. I subtract the amount of self-

[8] Farming and related occupations are also excluded from the analysis.

Table 3.3: Categorization of occupations - comparison of the categories used in Jaimovich and Siu (2013) and the corresponding SIPP categories

		Jaimovich and Siu categorization	SIPP categorization
Non-routine Cognitive		• Management, Business and Financial Operations • Professional and Related	• Managerial and Professional Specialty
Non-routine Manual		• Service	• Service
Routine	Cognitive	• Sales and Related • Office and Administrative Support	• Technical, Sales and Administrative Support
	Manual	• Production	• Precision, Production, Craft and Repair
		• Transport and Material Moving • Construction and Extraction	• Operators, Fabricators and Laborers

Source: Jaimovich and Siu (2013), p. 38, and author's own categorization

employed as my data set did not contain self-employment either. Tables 3.4 and 3.5 show the summary statistics for the two panels and the FRED data in the corresponding periods.

Table 3.4: Employment-to-population ratios - summary statistics

		Non-routine cognitive		Non-routine manual		Routine		Aggregate	
Panel		SIPP	FRED	SIPP	FRED	SIPP	FRED	SIPP	FRED
1991	Mean	0.186	0.242	0.097	0.122	0.415	0.406	0.698	0.770
	St. Dev.	0.002	0.002	0.003	0.003	0.004	0.006	0.006	0.009
2001	Mean	0.226	0.273	0.109	0.125	0.406	0.386	0.741	0.783
	St. Dev.	0.003	0.003	0.002	0.003	0.006	0.008	0.009	0.009

Source: SIPP, FRED Database; author's own calculations
Employment-to-population ratio: ratio of employment and working age population (ages 16-64)
Calculations for SIPP are including panel weights

At first sight one can observe that the employment-to-population ratios are closer to the FRED rates in the 2001 panel than in 1991 one. It is also apparent that the ratios for the routine group are always higher than the ones in the FRED data, while those of the non-routine categories are lower. This is especially obvious in the case of the employment shares. Table 3.5 shows that while in the whole economy the share of routine employment of the total is around 53% in the 1991-1993 period and 49% in the early 2000s, in the restricted SIPP samples these rates are closer to 60% and 55% respectively. This means that in the SIPP the ratio of routine workers in the working age population and in the total employment is higher than what the FRED data shows in both periods. As the rates for the SIPP are calculated with the inclusion of sample weights and the FRED data derives from the Current Employment Statistics (CES) survey executed by the BLS, the question arises that which ratio should be accepted as true for the economy. The CES is collected on a monthly basis and it is widely used (Hatch-Maxfield and Robertson, 2012), therefore I accept the ratios calculated from the BLS data as true and assume that the ratios calculated for the SIPP are hinting that the subsample I intend to use over-represents routine workers with higher panel weights.

Table 3.5: Employment shares - summary statistics

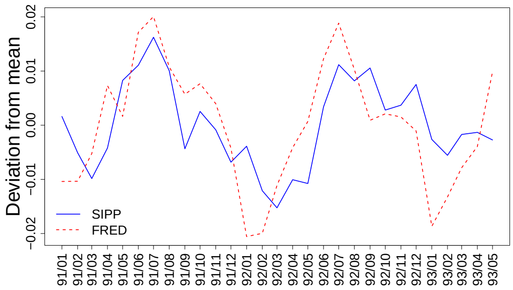
		Non-routine cognitive		Non-routine manual		Routine	
Panel		SIPP	FRED	SIPP	FRED	SIPP	FRED
1991	Mean	0.267	0.315	0.139	0.159	0.595	0.527
	St. Dev.	0.003	0.004	0.003	0.002	0.002	0.003
2001	Mean	0.305	0.348	0.147	0.160	0.547	0.492
	St. Dev.	0.001	0.004	0.003	0.005	0.002	0.005

Source: SIPP, FRED Database; author's own calculations
Employment share: ratio of employment in given category and total employment
Calculations for SIPP are including panel weights

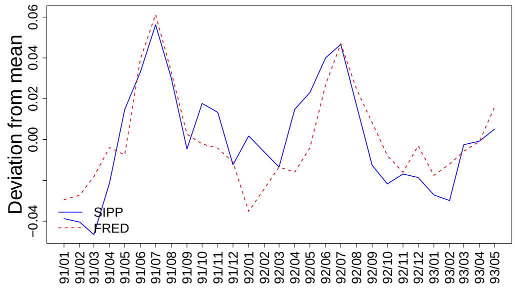
Apart from these observations, looking at the changes in the mean ratios one can see that both the employment-to-population ratio and the employment share of the routine occupations are smaller in the period of the second panel whereas the ratios for the non-routine categories show an increase. The same tendency holds for the FRED data. This behavior of the means can be considered as the sign of the long run polarization. The more interesting question is whether the polarization can be detected *within* the panels.

As polarization means that fewer and fewer people work in the routine sector while employment is increasing in the non-routine cognitive and manual occupations, therefore it should appear in the change of the above described ratios over the sample period. Thus plotting these ratios should show the presence of polarization in the samples if there is any. Figure 3.1 plots the deviations from the mean employment-to-population ratios normalized with their mean for the 1991 SIPP panel and for the BLS data in the same period. The graphs show strong seasonality for the ratios calculated from both datasets and a close co-movement is observable between the SIPP and the BLS ratios on the aggregate level. The co-movement for the non-routine manual and the routine occupational categories is also strong, however, the same cannot be said about the non-routine cognitive category. The ratios calculated for the SIPP data do not show a polarizing pattern, but as no polarization is clearly observable in the ratios for the BLS data, this does is not out of the ordinary. The lack of obvious polarization is not so surprising here because the process was less pronounced in the early 1990s.

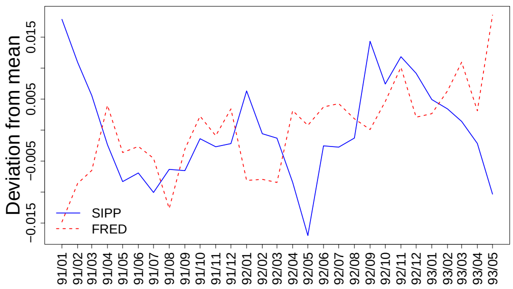
Figure 3.1: Employment-to-population ratios, 1991 panel



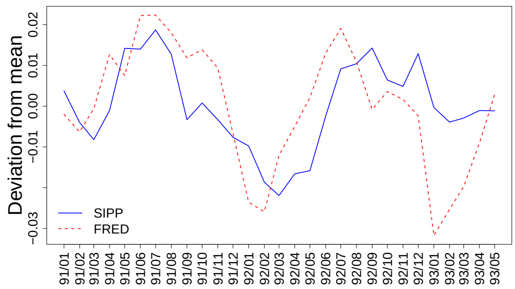
(a) Aggregate



(b) Non-routine manual



(c) Non-routine cognitive

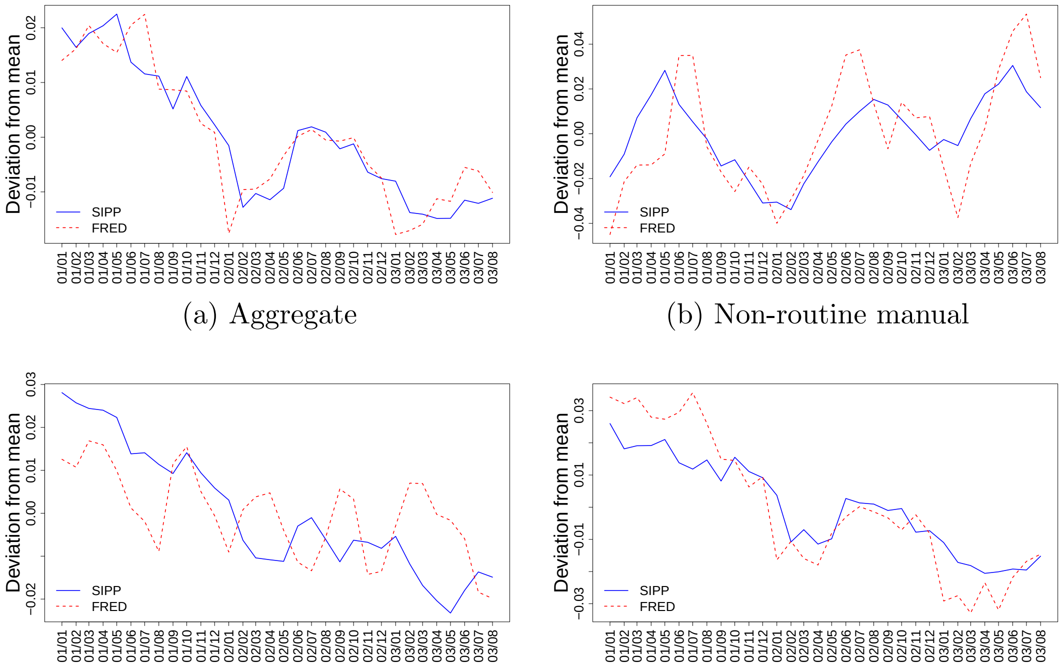


(d) Routine

The employment-to-population ratio is calculated as the number of observations in given occupational category divided by the total sample population. Longitudinal panel weights are used in the calculation. The graphs show the deviation from the mean, normalized with the mean for each occupational category and for the aggregate employment. The sample population includes those who are between the ages of 16 and 64, are neither self-employed nor working in farming occupations. For the FRED data the population is defined as the working age population excluding self-employed individuals.

Figure 3.2 shows the same ratios for the 2001 panel. In the period between January of 2001 and August of 2003 the polarization is clearly observable from the graphs. The aggregate employment-to-population ratio is declining almost throughout the whole era and the ratios for the non-routine cognitive and routine occupations follow this pattern. Only the non-routine manual occupations show a more or less solid growth, especially after the decline between May of 2001 and

Figure 3.2: Employment-to-population ratios, 2001 panel



(a) Aggregate (b) Non-routine manual

(c) Non-routine cognitive (d) Routine

The employment-to-population ratio is calculated as the number of observations in given occupational category divided by the total sample population. Longitudinal panel weights are used in the calculation. The graphs show the deviation from the mean, normalized with the mean for each occupational category and for the aggregate employment. The sample population includes those who are between the ages of 16 and 64, are neither self-employed nor working in farming occupations. For the FRED data the population is defined as the working age population excluding self-employed individuals.

March of 2002. This pattern is well-discussed in the literature, as the recession which had its through in November of 2001 is followed by a jobless recovery and also this was such an economic down turn in which the job loss of non-routine cognitive workers was more severe than in the previous recessions (Foote and Ryan, 2012). It is clear from the graphs that even though the recession reached its deepest point in November of 2001, the employment-to-population ratios continued to drop afterwards. This decline turns back only in the case of the non-routine manual occupations within the sample period, but that in itself does not turn around the negative trend in the aggregate ratio. Notably here the polarization “benefits” the non-routine manual workers, while the non-routine cognitive ones suffer just as much in the aftermath of the recession as the routine workers. In the case of Figure 3.2d) the co-movement between the SIPP and BLS ratios is observable for the routine category and for the aggregate ratios, while in the non-routine categories the SIPP ratios show smaller volatility than same ratios calculated from the BLS data.

Based on the two sets of figures and Tables 3.4 and 3.5 it is evident that the two periods differ from each other in terms of polarization. In the period of the 1991 panel polarization is not visible on the graphs, whereas one decade later it is observable. This implies that using these two panels to examine the influence of the polarization on the individual labor market decisions makes sense, as it is likely to find a difference in the effects of the aggregates between the two datasets.

4 Methodology

The graphs in Section 3 suggest that polarization on the aggregate is observable in the SIPP panels just as much as in the FRED data. However, my research question focuses on whether its presence and effects can be shown with econometric methods on the individual level; also as I stated it earlier, my main goal is to examine how the polarization affects the labor market decisions on the individual level, concretely the occupational changes, and how large is the feedback from the changes in the labor market aggregates (such as the employment-to-population ratios and employment shares) on the employment decisions of the sample members.

In the focus of my research are the movements between unemployment and the three occupational categories, the non-routine cognitive, the non-routine manual and the routine group. Unlike Foote and Ryan (2012) I do not try to deeply investigate the cyclical aspect of these flows, especially because the two SIPP panels I intend to use are not connected in time. By using two panels my goal is to make comparisons between them. As polarization is a structural phenomenon which is present over decades in the data, it is a natural way to gain more information about it by looking at its effects in two distinct time periods and search for similarities and possible differences. Therefore I explore how the transition probabilities between the above mentioned four categories look like in the 1991 and 2001 SIPP panels and examine whether the effect of polarization can be detected in these probabilities. To estimate these probabilities I use logistic regressions. In this section I shortly introduce the logit models I apply, focusing first on the general characteristics of these models. Then I describe the dependent and explanatory variables of the models I test on the data. Finally, I discuss some important properties of the data and the last transformations I administer to it before running the regressions.

4.1 The logistic regression – a short summary

A natural way to approach the estimation of transition probabilities is to use logistic regressions. In a logit model the dependent variable is binary and the aim of the regression is to estimate the probability with which it takes up the two possible outcomes it can have. More precisely the logistic regression helps to determine that with what probability the dependent variable takes up the value 1. The logit model has the general form described in Equation (1):

$$\mathbb{P}(y_i = 1) = \frac{\exp(X_i\beta)}{1 + \exp(X_i\beta)} \tag{1}$$

The functional form ensures that the value on the right hand side stays between 0 and 1; therefore it can be interpreted as a probability.

As in the case of a logistic regression the estimated model is not linear, the interpretation of the coefficients is not as straightforward as in, for example OLS regressions. Plugging in the linear predictors and their coefficients in Equation (1) results in a curve, on which one unit change in any independent variable will not have constant effect over the range. Naturally the most interesting is the maximum impact that can be induced by varying the explanatory variables, which can be reached where the slope of the logistic curve is the highest. As Gelman and Hill (2007) point out this is at the midpoint of the curve, where the slope of the function equals $\beta/4$. Therefore one can get the maximum ceteris paribus effect of one unit change in any independent variable on the outcome probability by dividing its coefficient by 4 (Gelman and Hill, 2007:82). Applying this approach is very convenient because it allows for interpreting the coefficients straight on the probabilistic scale, without further transformations.

The goodness of fit for the logistic regression is measured also somewhat differently from the method applied in the case of linear models. Rather than calculating R^2 the usual approach is to compare the null deviance and the residual deviance of the model. The null deviance is calculated for a model where the only explanatory variable is a constant, while the residual deviance is gained from the actual model which was estimated. If the estimated model has a better fit than the null model, then its deviance must be smaller as well (see Gelman and Hill, 2007:100).

4.2 The variables

After introducing the main econometric method I want to use to examine the effects of polarization on the individual employment decisions it is also necessary to talk about the various independent and dependent variables which I plan to include in the model.

There are four transition probabilities which I find important from the point of view of my paper. The first is the one for the movement from employment to unemployment, the others are the probability of acquiring a routine, a non-routine cognitive or a non-routine manual job after job change. My assumption is that as polarization became more apparent in the early 2000s than in the the beginning of the 1990s, therefore there could be a difference in these probabilities between the two panels. I expect to find a decline in the transition probabilities for routine workers towards routine jobs after job change, and an increase in their chances to move towards non-routine occupations based on Tables 3.4 and 3.5. For the probability of moving towards unemployment after change I expect that these probabilities should be higher for routine and non-routine workers in the 2001 panel, as that would be consistent with the Figure 3.2.

As I am interested in determining how the nature of the transition between the unemployment and the three main occupational categories is affected by the polarization in the panels, therefore I estimate transition probabilities of those individuals, who actually changed job during the sample period. In order to do so I exclude from the examination the sample members who did not have a job or worked for the same employer in the same occupation over the whole period. This restriction does not affect the form of Equation (1); however, it means that I estimate the probability of moving to unemployment or to an occupational category given that job change happened for the person. Similarly, when I estimate the probability of becoming unemployed after job change, I restrict the sample to those who had employment before the change. I do this because I am interested in the probability of not finding a job upon job change for those who have this possibility; and those, who do not have a job before the change always have one afterwards. This is true in the whole sample of job changers, as I originally excluded those observations from the sample where job change was indicated, but neither the employer, nor the occupation showed actual change.

After introducing the dependent variables now I discuss which variables are included in the logistic regressions and why. The model inputs can be divided into two main groups based on whether they are individual or aggregate variables. The first category incorporates personal characteristics which are generally considered important in determining a person's employment status. These are the individual's gender, race, age and educational level. Information about schooling is especially important, as it directly affects the occupational choice (e.g. through qualification requirements of certain jobs). However, not only these individual features influence the movement between unemployment and the different occupational categories; a person's previous occupation and his or her employment parameters before the job change can also be relevant in this regard. That is why I add such variables to the regressions which hold information about the previous job of the sample members.

I include a variable reflecting the union coverage of the sample members as being a union member or being covered by a union contract makes the movement towards unemployment less likely. Also, unions are more common in routine occupations (e.g. in the manufacturing sector), and deunionization was considered at least partly responsible for the job polarization (Autor, 2010). Therefore I expected that being a union member will have a negative effect on the probability of becoming unemployed upon job change.

I anticipate similar results from including variables measuring the weekly hours worked and the real hourly wage earned in the job before the change. The argument here is that someone that is high on the payroll is more important to his or her employers, therefore this person is less likely to lose his or her job because of firing and even in case of job change such people tend to move to jobs with similar pay, not to ones with lower salary. The wage at the previous workplace therefore can be considered as a crude proxy of the broad occupational category. The amount of weekly hours worked could be important in the regressions because people who work more are probably are more attached to their job and to the labor market in general, and thus they are also more likely to find a new job.

I also add two variables which contain information about the job change itself. One of them is a binary variable for being laid off on the previous workplace, the other one is a variable measuring whether there was an employer change. I assume that those who were laid off were more likely to be unemployed after the change, and that the employer change can increase the possibility of the occupational change. The summary statistics of these variables are included in Tables 4.1 and 4.2.

Apart from the variables described above I also include dummy variables for the occupational category to which the individual belonged to before the change; to explain the effect of polarization on the members of different skill groups these are between the most important variables.

In the specifications where these binary variables are present I also incorporate the aggregate variables. The aggregate variables are the changes in the employment-to-population ratios and employment shares calculated for every person who had a job change in the period when this alteration actually happened. The change in the ratios is defined as $dX_i = X_{i,t-1} - X_{i,t}$, where $t-1$ is the period

Table 4.1: Summary statistics for the 1991 and 2001 job changers datasets, continuous variables

Variables		1991 Mean	1991 St.dev.	2001 Mean	2001 St. dev.
Age		30.46	11.550	33.37	12.640
Real hourly wage before change		5.022	7.425	5.155	10.710
Hours before change		29.950	17.533	28.72	17.992
1000 × Change in employment share	Routine	-0.063	1.561	-0.188	1.422
	Non-routine cognitive	0.178	2.558	-0.170	1.106
	Non-routine manual	-0.114	2.659	0.358	1.870
1000 × Change in employment-to-population ratio	Routine	-0.135	3.536	-0.897	3.283
	Non-routine cognitive	0.084	1.547	-0.550	1.744
	Non-routine manual	-0.101	2.225	0.063	1.436
	Aggregate employment	-0.153	5.160	-1.384	5.129
Age when finished school		19.94	3.379		
Occupational experience (in months)				11.81	50.083

Source: SIPP; author's own calculations
Employment-to-population ratio: ratio of employment and working age population (ages 16-64)
Employment share: ratio of employment in given category and total employment
Calculations for SIPP are including panel weights

Table 4.2: Summary statistics for the 1991 and 2001 job changers datasets, binary variables

Share of ... in the sample:	1991	2001
Males	0.489	0.458
Whites	0.849	0.795
Labor force participants	0.901	0.870
High school drop outs	0.154	0.173
Union members	0.067	0.067
Laid off workers	0.158	0.043

Source: SIPP; author's own calculations

before the change and t is the period after the change and i denotes the individual who changed job in this period. Considering that my goal is to determine how do the changes in the aggregates affect the movement of the individuals, these variables are the most relevant ones in the analysis. Table 4.1 shows that the mean of these changes is very close to zero and that the standard deviation is very small. This results from the fact that mostly the job or occupational change happens from one month to the next, therefore the changes in the ratios are also small. My expectation was that the probability of being unemployed after job change will depend negatively on the changes in the aggregates, meaning that in periods when the employment-to-population ratios were increasing, individuals will be less likely to change to unemployment. I also wanted to see whether there is a difference in the effect of the ratios on the movement between the different skill groups.

4.3 More information about the data and a final transformation

Before implementing the regression models I narrow down the dataset to those individuals, who faced some sort of job change over the sample period as described in the previous subsection, furthermore I shorten the list of variables to those only, which contained important information about the change. By doing so I gain two smaller samples from the two panel files. To determine whether somebody had a job change in the 1991 panel I use the variable "WS1CHG", which takes up the value 1 in case of job change in the given month. Job change here covers both the alteration in occupation and in the employer. Unfortunately, in the 2001 panel people were asked only about employer changes; therefore I had to construct the job change variable for this panel myself. However, in the 2001 panel in the first wave people were asked about their work experience in the occupation they had in the survey month. From this information I could reconstruct the occupational experience over the sample period for the participants of the first wave. In the 1991 panel it is not possible to create the same variable as in this panel no question was asked about the occupational history of the sample members before the period of the survey. Therefore the regressions on the 1991 data do not contain a variable controlling for occupational experience.

Another significant difference between the two panels is the way they collected information about educational levels. In the 1991 panel it is possible to create a continuous variable which tells that at what age a person finished schooling (see Table 4.1), however, there is no data about the type of education the individual had. In the 2001 panel the educational variable contains information about the type of schooling but it is not possible to construct a continuous variable from that like for the other panel; instead here I use a multilevel categorical variable as a measure of schooling, which is summarized in Table 4.3.

Table 4.3: Summary statistics of the educational level of job changers in the 2001 panel

Ratio of people in the 2001 panel with ...	
Not even high school diploma	0.173
High school diploma only	0.310
With less than 3 years college education	0.330
With Bachelors degree	0.136
With more than 3 years of college education	0.051

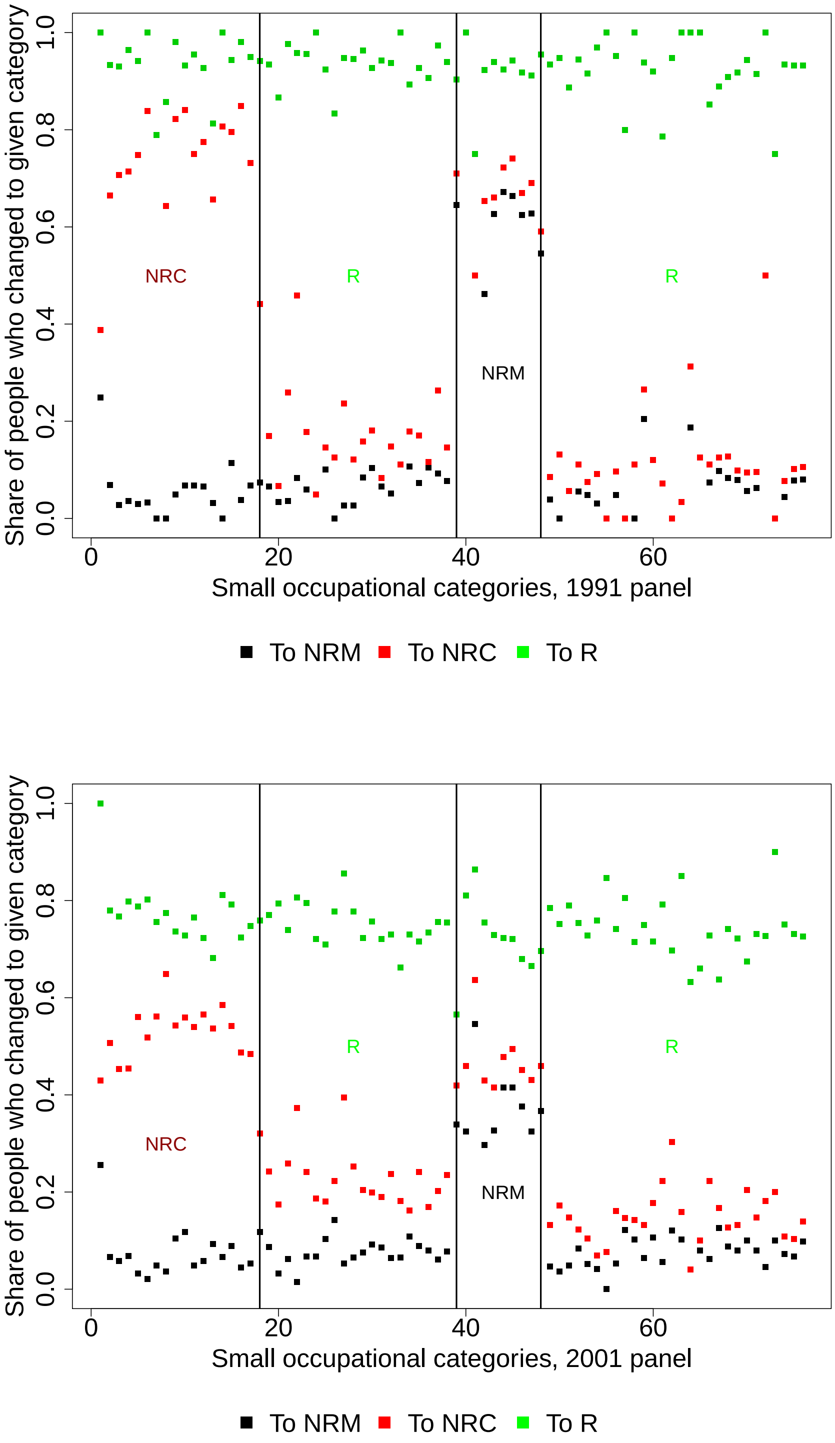
Source: SIPP; author's own calculations

The subset of job changers includes information about individual characteristics which are independent of the occupation such as gender, age or race but mainly it focuses on the parameters of the profession change. Based on their occupation I categorize how people moved between the three occupational categories and the unemployment. Apart from the individual parameters the dataset also incorporates the changes in the employment-to-population ratio and employment share of the main occupational groups in the period of the job change of each sample member.

Comparing the two datasets based on the summaries in Tables 4.1, 4.2 and 4.3, one can conclude that there is no large difference between the two periods in terms of personal characteristics. The average age is higher in the 2001 sample and the variation in the real hourly wages also increased compared to the 1991 period along with the share of high school drop outs, while in the 2001 sample significantly fewer job changes happened because of layoffs than in the 1991-93 period. It is also important to notice that as the 2001 survey had more participants, the number of job changers in this panel is almost five times larger than that of the 1991 sample.

Figure 4.1 plots the movements of job changers between the occupations. Here instead of plotting the movements between all the occupations in the sample, I aggregate the jobs into 76 smaller groups based on the smaller categories defined in the IPUMS code list. In each category the black marker denotes the share of those job changers, who moved to a non-routine manual occupation, similarly the red and green markers denote the share of those job changers in a given category, who got a routine or non-routine cognitive type of job respectively after the job change. The graphs are constructed in such a way, that the in each category the left-over share belongs to those, who turned out to by unemployed upon job change. The vertical lines denote the limit between the main occupational categories; between 0 and the first such line are the non-routine cognitive occupations, between the second and the third are the non-routine manual ones, and the remaining categories belong to the routine group. The Figure shows, that in all three main occupational categories people tend to change to the same category in the largest amount. It is also clear from the graphs, that in the second panel the share of those who moved to unemployment is higher in each category than in the 1991 panel and that this is true for all three skill groups.

Figure 4.1: Movements of job changers in the 1991 panel



On the horizontal axis there are the small occupational categories defined in the IPUMS code list. For each of these categories it is calculated that what share of those who were employed here moved to a routine (R), non-routine manual (NRM)or non-routine cognitive (NRC) occupation. These ratios are added up, and the difference between them and 1 gives the share of those who moved to unemployment in the given category. The vertical lines denote the limit between the skill groups. Categories 1-18 are non-routine cognitive, categories 40-48 are non-routine manual, the others are routine occupations. Category 0 denotes those, who were unemployed before the change.

A final transformation of the data was necessary before running the regressions, namely the standardization of the continuous variables. The method I applied was the one recommended in Gelman (2008) and it consists of dividing the mean-centered variables with two standard deviations. After this transformation the regression coefficients of the rescaled variables are more conveniently comparable with the coefficients of the binary variables than in the case when no rescaling is used. The interpretation of the coefficients in the rescaled model is then basically the comparison between values of the input variable at the mean minus one and at the mean plus one standard deviation (Gelman and Hill, 2007:56).

5 Empirical results

On both datasets I run several model specifications. In the simplest regression model I control only for personal characteristics then I expand it with variables which carry information about the general employment properties of the individual and about the changes in the employment shares and employment-to-population ratios in the period of the job change. In the following first I will examine the outcomes of those regressions where the dependent variable is the probability of being unemployed after the job change then I analyze the regressions where the explanatory variable is the probability of movement to one of the main occupational categories. To make the analysis more transparent, I describe the results of the two panels separately and then compare them.

5.1 Regression results in the 1991 panel

Specification 1) in Table 5.1 shows the results of the baseline regression for the job changers who were employed before the change in 1991. Interpreting the coefficients with the rule defined in Gelman and Hill (2007) one can say that being a white male decreases the probability of ending up without a job after change with 4.9 and 2.5% respectively, however, the coefficient of the gender variable is not significant. Not surprisingly higher levels of education significantly decrease the probability of job loss. However, interpreting the coefficient of education is not that straightforward, as the variable carries information about only the age at which the person finished schooling. The results also confirm that people who quit school before their 18^{th} birthday and therefore presumably did not finish high-school are 8% less likely to find a new job straight after the change. Similarly, the result shows that people whose age is the mean age plus one standard deviation (so approximately 41) are 3% less likely to have a job after the change than people whose age is under the mean age with one standard deviation.

Expanding the model with employment characteristics and labor market variables adds further details to the picture. First I include the aggregate employment-to-population ratio (Specification 2), then I exchange it with the employment-to-population ratios of the three occupational categories (Specifications 3 to 5). The results of these four specifications were similar to each other in the coefficients of the variables present in each of them. The probability of not having a job after a job change is higher with around 8.6% when employer change happened and this effect is significant on a 0.1% level. Based on the results people who were laid off and who were union members in their previous workplace are more likely to find a job after the change, however, only the effect of the union membership is significant here and that is on a 10% level. The regression results show that the more hours per week people worked on their previous work place, the less likely they were to be unemployed after the change. This effect is significant and causes a 13.6% difference in the job loss probability between those who worked one standard deviation less than the mean compared to those who worked one standard deviation more than the mean in their job before the change. The wage before the change also has a negative effect on the probability of not having a job after the change, but it is small and not significant.

The coefficients of the changes in the aggregate variables show a very interesting picture. The coefficients for the aggregate and the routine employment-to-population ratios are not significant, while the ratio of non-routine cognitive and manual occupations in the working age population have a significant effect with opposing signs. The sign of the non-routine manual rate turned out as expected, showing that an increase in the employment rate of this category decreases the probability of being unemployed after the change with almost 5%. The positive coefficient for the change in the non-routine cognitive employment-to-population is, however, implying that in periods when the more non-routine cognitive workers are employed, the probability of unemployment increases with 8%. This result can be explained with the help of Figure 3.1, where one can observe that the employment-to-population ratio of the non-routine cognitive occupational group moves in the opposite direction as the same ratio for the non-routine manual occupations; the correlation between them is -69.5%. The graph and the regression output together imply that in periods when the non-routine cognitive rate is increasing the probability of being unemployed after the change is higher because in these periods the non-routine manual employment-to-population ratio is decreasing and the ratio for the routine category is also lower, therefore the overall chances of finding a job are smaller.

Table 5.1: Regression output, with ℙ(no job after change) as dependent variable, 1991 panel

	1)	2)	3)	4)	5)
Intercept	−2.446***	−2.654***	−2.653***	−2.673***	−2.649***
	(0.103)	(0.124)	(0.124)	(0.124)	(0.124)
White	−0.196	−0.214*	−0.214*	−0.228*	−0.221*
	(0.103)	(0.103)	(0.103)	(0.103)	(0.103)
Male	−0.102	0.012	0.012	0.014	0.009
	(0.076)	(0.079)	(0.079)	(0.079)	(0.079)
Age	0.134	0.229**	0.230**	0.255***	0.248**
	(0.074)	(0.076)	(0.076)	(0.076)	(0.076)
School	−0.205*	−0.205*	−0.205*	−0.208*	−0.210*
	(0.094)	(0.096)	(0.096)	(0.096)	(0.096)
High school drop out	0.344**	0.316**	0.316**	0.309*	0.309*
	(0.121)	(0.122)	(0.122)	(0.122)	(0.122)
Employer change		0.349***	0.349***	0.378***	0.347***
		(0.095)	(0.095)	(0.095)	(0.094)
Laid off		−0.024	−0.024	−0.019	−0.023
		(0.104)	(0.104)	(0.104)	(0.104)
Wage before change		−0.026	−0.026	−0.038	−0.029
		(0.079)	(0.079)	(0.081)	(0.079)
Hours before change		−0.540***	−0.541***	−0.547***	−0.555***
		(0.106)	(0.106)	(0.107)	(0.107)
Union membership		−0.265	−0.265	−0.272	−0.266
		(0.160)	(0.160)	(0.160)	(0.160)
d Aggregate e-to-p		0.007			
		(0.078)			
d Routine e-to-p			−0.005		
			(0.078)		
d NR cognitive e-to-p				0.325***	
				(0.072)	
d NR manual e-to-p					−0.218**
					(0.078)
AIC	5416.9	5383.2	5383.2	5364	5375.4

	6)	7)	8)	9)	10)	11)	12)
Intercept	−2.660***	−2.659***	−2.680***	−2.652***	−2.659***	−2.659***	−2.655***
	(0.160)	(0.160)	(0.160)	(0.160)	(0.159)	(0.160)	(0.160)
White	−0.215*	−0.215*	−0.228*	−0.222*	−0.215*	−0.224*	−0.227*
	(0.103)	(0.103)	(0.103)	(0.103)	(0.103)	(0.103)	(0.103)
Male	0.010	0.010	0.012	0.008	0.010	0.009	0.006
	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)
Age	0.229**	0.229**	0.255***	0.247**	0.229**	0.252**	0.256***
	(0.077)	(0.077)	(0.077)	(0.077)	(0.077)	(0.077)	(0.077)
School	−0.203*	−0.203*	−0.205*	−0.209*	−0.204*	−0.208*	−0.213*
	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)
High school drop out	0.319**	0.319**	0.312*	0.311*	0.318**	0.312*	0.305*
	(0.122)	(0.122)	(0.123)	(0.122)	(0.122)	(0.122)	(0.123)
Employer change	0.350***	0.349***	0.379***	0.348***	0.349***	0.356***	0.354***
	(0.095)	(0.095)	(0.095)	(0.095)	(0.094)	(0.095)	(0.095)
Laid off	−0.026	−0.026	−0.021	−0.024	−0.026	−0.024	−0.022
	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)	(0.104)
Wage before change	−0.026	−0.026	−0.038	−0.029	−0.026	−0.032	−0.034
	(0.080)	(0.080)	(0.082)	(0.080)	(0.080)	(0.081)	(0.081)
Hours before change	−0.544***	−0.544***	−0.550***	−0.559***	−0.544***	−0.560***	−0.562***
	(0.108)	(0.108)	(0.108)	(0.108)	(0.108)	(0.108)	(0.108)
Union membership	−0.267	−0.267	−0.274	−0.268	−0.268	−0.269	−0.273
	(0.160)	(0.160)	(0.160)	(0.160)	(0.160)	(0.160)	(0.160)
d Aggregate e-to-p	0.007						
	(0.078)						
d Routine e-to-p		−0.005					
		(0.078)					
d NR cognitive e-to-p			0.325***				
			(0.072)				
d NR manual e-to-p				−0.218**			
				(0.078)			
d Routine emp share					−0.034		
					(0.078)		
d NR manual emp share						−0.260***	
						(0.076)	
d NR cognitive emp share							0.299***
							(0.078)
Routine occ. before change	0.016	0.015	0.017	0.012	0.016	0.013	0.013
	(0.117)	(0.117)	(0.117)	(0.117)	(0.117)	(0.117)	(0.117)
NR manual occ. before change	−0.009	−0.010	−0.008	−0.015	−0.010	−0.014	−0.016
	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)	(0.138)
AIC	5387.2	5387.2	5367.9	5379.3	5387	5375.7	5372.6

Source: SIPP data
d: change; emp share: employment share, e-to-p: employment-to-population ratio, NR: non-routine
Significance codes: '***' 0.001, '**' 0.01, '*' 0.05
Standard errors are in brackets under the corresponding coefficient.

Specifications 6 to 9 of Table 5.1 contains the same specifications as Specifications 2 to 5 extended with binary variables controlling for the occupational category to which the individuals belonged to before the change. I choose the non-routine cognitive category as a benchmark. The results of the regressions controlling for the occupation before the job change are very similar to those in the corresponding previous regressions and unfortunately the coefficients of the starting employment variables were both very small and not significant in either of the four specifications. Nevertheless, the sign of these coefficients is at least pointing in the right direction implying that those with a routine occupation before the change were less likely to have a job after the change compared to non-routine cognitive workers, while those employed in a non-routine manual occupation were more likely to stay employed. Exchanging the employment-to-population ratios to employment shares in the model led to similar conclusions, the summary of those regressions are in Columns 10 to 12. This again implies that in periods when non-routine manual employment is expanding, the probability of unemployment after job change declines, while the expansion of non-routine cognitive employment has a positive effect on the same probability. The reasoning behind this result is the same as for the case with the employment-to-population ratios, as here as well, an opposing movement is observable between the non-routine cognitive and manual rates (see Figure A.1).

After this exercise I run similar regressions on the whole sample of job changers where the dependent variable is the probability of being employed in a given occupational category out of the three main ones. For each of the three possible outcomes I examine one regression model. In each of them I control for the aggregate employment-to-population ratio change in the period of the job change and I include the dummy variables indicating the occupational group before the change, leaving the group which was unemployed before the change as the basis of comparison. The results of the regressions are summarized in Table 5.2.

Column 1 contains the results of the regression where the dependent variable is the probability of being employed in a routine occupation after the change. Being a white male here has a positive significant effect on the probability in question regardless of the type of the occupation before the change. Being laid off on the previous workplace increases the probability of getting a routine job, which is in accord with the findings of the previous regressions, while employer change has a negative coefficient which is not significant here. Interestingly higher levels of education seem to have a negative effect on the probability of getting a routine job and being a high school dropout seems to have a negative and significant effect as well. It is less surprising that those who were employed in a routine

Table 5.2: Probability of having a job after job change, 1991 panel

	Routine	NR cognitive	NR manual
Intercept	0.541***	−2.024***	−1.001***
	(0.094)	(0.141)	(0.114)
White	0.172**	0.130	−0.234***
	(0.059)	(0.093)	(0.070)
Male	0.294***	−0.159*	−0.405***
	(0.044)	(0.065)	(0.056)
Age	−0.085	0.475***	−0.413***
	(0.045)	(0.065)	(0.059)
School	−0.552***	1.697***	−0.434***
	(0.057)	(0.084)	(0.071)
High school drop out	−0.553***	0.373**	0.086
	(0.075)	(0.143)	(0.088)
Employer change	−0.076	−0.332***	0.169**
	(0.050)	(0.074)	(0.063)
Laid off	0.191**	−0.076	−0.280**
	(0.067)	(0.105)	(0.090)
Wage before change	0.001	0.030	−0.209*
	(0.049)	(0.058)	(0.091)
Hours before change	0.302***	−0.027	−0.113
	(0.067)	(0.097)	(0.085)
Union membership	0.216*	−0.129	−0.136
	(0.097)	(0.142)	(0.141)
d Aggregate e-to-p	−0.012	−0.021	0.045
	(0.043)	(0.063)	(0.054)
Before change:			
Routine	0.571***	−0.819***	−1.216***
	(0.091)	(0.137)	(0.116)
NR manual	−2.028***	−0.686***	1.765***
	(0.096)	(0.150)	(0.108)
NR cognitive	−2.024***	2.310***	−1.265***
	(0.113)	(0.145)	(0.154)
AIC	13265	6981	9232.7

Source: SIPP data
d: change; emp share: employment share, e-to-p: employment-to-population ratio, NR: non-routine
Significance codes: '***' 0.001, '**' 0.01, '*' 0.05
Standard errors are in brackets under the corresponding coefficient.

occupation before the change are 14% more likely to get a similar job after it, while non-routine workers are around 50% less likely to do so compared to the unemployed ones. The coefficient of the change in the aggregate employment-to-population ratio is not significant. Using the routine employment-to-population ratio instead the aggregate one does not change on the results in a significant way.

Columns 2 and 3 of the same table show the results for the cases where the probability of getting a non-routine cognitive and non-routine manual job are the dependent variables. The probability of being employed in a non-routine cognitive occupation is increased by 3% for whites (not significant coefficient), while being a male decreases the chances of employment in this category with almost 4%. In the case of non-routine manual occupations both being white and being male have a negative effect on the probability of finding a job and both effects are significant. There is also a difference in the influence of age and schooling on the two probabilities: in the non-routine cognitive case higher education has a large positive effect while for the non-routine manual case the higher education seems to be a drawback. Higher age (which can reflect higher experience) increases the probability of getting a non-routine cognitive job, while it decreases the probability of non-routine manual employment. Employer change on the other hand decreases the probability of having a non-routine cognitive job by 8% but has a positive effect on finding a non-routine manual job after the job change. In these two specifications the change in the aggregate employment-to-population ratio does not have a significant effect either, and similarly to the first specification, the probability of being employed in a non-routine cognitive or manual occupation is highest for those who had the same type of job before the change, while people employed in other categories previously have significantly smaller chance.

It is also worth to examine the chances of getting any type of job for those, who were unemployed before the change. In order to compare these probabilities across the three specifications in Table 5.2 I compare average predictive differences between these models. The average predictive difference is calculated for each specification separately in such a way that it compares the probability of moving to a the given occupational category upon change between two people who differ only along one characteristics, namely the occupational category they belonged to before the change. Table 5.3 shows that unemployed in the sample have a more favorable position when they change to a routine occupation, as here their probability advantage is the highest compared to those who had an occupation previously which was not routine. Also, the predictive difference is the least negative in the case of the first specification, when it comes to comparing the transition probability of the unemployed with those, who had a job in the same occupational category towards which the change happens. These results imply that for those, who were unemployed before the change, moving towards the routine occupations is easier than towards the non-routine ones.

The results of Table 5.2 are mostly in accordance with my original expectations, apart from the fact they suggest that the changes in the employment-

Table 5.3: Average predictive differences based on Table 5.2

$\mathbb{P}$(R after | U before)-$\mathbb{P}$(R after | NRM before)	45%
$\mathbb{P}$(R after | U before)-$\mathbb{P}$(R after | NRC before)	45%
$\mathbb{P}$(R after | U before)-$\mathbb{P}$(R after | R before)	-11%
$\mathbb{P}$(NRC after | U before)-$\mathbb{P}$(NRC after | NRM before)	6%
$\mathbb{P}$(NRC after | U before)-$\mathbb{P}$(NRC after | R before)	6%
$\mathbb{P}$(NRC after | U before)-$\mathbb{P}$(NRC after | NRC before)	-40%
$\mathbb{P}$(NRM after | U before)-$\mathbb{P}$(NRM after | NRC before)	14%
$\mathbb{P}$(NRM after | U before)-$\mathbb{P}$(NRM after | R before)	14%
$\mathbb{P}$(NRM after | U before)-$\mathbb{P}$(NRM after | NRM before)	-38%

The regression results in Table 5.2 are evaluated on the sample such that the occupational category before the change is fixed for each observation, then the mean of the difference between two predictions is calculated. NRM: non-routine manual occupation, NRC: non-routine cognitive occupation, R: routine occupation, U: unemployed

to-population ratios do not have a significant effect on the flows between the different occupational categories in the 1991 sample. The signs of the coefficients for the binary variables capturing the occupational category before change are in line with the intuition that experience in the same category is more valuable than experience in another occupational group. These results do not show the presence of polarization in the sample, and this is what one can expect based on Figure 3.1.

5.2 Regression results in the 2001 panel

Almost the same regression models are used on the dataset of job changers in the 2001 panel, there are only two significant alterations in the modeling structure. First of all in the 2001 panel *School* is a multilevel categorical variable, which can take up values between 1 and 5, 1 being “Less than high school degree”, 5 being “More than 3-year college degree”. This form of specification makes the usage of the high school dropout variable unnecessary, as School already accounts for this category. The other difference is that the variable for occupational experience also appears in the model, which would cause multicollinearity, if the age and schooling variables were both continuous variables. However, as School is a categorical variable, this problem does not emerge here.

Table 5.4 summarizes the results of the regressions were the dependent variable is the probability of not having a job after the job change. The first Specification contains the coefficients of the baseline model which includes personal characteristics only. All coefficients are significant, and they differ from the ones found in the 1991 specification. Being white and male both have now a diminishing effect of around 6% on the probability of post-change unemployment, while higher levels of education compared to the high school drop outs also decreases this probability by more than 4% with each additional level. Similarly to the 1991 panel the probability of not having a job after the change is increasing in age.

Specifications 2 to 5 contain the results for the regressions which controlled for employment properties and changes in the different employment-to-population ratios as well. Here the coefficients of the variables are again very close to each other across specifications. In this sample, however, employer change has a negative effect on the probability of job loss, while being laid off in the pre-change occupation increases it. The coefficients of these variables in the 1991 sample had the opposite sign, which can imply a change in the behavior of the workers, but this cannot be determined without further analysis. In this sample the level of real hourly wages has a negative effect on the dependent variable as well, namely those workers, whose wages are with one standard deviation higher than the mean real hourly wage are more than 10% less likely to turn out to be unemployed after the job change compared to job changers with wage equal to the mean wage minus one standard deviation. Interestingly higher levels of occupational experience indicate higher probability of being unemployed after job change, which would be natural if the job change meant occupational change as well, as larger experience in one occupation can be a hindrance in finding job in another one. The effect of the change in the aggregate employment-to-population ratio is small and not significant; however, this cannot be said about the employment-to-population ratios of the three main occupational categories. A two standard deviation increase in this rate of the non-routine cognitive occupations seems to increase the probability of unemployment after change with as much as 4%, while an increase of similar size in the ratio for non-routine manual occupations decreases the probability with almost 9%. These results are similar to the ones found in the other panel and the explanation is also similar.

Looking at Figure 3.2 one can observe that again, the non-routine cognitive and manual occupations move in opposite directions most of the sample period. In this period the ratio for the routine occupations also moves together with that of the non-routine cognitive ones, therefore the positive sign of the coefficients of the change in this variables is also not that surprising. Now again, in the periods when the employment-to-population ratio increases for the routine and non-routine cognitive occupations, the ratio for the non-routine manual category

Table 5.4: Regression output, with ℙ(no job after change) as dependent variable, 2001 panel

	1)	2)	3)	4)	5)
Intercept	−0.238***	0.307***	0.311***	0.320***	0.305***
	(−0.035)	(−0.038)	(−0.038)	(−0.038)	(−0.038)
White	−0.241***	−0.236***	−0.237***	−0.240***	−0.237***
	(−0.025)	(−0.027)	(−0.027)	(−0.027)	(−0.027)
Male	−0.256***	−0.185***	−0.185***	−0.185***	−0.186***
	(−0.021)	(−0.023)	(−0.023)	(−0.023)	(−0.023)
Age	0.228***	0.225***	0.226***	0.229***	0.231***
	(−0.021)	(−0.024)	(−0.024)	(−0.024)	(−0.024)
School	−0.187***	−0.198***	−0.198***	−0.198***	−0.197***
	(−0.01)	(−0.011)	(−0.011)	(−0.011)	(−0.011)
Employer change		−1.679***	−1.688***	−1.707***	−1.695***
		(−0.033)	(−0.033)	(−0.033)	(−0.033)
Laid off		0.240***	0.241***	0.240***	0.242***
		(−0.07)	(−0.07)	(−0.07)	(−0.07)
Wage before change		−0.424***	−0.423***	−0.419***	−0.421***
		(−0.039)	(−0.039)	(−0.039)	(−0.039)
Hours before change		−0.609***	−0.608***	−0.604***	−0.606***
		(−0.033)	(−0.033)	(−0.033)	(−0.033)
Union membership		−0.045	−0.044	−0.042	−0.046
		(−0.041)	(−0.041)	(−0.041)	(−0.041)
Occ. Experience		0.142***	0.142***	0.140***	0.142***
		(−0.027)	(−0.027)	(−0.028)	(−0.028)
d Aggregate e-to-p		0.004			
		(−0.022)			
d Routine e-to-p			0.056*		
			(−0.023)		
d NR cognitive e-to-p				0.156***	
				(−0.023)	
d NR manual e-to-p					−0.261***
					(−0.021)
AIC	54332	49910	49904	49863	49756

	6)	7)	8)	9)	10)	11)	12)
Intercept	0.307***	0.310***	0.319***	0.302***	0.306***	0.313***	0.317***
	(−0.052)	(−0.052)	(−0.052)	(−0.052)	(−0.052)	(−0.052)	(−0.052)
White	−0.234***	−0.235***	−0.238***	−0.235***	−0.235***	−0.239***	−0.240***
	(−0.027)	(−0.027)	(−0.027)	(−0.027)	(−0.027)	(−0.027)	(−0.027)
Male	−0.180***	−0.180***	−0.181***	−0.182***	−0.180***	−0.182***	−0.183***
	(−0.023)	(−0.023)	(−0.023)	(−0.023)	(−0.023)	(−0.023)	(−0.023)
Age	0.225***	0.227***	0.230***	0.232***	0.231***	0.237***	0.231***
	(−0.024)	(−0.024)	(−0.024)	(−0.024)	(−0.024)	(−0.024)	(−0.024)
School	−0.197***	−0.197***	−0.197***	−0.196***	−0.196***	−0.195***	−0.197***
	(−0.012)	(−0.012)	(−0.012)	(−0.012)	(−0.012)	(−0.012)	(−0.012)
Employer change	−1.681***	−1.689***	−1.708***	−1.696***	−1.698***	−1.726***	−1.710***
	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)
Laid off	0.246***	0.247***	0.246***	0.247***	0.250***	0.250***	0.241***
	(−0.07)	(−0.07)	(−0.07)	(−0.07)	(−0.07)	(−0.071)	(−0.071)
Wage before change	−0.424***	−0.422***	−0.419***	−0.420***	−0.421***	−0.417***	−0.417***
	(−0.039)	(−0.039)	(−0.039)	(−0.039)	(−0.039)	(−0.039)	(−0.039)
Hours before change	−0.600***	−0.598***	−0.595***	−0.597***	−0.597***	−0.592***	−0.593***
	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)	(−0.033)
Union membership	−0.043	−0.043	−0.040	−0.044	−0.044	−0.043	−0.040
	(−0.041)	(−0.041)	(−0.041)	(−0.041)	(−0.041)	(−0.041)	(−0.041)
Occ. Experience	0.143***	0.142***	0.140***	0.142***	0.142***	0.140***	0.140***
	(−0.027)	(−0.028)	(−0.028)	(−0.028)	(−0.028)	(−0.028)	(−0.028)
d Aggregate e-to-p	0.004						
	(−0.022)						
d Routine e-to-p		0.056*					
		(−0.023)					
d NR cognitive e-to-p			0.156***				
			(−0.023)				
d NR manual e-to-p				−0.261***			
				(−0.021)			
d Routine emp share					0.154***		
					(−0.022)		
d NR manual emp share						−0.300***	
						(−0.022)	
d NR cognitive emp share							0.300***
							(−0.022)
Routine occ. before change	−0.024	−0.023	−0.023	−0.020	−0.020	−0.018	−0.025
	(−0.031)	(−0.031)	(−0.031)	(−0.031)	(−0.031)	(−0.031)	(−0.031)
NR manual occ. before change	0.038	0.039	0.039	0.039	0.041	0.041	0.036
	(−0.038)	(−0.038)	(−0.038)	(−0.038)	(−0.038)	(−0.038)	(−0.038)
AIC	49910	49903	49863	49755	49860	49725	49722

Source: SIPP data
d: change; emp share: employment share, e-to-p: employment-to-population ratio, NR: non-routine
Significance codes: '***' 0.001, '**' 0.01, '*' 0.05
Standard errors are in brackets under the corresponding coefficient.

is declining leading to higher chances of unemployment in the whole sample. The pattern of the signs remains the same after controlling for the occupational category before the change (see Specifications 6 to 9). The coefficients of the occupational category variables are not significant, just as in the case of the 1991 panel, however, their sign is the opposite of the ones found in Table 5.1. In other words routine workers in the 2001 panel who changed job were less likely to become unemployed upon job change compared to non-routine cognitive workers, while non-routine manual employees were more likely to end up without a job after the change. Using changes in employment share instead of employment-to-population ratios does not change on this result as it can be seen in Specifications 10 to 12. The coefficients for the employment share variables are significant and they do not give different information as the coefficients of the employment-to-population ratio variables. These results are in line with Figure A.2, where the non-routine cognitive and the routine employment-to-population ratios move together more closely and the employment share of the non-routine manual occupations goes in the opposite direction. The coefficients of the dummies for the occupational categories before the change are not significant in these specifications either.

Table 5.5 contains the main results of the other three regressions where the dependent variables were the probabilities of moving to one of the three occupational categories after the change. The coefficients are quite similar to the ones for the 1991 panel, females are more likely to move towards non-routine occupations than males, and people with higher education have a higher chance of being employed in non-routine cognitive occupations. Union membership has a positive effect on the probability of having a non-routine manual or a routine job after the change. The occupational experience, which is an additional variable compared to the 1991 panel has a positive effect in the second specification only, so it increases the probability of moving to a non-routine cognitive occupation after the job change.

In none of the three specifications of Table 5.5 was the coefficient of the change in the aggregate employment-to-population ratio significant, while the coefficients of the binary variables reflecting the occupational category before the change always were. In these coefficients it is possible to see the presence of polarization, unlike in the case of the 1991 panel. The probability of being employed in a routine occupation after the job change is smaller now for people who were employed in any of the main three categories than for individuals who were unemployed before the change. It is still more likely to have a routine job after the change for the ones employed as a routine worker previously than for the ones who were non-routine workers; however, compared to the ones who

Table 5.5: Probability of having a job after job change, 2001 panel

	Routine	NR cognitive	NR manual
Intercept	0.523***	−3.256***	−0.400***
	(0.045)	(0.066)	(0.058)
White	0.135***	0.202***	−0.136***
	(0.022)	(0.033)	(0.029)
Male	0.366***	−0.210***	−0.331***
	(0.018)	(0.026)	(0.026)
Age	−0.078***	0.093***	−0.220***
	(0.019)	(0.028)	(0.027)
School	−0.143***	0.745***	−0.268***
	(0.009)	(0.013)	(0.013)
Employer change	0.660***	0.081*	0.610***
	(0.022)	(0.032)	(0.029)
Laid off	0.201***	−0.021	−0.572***
	(0.048)	(0.066)	(0.075)
Wage before change	0.057**	0.131***	−0.011
	(0.021)	(0.026)	(0.032)
Hours before change	0.306***	0.446***	−0.100*
	(0.030)	(0.041)	(0.042)
Union membership	0.069	−0.163**	0.048
	(0.037)	(0.051)	(0.060)
Occupational experience	−0.100***	0.155***	−0.071*
	(0.020)	(0.025)	(0.032)
d Aggregate e-to-p	0.003	0.027	−0.031
	(0.018)	(0.025)	(0.025)
Occ. before change:			
Routine	−0.536***	−1.193***	−1.352***
	(0.039)	(0.056)	(0.053)
NR manual	−1.927***	−1.213***	0.477***
	(0.042)	(0.062)	(0.049)
NR cognitive	−1.713***	0.159**	−1.353***
	(0.046)	(0.058)	(0.067)
AIC	71989	41981	43201

Source: SIPP data
d: change; emp share: employment share, e-to-p: employment-to-population ratio, NR: non-routine
Significance codes: '***' 0.001, '**' 0.01, '*' 0.05
Standard errors are in brackets under the corresponding coefficient.

move from unemployment to routine jobs, their chances are 13% smaller. This outcome allows drawing conclusions about the effect of polarization on the level of individuals, as it implies that moving towards the routine occupations is more likely to happen after some time spent in unemployment than straight from some other job. In the other two specifications of Table 5.5 there is no similar result; the coefficients show the same pattern as in the 1991 sample.

One can now calculate average predictive differences for the specifications in Tables 5.2 and 5.5 and compare the values to measure the change in the

Table 5.6: Comparison of average predictive differences between the two panels

	1991	2001
$\mathbb{P}$(R after | R before) - $\mathbb{P}$(R after | U before)	11%	-12%
$\mathbb{P}$(R after | R before) - $\mathbb{P}$(R after | NRM before)	56%	30%
$\mathbb{P}$(R after | R before) - $\mathbb{P}$(R after | NRC before)	56%	26%
$\mathbb{P}$(R after | NRM before) - $\mathbb{P}$(R after | NRC before)	0%	-4%
$\mathbb{P}$(NRC after | R before) - $\mathbb{P}$(NRC | U before)	-6%	-15%
$\mathbb{P}$(NRC after | R before) - $\mathbb{P}$(NRC after | NRM before)	-1%	0%
$\mathbb{P}$(NRC after | R before) - $\mathbb{P}$(NRC after | NRC before)	-47%	-18%
$\mathbb{P}$(NRC after | NRM before) - $\mathbb{P}$(NRC after | NRC before)	-46%	-18%
$\mathbb{P}$(NRM after | R before) - $\mathbb{P}$(NRM after | U before)	-14%	-16%
$\mathbb{P}$(NRM after | R before) - $\mathbb{P}$(NRM after | NRM before)	-52%	-26%
$\mathbb{P}$(NRM after | R before) - $\mathbb{P}$(NRM after | NRC before)	0%	0%
$\mathbb{P}$(NRM after | NRM before) - $\mathbb{P}$(NRM after | NRC before)	53%	26%

Source: SIPP data
NRM: non-routine manual occupation, NRC: non-routine cognitive occupation,
R: routine occupation, U: unemployed

probability of getting a routine, non-routine cognitive and non-routine manual job between the two samples. Then these differences can be compared across samples. Table 5.6 contains the average predictive differences for specifications in Tables 5.2 and 5.5. It is clear from the table that the average probability advantage of getting a routine job after the job change declined a lot between the two periods for those who had a routine job before the change compared to all other groups. In the case of the other two specifications the differences in the probabilities are similar in sign between the two samples, although they are generally smaller in absolute value in the 2001 period. The average probability disadvantage of routine workers in finding a non-routine occupation, however, declined between the two period, implying that even though routine workers were less likely to acquire a routine job on average after the change, their chances at working as a non-routine worker are less bad as previously. Those who had a non-routine manual job before the change also witnessed an increase in their probability of moving to a non-routine cognitive job after the change, while their chance to work as a routine worker upon change also declined. These results strongly suggest by the time the 2001 SIPP panel was executed, the polarization was present on the labor market.

While the logistic regressions on the dataset of job changers in the 1991 panel do not show the presence of polarization, the same regressions on the data from one decade later give the opposite result in this regard. The fact that polarization can be detected only in the 2001 dataset is not so surprising, as the graphs in Section 3 suggested that the phenomenon was more strongly apparent in the 2001 panel of the SIPP and it was almost not observable in the 1991 panel. It is therefore reassuring to see that the regression models are in line with the graphs. The outcome of the last three regressions is also robust to changes in the control variable for the labor market movements in the period of the job changes. All in all the regressions show that the movement in the labor market aggregates tend to have an effect on the individual movements between the different occupational categories in both panels, however, the regressions shed only a little more light on the deeper nature of these movements. The most important result of the logit models is that between the two periods the probability of getting a routine job after the job change became less likely for those, who had the same type of employment before the change. This result implies that movement from any other occupational category towards the routine one is more likely to happen after some time is spent in unemployment before reaching the routine occupation.

6 Conclusion

The goal of this essay was to investigate the topic of the labor market polarization in the U.S. economy from the perspective of its effects on the labor market decisions of the individuals. So far the research papers dealing with polarization focused mostly on the causes of the phenomenon and on its links with the business cycle. The main findings of this literature was that the principal cause of polarization is the skill-biased technological change which resulted in the 1980s and 1990s in higher relative demand for more educated workers, while later it also lead to an increase of the employment share of non-routine manual or low-skilled occupations (Autor et al., 2003; Autor and Dorn, 2013). The debate about the relationship of the polarization and the business cycles is not settled so far, as Jaimovich and Siu (2013) argue that the polarization leads to the joblessness of the most recent recessions, while Foote and Ryan (2012) find that the recessions of 2001 and 2007 are not polarizing and workers suffer from higher separation rates regardless of their occupational category. Foote and Ryan (2012) also show that workers tendentiously move back to the same type of job that they had before they became unemployed.

My research differs from the ones quoted in this paper is on two major points. First of all I use the SIPP as data source instead of the CPS which is generally used in the polarization literature; secondly I examine the feedback effect from the aggregate level on the employment decisions of the individuals with special interest on the movement between the three occupation categories defined in the literature (routine, non-routine cognitive and non-routine manual) and the unemployment. As I focus on the movements of the individuals between these four possibilities, the SIPP is a better choice than the CPS, as it has data about each sample member over a longer period of time. I use the 1991 and 2001 panels of the survey, both of which collect data for two and a half years. Using two distinct panels makes it possible to compare the effects of polarization in two separate stages of the polarization. While in the beginning of the 1990s polarization was less strongly present (see Figure 3.1), in the early 2000s it is more pronounced in the SIPP data.

As I am interested in the effect of polarization on the movements of the individuals between the three occupational categories and the unemployment, I estimate logistic regression on the subsample of those individuals, who experienced job change over the sample period. Job change is defined in both samples as occupational or employer change. In the regressions I control for the personal and employment characteristics of the sample members, including the parameters of the job changes. I include in the regression variables which contain information about the wage and hours worked on the workplace before the change, assuming that higher values in these variables result in smaller probabilities for moving to unemployment. I also control for the union membership and that whether somebody was laid off or not on his or her previous workplace. The most important variables are, however, the ones which carry information about the occupational category to which the person belonged before the change and the ones which contain the change in the employment-to-population ratios and in the employment shares of the three occupational categories in the period when the job change happened. As in most cases the change happened from one month to the next, therefore the changes in the aggregate rates are relatively small.

The regression outputs introduced in Section 5 show that there is a difference in the transition probabilities between the two periods. The results show that neither in the 1991 nor in the 2011 panel was the probability of being unemployed after the job change significantly affected by the occupational category to which a person belonged to before the job change. However, the change in the employment-to-population ratios and the employment shares of the non-routine occupations have a significant effect on these probabilities in both periods. The sign of the coefficients for the non-routine manual and non-routine cognitive ratios are opposing each other, which is surprising in the sense that I expected that an increase in all of these rates will have a diminishing effect on the probability of not having a job after the job change, as an increase in these ratios would suggest an increasing tendency in the aggregate employment. Comparing this result with Figures 3.1 and 3.2 shows that the positive coefficient for the change in the non-routine cognitive rates is actually not out of the ordinary, because this rate moves in the opposite direction as those of the non-routine manual occupations, implying that in periods when the non-routine cognitive employment expanded, the non-routine manual occupations faced a decline, therefore the overall job finding probabilities were less favorable as one would conclude from the non-routine cognitive rates only.

The transition probabilities between occupational categories reveal that polarization is present in the 2001 panel, while in the 1991 panel its effects are not observable. The results for the 1991 panel show that the probability of moving to a given occupational category was the highest for those who worked in the same type of occupation before the job change compared to the chances of getting a job in this category by someone who was unemployed beforehand. People who worked in other types of occupations had an even smaller chance to do so. These probabilities are different in the 2001 panel, especially the probability of moving to a routine occupation after the job change. The probability of getting a routine job is smaller for every individual who worked before the job change than for those who were unemployed before the change. This means that the advantage of occupational experience is lost for those who worked in a routine occupation before the change, while non-routine workers still have this advantage in probability over those who get a job after unemployment in their categories. This result is important, because it implies that the movement towards routine occupations indeed became more difficult in the 2001 panel, and that it is more likely that people who got a routine job in its period were unemployed beforehand.

Comparing average predictive differences across the two panels also strengthens this result and additionally it reveals that the probability disadvantage of routine workers in transitioning to non-routine occupations declined between the two samples.

All in all the regression results show that polarization had an effect on the transition probabilities in the 2001 panel, and that the changes in the employment-to-population ratios and employment shares have an effect on the probability of transiting to unemployment from employment. The feedback effect from the aggregates, however, is not significant on the movements between the three occupational categories in both samples. Based on the regression I draw the conclusion that when polarization is present in the sample than it appears in the transition towards the routine occupations by making it less likely as in the 2001 sample it is more likely to get a routine job after some time spent in unemployment. This result is actually the one I am interested in on the first place as it shows, that the movement between the different occupational categories includes some time spent in unemployment at least when the destination of the transition is a routine occupation.

Obviously there are plenty of directions in which this research line could be continued. Here I would like to mention only two of these possibilities. First of all, one could examine the same transition probabilities in other SIPP panels, which could lead to a more precise picture about polarization. It would be very interesting to see how these probabilities turned out after the recession in 2007. One could also include variables related to the program participation of the sample members in order to examine the welfare background of those who are affected by the polarization. This could be helpful in drawing up policy reaction to polarization. The other possible extension I would like to mention here is that one could incorporate such aggregate variables in the model which can have an effect on labor market decisions. Autor et al. (2003) for example control for capital deepening and investment in computer capital on the industry level and show that the latter variable has an effect on the changes in the relative demand for routine workers. It would be worthwhile to include similar variables in the model in order to see whether the effect of the technological investment is still present in the data, or it was only important in the earlier periods of the polarization.

References

ACEMOGLU, D. (1999): "Changes in unemployment and wage inequality: An alternative theory," *American Economic Review*, 89, 1259–1278.

AUTOR, D. (2010): "The Polarization of Job Opportunities in the U.S. Labor Market: Implications for Employment and Earnings," .

AUTOR, D. H. AND D. DORN (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

AUTOR, D. H., L. KATZ, AND M. KEARNEY (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90, 300–323.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The Skill Content Of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

FOOTE, C. L. AND R. W. RYAN (2012): "Labor-market polarization over the business cycle," Tech. rep.

FUJITA, S., C. J. NEKARDA, AND G. RAMEY (2007): "The cyclicality of worker flows: new evidence from the SIPP," Tech. rep.

GELMAN, A. (2008): "Scaling regression inputs by dividing by two standard deviations," *Statistics in Medicine*, 27, 2865–2873.

GELMAN, A. AND J. HILL (2007): *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge, UK: Cambridge Univerisity Press, third ed.

GOOS, M., A. MANNING, AND A. SALOMONS (2010): "Explaining Job Polarization in Europe: The Roles of Technology, Globalization and Institutions," .

HATCH-MAXFIELD, J. AND K. W. ROBERTSON (2012): "Data Collection in the U.S. Bureau of Labor Statistics' Current Employment Statistics Survey," Tech. rep.

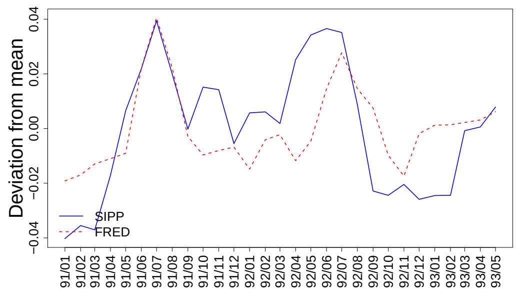
JAIMOVICH, N. AND H. E. SIU (2013): "The Trend is the Cycle: Job Polarization and Jobless Recoveries," .

TÜZEMEN, D. AND J. WILLIS (2013): "The vanishing middle: job polarization and workers' response to the decline in middle-skill jobs," *Economic Review*, 5–32.

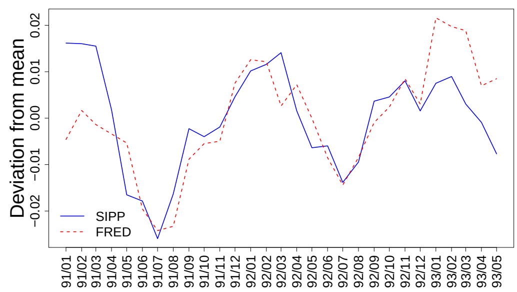
WESTAT (2001): *Survey of Income and Program Participation Users Guide.*, Washington, D.C.: U.S. Census Bureau, third ed.

A Appendix

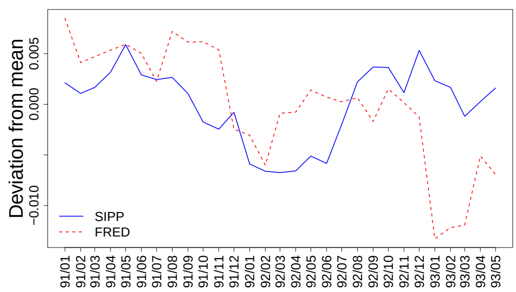
Figure A.1: Employment share, 1991 panel



(a) Non-routine manual



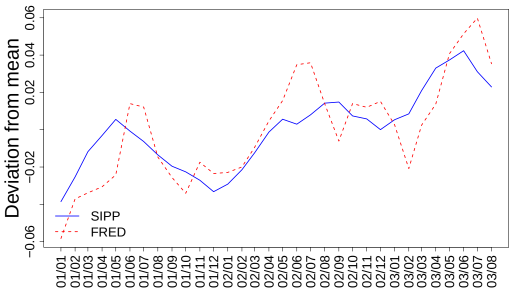
(b) Non-routine cognitive



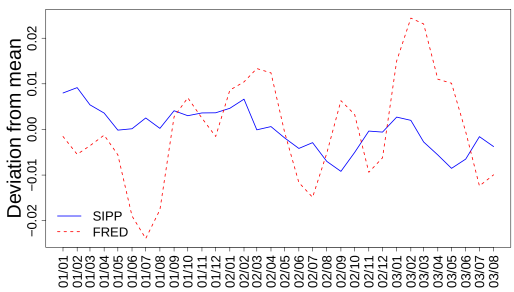
(c) Routine

The employment share is calculated as the number of observations in given occupational category divided by the total number of employed observations in the 1991 panel. Longitudinal panel weights are used in the calculation. The graphs show the deviation from the mean, normalized with the mean for each occupational category and for the aggregate employment. The sample population includes those who are between the ages of 16 and 64, are neither self-employed nor working in farming occupations. For the FRED data the population is defined as the working age population excluding self-employed individuals.

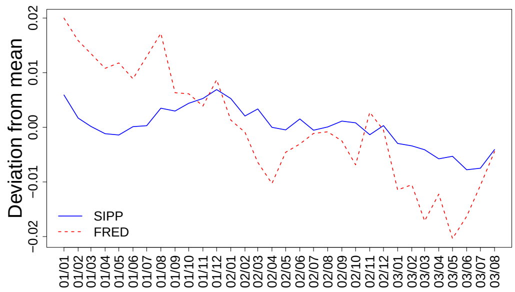
Figure A.2: Employment share, 2001 panel



(a) Non-routine manual



(b) Non-routine cognitive



(c) Routine

The employment share is calculated as the number of observations in given occupational category divided by the total number of employed observations in the 2001 panel. Longitudinal panel weights are used in the calculation. The graphs show the deviation from the mean, normalized with the mean for each occupational category and for the aggregate employment. The sample population includes those who are between the ages of 16 and 64, are neither self-employed nor working in farming occupations. For the FRED data the population is defined as the working age population excluding self-employed individuals.