DISCUSSION PAPER SERIES

IZA

IZA DP No. 10129

Does Rosie Like Riveting?
Male and Female Occupational Choices

Grace Lordan
Jörn-Steffen Pischke

August 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Does Rosie Like Riveting?
Male and Female Occupational Choices

Grace Lordan
LSE and IZA

Jörn-Steffen Pischke
LSE and IZA

Discussion Paper No. 10129
August 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 10129
August 2016

ABSTRACT

Does Rosie Like Riveting? Male and Female Occupational Choices*

Occupational segregation and pay gaps by gender remain large while many of the constraints traditionally believed to be responsible for these gaps have weakened over time. Here, we explore the possibility that women and men have different tastes for the content of the work they do. We run regressions of job satisfaction on the share of males in an occupation. Overall, there is a strong negative relationship between female satisfaction and the share of males. This relationship is fairly stable across different specifications and contexts, and the magnitude of the association is not attenuated by personal characteristics or other occupation averages. Notably, the effect is muted for women but largely unchanged for men when we include three measures that proxy the content and context of the work in an occupation, which we label 'people,' 'brains,' and 'brawn.' These results suggest that women may care more about job content, and this is a possible factor preventing them from entering some male dominated professions. We continue to find a strong negative relationship between female satisfaction and the occupation level share of males in a separate analysis that includes share of males in the firm. This suggests that we are not just picking up differences in the work environment, although these seem to play an independent and important role as well.

JEL Classification: J16, J4

Keywords: occupational choice, job content, gender, preferences

Corresponding author:

Grace Lordan
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
United Kingdom
E-mail: g.lordan@lse.ac.uk

* We thank Alan Manning, Kelly Bedard, Dan Black, Alex Bryson, Deborah Cobb Clark, Stephen Machin, David Neumark, Ron Oaxaca, Barbara Petrongolo, and David Ribar for helpful discussions and participants of the CEP Labor Market Workshop at the LSE, 14th IZA SOLE transatlantic workshop, 18th IZA European Summer School in Labor Economics, 2016 Royal Economic Society meeting, and 2016 Society of Labor Economics Meeting for their comments. We are grateful to the ESRC for support through the Centre for Economic Performance at the LSE.

Introduction

"And finally, in our time a beard is the one thing that a woman cannot do better than a man." - John Steinberg, Travels with Charley: In Search of America.

Women's progress in the labor market has been dramatic since the 1960s. The female employment rate has risen, the pay gap with men has declined, and occupational segregation has decreased. Figure 1 shows the share of males in the occupations in which women work. Despite all the progress, it is striking that female convergence has slowed and possibly stopped since about the turn of the millennium, while sizeable differences remain between the jobs done by women and men. One particular concern is that females are underrepresented in many high paying professional and managerial occupations (see Figure 2 and Goldin, 2014), while average female earnings still lag behind male earnings (Blau and Kahn, 2000, 2006). Since occupational earnings differences are large, the underrepresentation of women in many high-paying, male dominated professions remains a major candidate for the residual gender wage gap (Macpherson and Hirsch, 1995; Bielby and Baron, 1984; Bayard et al., 2003). For example in 2014, the average hourly wage of individuals who work in majority male occupations (proportion of males >=0.70) is \$23.67, versus \$19.30 for those in minority male occupations (proportion of males <=0.30).[1]

The traditional explanations for these wage gaps are discrimination, labor supply, and human capital investments, as well as barriers, which make it difficult to combine work and family.[2] More recently, the literature has turned towards the role of attitudes, personality traits, and gender identity as possible explanations for different labor market choices and outcomes of men and women (e.g. Croson and Gneezy, 2009 and Bertrand, 2010). However, the role of many of the variables suggested as explanations for lower female earnings remain empirically elusive (Manning and Swaffield, 2008).

In this paper we focus on tastes for particular job attributes as an explanation for the remaining occupational segregation. We estimate job satisfaction equations for men and women. We complement these with regressions for leaving a particular occupation between two periods. One of our key regressors is the share of men in an occupation. Like Usui

1. Based on the 2014 Current Population Survey (CPS) monthly outgoing rotation group data.
2. See Altonji and Blank (1999). Women's "Second Shift" when they combine market work with home making is portrayed by Hochschild and Machung (1989).

(2008), who uses the National Longitudinal Survey of Youth 1979 (NLSY79) from 1979-1982, we find that that women are less satisfied in male dominated jobs, while males either like or are indifferent to the share of males in an occupation**.** In addition, we find that women are more likely to leave occupations with a higher share of males, whereas males are more likely to stay. This mirrors findings by Hunt (2016) for US college graduates. We document these relationships for the US, Britain, and Russia. The basic patterns are robust to including many other occupation and individual characteristics, as well as individual fixed effects.

The core of our analysis links job satisfaction to attributes of the work done in various occupations. The idea is that women may not like the nature of male dominated jobs. This hypothesis underlies work by Pinker (2008), who argues that females and males have different tastes, which lead them to gravitate towards different occupations. She argues that women tend to prefer jobs that require empathy and interacting with people. Conversely, men like work that requires them to 'make things.' Pinker (2008) sees this as the reason why women are less likely to become aerospace engineers and are more likely to enter teaching. To empirically examine this hypothesis, we parsimoniously summarize occupational content in three latent factors, distilled from descriptions in the ONET database, which we label 'people,' 'brains,' and 'brawn.' The occupational content measures matter for both male and female job satisfaction. We find that female job satisfaction is higher in occupations that have high 'people' and 'brain' content but is lower for 'brawn.' Conversely, males are indifferent to jobs that have high 'brawn' content. Importantly, including these measures reduce the coefficient on the share of men in the occupation by a third or more for women, while it does little to the coefficient in the male job satisfaction regressions.

While we argue that these results point to differences in tastes for job attributes, the share of males may also proxy for differences in the work environment, which are perceived differentially by men and women. In order to probe this possibility, we use data from the UK Workplace Employment Relations Survey (WERS) which allows us to relate job satisfaction to both the share of males in the occupation and to characteristics of the firm (captured either by the share of males in the firm or firm fixed effects). We find that the share of males in the occupation matters even in the presence of firm level controls. However, the share of males in the firm also matters separately from the share of males in the occupation, pointing to the potential importance of factors related to work environment. We also find that the time variation in the share of males matters in our job satisfaction regressions. Compared to the cross sectional variation in the share of males that we focus on in our main analysis, we suspect that the time variation is also likely to pick up factors related to work environment or organization.

While our results are basically descriptive, they suggest that differential tastes by gender may be an important ingredient in explaining the occupational choices of men and women. We remain agnostic regarding the origin of these differences in preferences, which could be biological, evolutionary, or caused through socialization. While direct preference based explanations have been gaining prominence in some quarters they have not featured prominently in the economics literature so far. A recent exception is Wiswall and Zafar (2016), who find that attitudes towards job attributes differ between male and female college students in hypothetical choice experiments. These differences predict both major choices in college and job choices. However, the job attributes they focus on are mostly more traditional factors like earnings, hours, flexibility, and job security rather than tastes for the content of work per se.

Also related is a literature which has focused on biological differences between men and women. Baker and Cornelson (2016) link the share of men in an occupation to DOT codes that capture the sensory, motor, and spatial skills required in particular occupations. They find that occupational segregation would have been about 25% lower if these skills did not vary by gender but that the skills did not play a role in the narrowing of the occupation gap during the past 40 years. We suspect that their skills pick up some related variation to our 'people,' 'brains,' and 'brawn' factors; however, Baker and Cornelson (2016) do not relate their skills to job satisfaction.[3]

Data

US NLSY79:

We use the NLSY79, a panel of 12,686 individuals who were between 14 and 22 years old when first surveyed in 1979. These individuals were interviewed annually through 1994 and then on a biennial basis. The NLSY79 sample spans 1979 to 2012.

3 Weinberg (2000) is an earlier analysis along these lines.

The question on job satisfaction was asked in every wave. Specifically, respondents were asked, “How do you feel about the job you have now?” and were given the following response option: ‘I like it very much’; ‘I like it fairly well’; ‘I dislike it somewhat’; ‘I dislike it very much’. We coded responses so that higher values represent higher satisfaction. Our analysis is restricted to an unbalanced panel of employees who responded to this job satisfaction question. The NLSY79 uses the US Census Bureau occupation definitions. Specifically, the 1982-2000 and 2002-2012 waves use the 1980 and 2000 codes respectively. Our analysis sample spans the years 1982 to 2012.

We create an additional dependent variables that captures movements in the labor market.[4] This variable is defined equal to 1 if a person has the same three digit occupation code in t+2 compared to the occupation that they held in t. Conversely, the variable is defined equal to 0 if an individual has a different occupation code in t+2 or has left employment. We call this variable ‘stayers.’ The variable is defined on a biennial basis given the interview schedule of the NLSY79 post 1994.

We use the pooled monthly CPS samples from 1983-1991 and 2003-2010 to calculate the proportion of males in each occupation for the 1980 and 2000 three-digit occupation codes respectively.[5] In particular, the share of males (SOM) is the count of men within occupation j divided by the total number of workers in the same occupation. Additionally, we calculate averages of the hourly wage, hours, the proportion college graduates, and age for each occupation.[6] We match the CPS averages derived from the 1980 occupation codes to the 1982-2000 NLSY data and the averages derived from the 2000 occupation codes to the 2002-2012 NLSY data. There is a single average for all the years within the sub-periods when occupation codes are unchanged. Hence, we exploit cross-sectional variation and variation due to occupation switchers but not variation over time in these averages in the estimation (except in Table 10, where we analyze the time variation separately). In order to allow for the break in the occupation coding, we control for individual times sub-period specific fixed effects in some of our regressions. We also utilize sampling weights that reflect that the NLSY79 oversampled blacks, Hispanics, and the economically disadvantaged (see Appendix D for the unweighted results).

---

4. Give that this outcome relies on comparing occupation codes across periods, this analysis omits the year 2000 from the analysis given the change in occupation coding

5. From 1992- 2002 the CPS uses 1990 occupation codes.

6. Hourly wages are calculated from the Merged Outgoing Rotation Groups.

British Household Panel Survey (BHPS):

We use all 18 waves of the original sample of the British Household Panel Survey (BHPS), a longitudinal study of around 5,500 households and over 10,000 individuals in England, Wales and Scotland that began in 1991. This main sample was supplemented in later years with a Welsh extension from 1999 (about 1500 households), a Scottish extension from 1999 and a Northern Ireland extension from 2001 (about 1900 households). We present unweighted results from the unbalanced panel of all individuals including the extensions between 1991 and 2008.[7]

The BHPS contains a number of different job satisfaction questions, which are available for the full 18 waves. We use the two questions asking respondents how satisfied or dissatisfied they are with i) their current job overall and ii) the actual work itself. Answers are on a 7-point scale. The BHPS uses occupation codes based on the Standard Occupational Classification 1990 (SOC90) up to 2001; in 2002 this was replaced with SOC 2000 (SOC00).

We again create an additional binary dependent variable that captures whether a person stayed in the same occupation. We measure mobility in the BHPS between two consecutive years.[8]

We calculate the SOM and other occupation averages in a three-digit occupation using the 1993-2012 Quarterly Labor Force Survey (QLFS). The QLFS is the main survey of individual economic activity in the Britain, and provides the official measure of the national unemployment rate. It uses SOC90 codes from 1993 through 2000 and SOC00 from 2001. Thus, we calculate the same occupation averages as for the NLSY for each sub-period when the SOC90 and SOC00 were in use. We then match the occupation averages to the BHPS data. We allow for individual sub-period specific fixed effects in some of our regressions.

---

7. We have investigated the sensitivity of our results to i) unweighted regressions of the original BHPS sample only ii) weighted regressions of the main BHPS sample, where the weights are the longitudinal weights described in Taylor et al (2010) (these are the weights recommended for use in longitudinal analysis, however we lose a significant amount of our sample owing to these weights only being provided when an individual was present in all waves. The conclusions in this work are robust to these changes. See Appendix D for these results.

8. This outcome relies on comparing occupation codes across periods, therefore this analysis omits the year 2002 from the analysis given the change in the occupation codes.

British Workplace Employment Relations Survey (B-WERS):

We use the 2004 and 2011 years of the British Workplace Employment Relations Survey (B-WERS), which included an individual's three-digit occupation code (previous versions did not), along with a series of questions concerning various elements of job satisfaction. This is a national survey of people at work in Britain, which collects data on employees, employee representatives, and employers in a representative sample of about 2500 firms. We utilize the employee responses to the question about satisfaction with the work itself. Response options are on a 5-point scale. We also utilize a variable that is reported by the employer or senior manager that quantifies the SOM in their firm. We calculate the SOM in the employee's occupation along with other occupation averages from the QLFS in the same manner described for the BHPS data and match this to B-WERS based on the employee's three-digit occupation code.

Russian Longitudinal Monitoring Survey (RLMS):

Our measure of job satisfaction for Russia comes from the Russian Longitudinal Monitoring Survey (RLMS). This is a series of nationally representative annual surveys, with data available from 1994-2012. However, job satisfaction data is only available from 2002-2012. We restrict our sample to employees who answer the question: 'How satisfied or unsatisfied are you with your job in general?' Response options are absolutely satisfied, mostly satisfied, neutral, not very satisfied and absolutely unsatisfied. We code responses so that higher values represent being more satisfied. We create a binary dependent variable that captures whether a person stayed in the same occupation over two consecutive years.

We do not have a large labor force survey that allows us to calculate occupation averages for Russia, like the US CPS or British QLFS. Instead, we rely on merging the RLMS from 1994-2012 with two other data sources, the International Social Survey Program (ISSP) 1995-2011[9] and the European Social Survey (ESS) from 2002-2012.[10] Pooling the ISSP 1995-2011, the ESS 2002-2012 and the RLMS 1994-2012, we calculate the SOM in each occupation, along with the other occupation averages, age, hours and proportion of college graduates. Only the RLMS reports individual earnings and, as a result, we calculate the average wage from this data source only. Our RLMS regressions use weights that allow for

[9] http://www.issp.org/page.php?pageId=4
[10] http://www.europeansocialsurvey.org/

the complex design of the RLMS where many observations are derived from following the housing unit rather than the person, as well as having oversamples from the first wave to allow for forecasted attrition. However, the overall conclusions are not sensitive to weighting, and we show unweighted regressions in Appendix D.

Methods and Results

Our starting point is a linear regression for job satisfaction or mobility of the form

$$Y_{ijt} = \delta SOM_{js} + X_{js}\beta + X_{ijt}\gamma + \mu_t + \varpi_\alpha + \varepsilon_{ijt} \quad (1)$$

where Y_{ijt} is either job satisfaction or a binary variable which indicates whether a person stayed in the same occupation in the next period for individual *i* in occupation *j* and year *t*, SOM_{js} is the proportion of males in a particular occupation, X_{js} is a vector of other occupational averages, X_{ijt} is a vector of individual-level control variables, μ_t are wave effects, and ϖ_a are region effects.[11] In the baseline specification, X_{js} contains average wages, hours, age, and the proportion college graduates, while X_{ijt} contains age and age squared. We calculate standard errors using two-way clustering by individual and occupation.[12]

The coefficient of interest in equation (1) is δ. For example, in stayer regressions, a positive coefficient implies that a higher SOM in an occupation is associated with a higher tendency to stay in that occupation. For the job satisfaction regressions, a positive coefficient implies that a higher SOM in an occupation is associated with higher levels of job satisfaction. To make the interpretation of δ more intuitive in the job satisfaction regressions (given that the job satisfaction scales differ across country) we follow van Praag and Ferrer-i-Carbonell (2008) and normalize the job satisfaction variables by using the fitted values from an ordered probit on the raw sample fractions. We estimate equation (1) separately for males and females.

[11] For the BHPS this amounts to the inclusion of 19 fixed effects representing the following regions: inner London, outer London, rest of the South East, South West, East Anglia, East Midlands, West Midlands Conurbation, Rest of the West Midlands, Greater Manchester, Merseyside, Rest of the North West, South Yorkshire, West Yorkshire, Rest of Yorks and Humberside, Tyne and Wear, Rest of the North, Wales, Scotland and Northern Ireland. For the United States, regions are at a higher level, so we control only for whether the respondent resides in the North East, North Central, South or West. For Russia we include eight individual residential site indicators.

[12] See Cameron, Gelbach, and Miller (2011). Practically this is implemented using ivreg2 and xtivreg2 as appropriate in Stata.

Table 1 displays our baseline results for job satisfaction. The SOM in an occupation is consistently associated with lower levels of job satisfaction for women, and the magnitudes seem sizeable. For the US, the coefficient on the SOM is -0.209. This implies that a 10-percentage point increase in the SOM (approximately the effect of moving from more female accounting to more male pharmacy (see Figure 2)) is associated with 2.1% of a standard deviation lower job satisfaction. For the BHPS, a 10-percentage point increase in the SOM has a lower association at 1.2%. For Russia a 10-percentage point increase in the SOM is associated with 1.3% of a standard deviation lower job satisfaction. For the BHPS, in addition to overall job satisfaction, we also have a measure of satisfaction with work itself. A 10-percentage point increase in the SOM in an occupation is associated with a larger 2.8% of a standard deviation decline in satisfaction, compared to the 1.2% lower effect for overall job satisfaction.

The results for males, on the other hand, are much smaller in magnitude, not significant and centered closely around zero. These regressions control for a number of other occupation averages: the log of wages, hours, age, and the fraction of college graduates. Particularly age, wages, and the fraction of educated workers are important correlates with job satisfaction but for women the SOM certainly plays a sizeable role in explaining job satisfaction**.**

There are other data sources which allow us to consider the association between the SOM and overall job satisfaction. We find the same pattern of results in the (US) General Social, the European Social Survey (ESS) and the International Social Survey Programme (ISSP). That is, females are less satisfied working in occupations with higher SOM. Conversely, the coefficient for males is centered around zero. These results are documented in Appendix E.

Table 2 shows the results from the stayer regressions and compares them to the results for job satisfaction. For brevity, coefficients for occupation averages other than the SOM are not shown in the table. The pattern of results for mobility are very similar to those for job satisfaction. For all three countries, higher shares of males in period *t* increase the likelihood that women change occupation in the next period. For males, our findings for all three countries suggest the opposite. For example, the associations for the US imply that a 10-percentage point increase in the SOM in an occupation increases the probability of a female changing her occupation by 2.3 percentage points. Conversely, for males the same increase implies that they are about 1-percentage point less likely to change their occupation. These effects may seem large but turnover is large in the data. For example, 36% of respondents in our NLSY sample move out of their occupation within two years on average. This result is maybe unsurprising since the SOM ultimately reflects the occupational choices of women and men. It suggests an adjustment process, where women tend to leave male dominated occupations over time if they find themselves starting out in one of these.

How workers sort into heterogeneous occupations is an important issue in interpreting the results from a regression like (1). The standard compensating differentials framework suggests that workers pick among packages of wages and job attributes while employers offer such packages in order to attract workers. To the degree that workers differ, they will sort into the type of jobs they prefer in equilibrium. Wages adjust to eliminate any excess supplies and demands, so that occupation wage differentials reflect the compensating differentials required by marginal workers who are indifferent between two alternative jobs. This framework predicts that men and women may end up working in different jobs in equilibrium if they have different preferences for job attributes or if they face different constraints (say in terms of hours choices or flexible schedules an occupation offers). In this scenario, it is unlikely that job satisfaction will reflect preferences. One reason is that most of the variation in (1) is cross-sectional, and it is unclear whether the answers to job satisfaction questions are comparable across individuals. The fixed effects specifications we explore below address this issue. Another reason is that in the competitive compensating differentials model everybody works in their most preferred occupation, given equilibrium wages, and hence should report their maximum job satisfaction attainable.

The frictionless, full information framework underlying the standard model is unlikely to be a good representation of actual labor markets, where individuals often make choices subject to constraints, imperfect information regarding what an occupation's content is in practice, and other frictions. Occupations are also bundles of attributes but not all possible combinations may be on offer fitting all individual tastes. Modeling occupational choices and wage differentials in a framework with frictions can lead to very different equilibrium outcomes (see e.g. Manning, 2003). One implication is that wages no longer reflect compensating differentials. Rather, employers with wage setting powers will use wage-amenity packages to attract workers, and wages and amenities may be positively correlated in equilibrium.

Furthermore, workers may end up in jobs other than their preferred one, but they will switch jobs in future periods in search of better matches. This "frictional disequilibrium" constitutes a natural source for interpreting the results from job satisfaction equations like (1). As there are good jobs and bad jobs, as well as high and low quality job matches for particular individuals in this framework, the coefficients on occupation characteristics have a more natural interpretation as individual preferences for these characteristics. Frictions also offer a natural point of departure for interpreting the stayer regressions, as there is no reason for systematic job changes in the frictionless model. However, the caveat that within person comparisons should be more accurate still applies in the model with frictions as well.

Therefore, we add individual fixed effects to equation (1) and show results in Table 3. Including fixed effects amounts to identifying the effect of the SOM from occupation switchers, while controlling for time invariant individual differences.[13] Recall that the occupation coding changed in the US and British data sets over time. In order to exploit only variation within periods with consistent occupation codes, we interact the individual fixed effects with indicators for the sub-periods without coding changes. Denoting these sub-periods by s, we estimate:

$$Y_{ijt} = \alpha_{is} + \delta SOM_{js} + X_{js}\beta + X_{ijt}\gamma + \mu_t + \varpi_a + \varepsilon_{ijt} \qquad (2)$$

From Table 3, including the fixed effects yields somewhat weaker results of the SOM for women in the US, stronger results in Britain, and has little impact in Russia. For males, the SOM coefficients are still centered around zero just as in Table 1. Overall, accounting for fixed effects fails to explain the negative correlation between job satisfaction and the SOM.

The decision to stay in an occupation also remains strongly related to the SOM in all three countries. In particular, a woman who switches into an occupation with a higher SOM is less likely to stay in that occupation. Conversely, a male is more likely to stay.

We have also considered adding a number of individual factors to the specification in equation (2). In particular we include covariates that are traditional in the job satisfaction

[13] Including individual fixed effects also controls for personality traits that are fixed to the individual over time, which have been highlighted as important determinants of labor market outcomes, including locus of control (Fortin 2008) and willingness to compete (Gneezy, Niederle and Rustichini, 2003 and Niederle and Vesterlund, 2007).

literature in the spirit of Clark (1996) and Clark and Oswald (1996), the log of own income, own working hours, household size, number of children, a dummy for college graduates, and marital status. In addition, we created measures for the flexibility of hours in an occupation. The importance of flexibility for females in the workplace has been emphasized in the literature (Goldin, 2014; Goldin and Katz, 2008; Goldin and Katz, 2011; Goldin and Katz, 2016), with some suggesting that career women are 'opting elsewhere' in choosing occupations that allow them to accommodate family responsibilities (Polachek, 1981; Belkin, 2003; Stone, 2007). Adding these variables leaves the coefficients on the SOM unchanged or increases them slightly in absolute value when compared to Table 3 (see Appendix A).

People, brains, and brawn

Why do women report lower job satisfaction when they work in occupations with a high share of males? One hypothesis is that men and women have different preferences for characteristics and attributes of jobs, as well as the environment in which they work. Individuals sort to some degree into jobs according to these preferences, and the observed SOM may therefore reflect male and female tastes. Differences in tastes by gender for particular occupational traits may also explain why our coefficient on the SOM is not significant in the male satisfaction regressions documented in Tables 1 and 3, but for females it is mostly negative and significant. That is, over the past half century females have entered roles that were previously male dominated so they now have some presence in these relatively male jobs. If these are bad matches this will show up in reports of low job satisfaction. However, the gender revolution has been an asymmetric one (see, for example, Figure 2). That is, it is a revolution in which females have increasingly assumed male jobs, but males have not to the same extent moved into traditional female jobs (like nursing and teaching). So, males remain less likely to find themselves working in occupations that require the empathy they lack, while females may more often find themselves in environments that are focused on tasks they may care less about.

In order to probe the possibility that the SOM at the occupation level is picking up the content of the work being done by workers in that occupation, we would like to control for the occupation characteristics which are related to such preferences directly. Therefore, we turn to the ONET database version 5.[14] ONET provides a diverse set of information on occupational attributes, requirements, and characteristics of the workers in an occupation; all in all, it offers about 249 distinct items. Out of these, we start with the 79 items describing the work activities and context of a person's occupation at the US 2000 SOC level. For each individual item, an incumbent in randomly selected firms reports a level from 1 to 7. For example, in activities, an item might describe to which degree an occupation involves 'assisting and caring for others,' 'analyzing data or information,' or the 'repairing and maintaining of mechanical equipment.' Examples for context are the level of 'contact with others,' 'the importance of being exact or accurate,' and 'being exposed to hazardous conditions' (see Appendix C Table C.1 for all attributes). We standardize each of these variables to have a mean of 0 and a standard deviation of 1.

We utilize a crosswalk provided by the Bureau of Labor Statistics to assign a Census 2000 occupation code to each occupation in the ONET file.[15] We then rely on the crosswalks from Autor and Dorn (2013) and Dorn (2009) in order to create a consistent set of occupations, which can be matched to both Census 1980 and 2000 codes. This lets us match the ONET data to the CPS from 1983-2012. Unlike for the SOM, the ONET variables are constant throughout the entire period from 1983-2012 based on the consistent set of occupations created from the crosswalks. This gives us a data set that represents the distribution of occupation characteristics for the US for this period.

We could add the 79 context and activities variables to our regressions directly. However, we are worried about over-fitting, so we follow the psychometric literature (Gorsuch, 1983; Thompson, 2004) and use exploratory factor analysis to reduce the dimensionality of the ONET variables first. To extract the underlying latent factors, we first determine the number of factors to retain based on a scree plot from an orthogonal exploratory analysis and the eigenvalue of each individual factor. A clear structure emerges in the first rotation (see Appendix C Table C.1). We then follow Heckman et al. (2012) and drop ONET items that are weakly associated with the factors or those that are not associated with more than one factor. For the former, we remove items with a loading of 0.4 or less. For the latter we remove items that have a loading that is greater than 0.4 on more than one factor. We then

[14]. We choose to work with this version as we have a crosswalk between the US and British occupation codes mapping ONET 5 to the SOC00. However, we have assessed robustness to using the most recent version of ONET for the NLSY data and these results are reported in Appendix C.
[15]. https://www.census.gov/people/eeotabulation/documentation/jobgroups.pdf

repeat the factor analysis using the remaining ONET items and extract the final latent variables, which we allow to be freely correlated. We loosely label the three factors we obtain as 'people,' 'brains,' and 'brawn,' based on the items that load onto each factor (see Appendix C Table C.2 for full details of the items that load on each factor). We next match the occupation specific factors to the NLSY data. We note that we have explored variations of how to extract these factors. In all cases, our analysis suggested three main factors and results are robust to exactly how we obtained these factors.

Our approach differs from that taken by Beaudry and Lewis (2014), who use the DOT (the predecessor to ONET) to manually pick attributes they view as being associated with physical, cognitive, and people skills in an occupation. We rely on a more mechanical method to reduce the dimensionality of the data to avoid handpicking occupational attributes, which may or may not fit our prejudices. Nonetheless, we arrive at a roughly similar classification. Table 4 lists the top and bottom ten occupations for each of the three factors. In addition, Table 5 documents the scores for a number of occupations, which we find useful for thinking about occupational segregation, together with the share of men in 1930 and now. The factors have a mean 0 and standard deviation 1, so chemical engineers, for example, score about half a standard deviation below the mean on 'people,' one standard deviation above the mean on 'things,' and a bit more than 3 standard deviations above the mean on 'brains.' We note that women have always been dominant in occupations which score high on 'people' (e.g. teaching, nursing, and social work), and are underrepresented in occupations high on 'brawn.' This latter factor does not simply capture physical strength but occupational content related to making and manipulating things as well (see Table 4). The 'brawn' factor is strongly associated with traditional blue-collar occupations but also with engineering fields and isolated other occupations like nursing. As expected, professional and technical jobs tend to be associated with positive 'brains.' The most cerebral occupations are the hard sciences, engineering, and mathematics but also financial managers. This group of occupations also tends to be characterized by having low 'people' content (see Table 4).

In order to investigate these issues more closely, we return to our three country specific micro-datasets. For the British analysis, we match the US SOC00 codes in the ONET data directly to the British SOC00 in the QLFS data using a crosswalk provided by Anna

Salomons.[16] We then proceed as before and extract the underlying latent factors. These differ only from the US analysis in the fact that the distribution of workers across occupations is slightly different in Britain. Unsurprisingly, we again obtain three latent factors corresponding to 'people,' 'brains,' and 'brawn' from the QLFS analysis, which we match to the BHPS. For the Russian data (complementing the RLMS with ISSP and ESS data in order to get more observations in the occupation cells) we match the ISCO code to the US SOC00 using a crosswalk provided by the Bureau of Labor Statistics. The factor analysis again yields the three familiar factors labeled 'people,' 'brains,' and 'brawn' (PBB).

In Table 6, we present a linear regression of the SOM on the three latent factors along with the same other occupational averages as before, time dummies, and area dummies (though we run this at the individual level, note that this is essentially an occupation level regression and the individuals here only serve to give different weights to different occupations). These regressions use the CPS, QLFS, and RLMS. This table highlights that there is substantial sorting in all three countries along the PBB dimension. Women are overrepresented in 'people' jobs, particularly in Russia, men in 'brawn' jobs, and they share 'brain' jobs. Unlike in the examples in Table 5, the 'brawn' component seems to be the more potent predictor of sorting by gender than the 'people' factor, stressed by writers like Pinker (2008). We suspect that this is due to the role of blue-collar jobs in the occupation distribution at large, and the fact that some of the occupations in Table 5 are small (so they get little weight in Table 6). The important take-away from Table 6 is that the three factors explain some of the differential occupational sorting of men and women.

We now return to or main regressions on job satisfaction and occupational mobility, adding the three latent factors denoted PBB_j to equation (2). We estimate:

$$Y_{ijt} = \alpha_{js} + \delta SOM_{js} + PBB_j \chi + X_{js}\beta + X_{ijt}\gamma + \mu_t + \varpi_\alpha + \varepsilon_{ijt} \tag{3}$$

Including PBB_j allows us to investigate how the negative correlation between the SOM and overall satisfaction in the female regressions is related to job content. Hence, we are both interested in the coefficients associated with 'people,' 'brains,' and 'brawn' as well as with

16. For the years in the LFS where the UK SOC90 code is used, we use a translation to SOC00 that is implicitly provided by the BHPS. That is, SOC00 appears for the respondent's primary occupation post 2000 and SOC90 appears for all waves of the survey. So we, have a translation between the two coding systems.

the impact the PBB_j factors have on the coefficient on the SOM.

Table 7 shows the results from an analysis that adds the three latent factors to the job satisfaction regressions. The addition of the PBB variables does not change the coefficients for the SOM in the male job satisfaction regressions, they remain small and centered around zero. On the other hand, the inclusion of these variables has reduced the magnitudes of the SOM coefficients in the female regressions. Specifically, in the US and Russia the SOM coefficient shrinks to close to zero while in Britain it falls to about two thirds its original value. The reductions in all three countries are both sizeable and statistically significant, as indicated by a generalized Hausman test in the last row comparing the SOM coefficients with and without the PBB regressors. These findings suggest that job attributes play an important role in explaining female job satisfaction, and the SOM captures some this when PBB is not included in the regression.

The coefficients on the PBB variables indicate that women tend to be happier in 'people' and in 'brain' jobs. In Britain, this is particularly the case for the 'satisfaction with work itself' variable but less so for overall job satisfaction, consistent with our interpretation that the ONET variables capture something about the content of work. For males, PBB matters less. The coefficients are closer to zero and typically insignificant but the pattern is not that different from women. The sexes bifurcate most with respect to 'brawn:' women tend to dislike these jobs and the effects are always strong and significant for women.[17]

The job satisfaction regressions paint a picture, which is consistent with the idea that men and women have different preferences for job attributes (or women have such preferences while men care much less about what they do at work). The same is not born out in the mobility regressions in Table 8. Here, coefficients on the SOM hardly change when we add the PBB variables. Across the three countries, females are less likely to stay in jobs that are high in 'brawn' and more likely to stay in jobs that are high in 'people.' However, in Britain males

[17] In order to consider whether the finding that the sexes bifurcate most on 'brawn' is driven solely by females being less physically able than males, we also add to the same regression a measure of occupation physicality. This measure averages the values of 'handling and moving objects' from the activities category. This attribute was chosen as the characteristic that requires females to carry weight, which may place them at a disadvantage if their body strength is lower. While adding these to the regressions in Table 7 does change the coefficient on brawn, it is still -0.023 and significant for females for the US. The male coefficient is -0.007 and not significant. In the UK the brawn coefficient decreases to -0.005 and -0.010 in the overall satisfaction and work itself regressions, however for males the coefficients are now significant at the 1% level and positive in these regressions. Specifically, the new coefficients are 0.020 and 0.015 respectively. Finally, for Russia the coefficient has reduced to -0.022 and the male coefficient has risen to 0.015. This leads us to conclude that that aversion to the job content of brawn jobs may have something to do with physical strength but that is not the full story. Importantly, the finding that the sexes seemingly differ in their tastes for this component of job content remains.

are equally likely to leave ‘brawn’ jobs, and in the US males are equally likely to stay in ‘people’ jobs. Some of this pattern seems to be due to the decline in blue-collar jobs.

How should we interpret the different results in Tables 7 and 8? Economists tend to favor evidence based on revealed behaviors over stated preferences. Nevertheless, there are reasons to look to the job satisfaction regressions at least as a complement to the mobility regressions, and possibly as a more informative analysis. The SOM variable itself captures the sorting of men and women into occupations, creating the potential for a mechanical relationship with the occupational mobility variable on the left hand side. The lower the SOM, the fewer opportunities there are to change occupations and end up in an occupation with an even lower SOM. Suppose that the SOM captures occupational characteristics, which men and women care about. If the SOM is lower in an occupation then it is more difficult for women to find another occupation, which is very different because they basically only have higher SOM occupations to choose from. This is not true for women who work in occupations where the SOM takes on intermediate values, as these women can go in both directions when changing jobs. This problem due to the boundedness of the SOM distribution is symmetric, and equally applies to high SOM occupations. Of course, women are mechanically overrepresented in low SOM occupations. This may be a strong force leading to a negative coefficient in the stayer regressions. This problem doesn’t arise for the job satisfaction regressions, where the left and right hand side variables are distinct.

Another explanation for the disparity between the job satisfaction and mobility results is that it is easy to be dissatisfied while it may be more difficult to switch occupations. Occupations may not be ideal bundles of attributes for everyone, and many women may have to make tradeoffs and work in jobs with which they are not completely satisfied. They stay in these jobs because they offer other amenities, including income or job flexibility.[18]

Work environment

The results we have presented so far are consistent with the idea that tastes for the content of work influence occupational choices of women and men. However, the PBB variables are crude measures of work content, and they may well pick up other attributes about the workplace. In particular, they may proxy for environmental or organizational factors, which affect men and women differently. We have tried to guard against this by using a time

[18] Hunt (2016) finds that female college graduates leave male dominated fields of study primarily because of the pay and promotion opportunities in the field.

invariant measure of the SOM (as well as the PBB variables), and our fixed effects regressions are solely identified from occupation switchers. This should make it more likely that these variables pick up occupation characteristics rather than the work environment, which may change over time as women enter some occupations.

A lot of aspects related to the work environment should be specific to a workplace and shaped by the co-workers, i.e. a firm level characteristics rather than a characteristic of the occupation per se. None of the datasets we have analyzed lets us get at this. We therefore turn to the British Workplace Employment Relations Survey (WERS). This data set samples workplaces, and within these workplaces surveys managers, worker representatives, and a subsample of employees. Of interest are the 2004 and 2011 versions of the survey, which included an individual's three-digit occupation code (previous versions did not), along with a series of questions concerning various elements of job satisfaction. We use the question on 'satisfaction with work itself.' The employer questionnaire asks about the gender composition of the workplace.

Our analysis of the WERS data returns to cross-sectional specifications, as we do not observe individual employees switching firms. Although we have focused on the individual fixed effects results above, we recall that the results from simple cross-sectional regressions in Table 1 were fairly similar. As a result, the analysis of the WERS should be informative for our purpose. With the WERS data, we can either add the SOM in the firm in addition to the occupation level SOM to the job satisfaction regressions, or control for firm fixed effects directly. Denoting firms by the index *f*, we estimate the following specifications

$$Y_{ifjt} = \delta SOM_j + \phi SOM_f + X_j\beta + X_{ifjt}\gamma + \varepsilon_{ifjt} \qquad (4a)$$

$$Y_{ifjt} = \delta SOM_j + F_f + X_j\beta + X_{ifjt}\gamma + \varepsilon_{ifjt}. \qquad (4b)$$

Our working assumption is that the work environment is determined primarily at the workplace rather than at the occupation level but certain occupations will surely cluster in particular firms. Hence, if the PBB variables and the occupation level SOM only pick up aspects of the work environment then the importance of these regressors should vanish once we include the firm level SOM or firm fixed effects.

The results documented in Table 9 support our *a priori* expectations. That is, overall we find that the firm SOM (the workplace environment) matters independently for women but the effect of the occupation SOM, though reduced, remains important. Moreover, adding PBB still explains some of the occupation SOM effects. In fact, PBB also explains some of the firm SOM effects. Replacing the SOM in the firm by firm fixed effects alters the estimates for the occupation SOM little but has some impact on the PBB coefficients. Overall, our previous conclusions carry over to the WERS analysis: PBB captures an important aspect of female satisfaction, which shows up in the occupation SOM effects when PBB is not separately controlled. The WERS results bolster the case that this relates to occupational content rather than work environment.

It is interesting, however, that the firm SOM has a sizeable negative relationship with female satisfaction (and a somewhat smaller but also significant one with male satisfaction). Workplaces with more women seem to be more pleasant workplaces for both genders. It is impossible to tell from these regressions whether this is due to the presence of the women per se, some underlying attributes of the workplaces which attract the women to work there in the first place, or changes in the environment, culture, and organization which take place at workplaces once they employ more women. There are variables in the WERS, which should in principle capture aspects of the work environment more directly (like whether there have been cases of sexual harassment), but none of these variables affects the more direct revealed preferences captured by the SOM variables.[19]

Because the firm SOM effects are large, and point to other sources of female satisfaction than our earlier results, we explore these issues further by returning to our earlier data sets, the NLSY and BHPS, in Table 10.[20] Here, in addition to the average share of males in the occupation, which we have analyzed throughout, we now add time variation in the share of males. Women have been entering some occupations during our sample periods but not others, so there is some variation in the SOM over time. This variation is unlikely related to the nature of the work, which we have argued the average SOM in the occupation might reflect. Most likely, changes in the SOM over time also pick up changes in the work environment.

[19] See Appendix B for full details of this analysis and the overall results.

[20] We do not have a large enough data set to calculate year-by-year averages of the SOM in an occupation precisely enough for Russia, so we do not use the RLMS in this analysis.

In Table 10, we find negative and consistently large effects of the time varying share of males on female satisfaction in both countries. These estimates are identified from job stayers and the changes in who works in their occupations over time. The results suggests that women get happier in their jobs as more female colleagues enter their line of work. Standard errors are larger than for the average SOM in the occupation because there is less relevant variation in the changes of the SOM over time. Effects for men are less consistent and not significant. Interestingly, for the BHPS, where we have separate variables for overall satisfaction and satisfaction with the work itself, we find large effects for the overall satisfaction of females, and much more muted effects for satisfaction with work itself. This is the opposite pattern from the average SOM. Hence, these results are again consistent with the idea that the average SOM picks up elements of the content of work, while the time varying SOM picks up elements of the work environment.

Discussion

Stigler and Becker (1977) have famously cautioned economists against relying on variation in preferences to explain economic outcomes, suggesting that the most worthwhile focus is on the comparative statics induced by variation in constraints. The literature on differences in labor market outcomes and behaviors between men and women has indeed for a long time adopted this approach, and studied the impact of discrimination, human capital investments, and labor supply. Less than two decades ago, Altonji and Blank (1999) devoted two paragraphs of their handbook chapter on race and gender to differences in preferences before moving on to the traditional constraint based explanations.

But stubborn differences in male and female pay and occupational segregation persist while many of the constraints faced by women in the workplace seem to have diminished (which does not mean that these constraints are all gone). At the same time, economists have grown more relaxed about thinking about differences in tastes. The handbook chapter by Bertrand (2010), a mere ten years after Altonji and Blank, focuses almost entirely on explanations based on differences in psychological traits between men and women, as well as gender identity. A powerful form in which such psychological differences manifest themselves is in different tastes of men and women for the content of the work they do. We argue that economists should be open-minded towards this explanation, and subject it to scrutiny.

Here we have offered an initial attempt at this by analyzing the differences in job satisfaction of women in male and female dominated jobs. We find that women are less satisfied in occupations with high shares of males. Adding variables which might proxy job content ('people', 'brains,' and 'brawn') to our main regressions, we find that women are happier in the jobs relatively low in 'brawn' and relatively high in 'people' and 'brains' while at the same time the explanatory power of the SOM variable is reduced. In fact, in the US and Russia the PBB variables are the only ones which seem to matter. Although regression models for job mobility paint a somewhat different picture we have argued that the job satisfaction results may be more indicative of women's actual tastes for job content.

A subsidiary finding of our analysis is that elements of the work environment or organization also seem to matter greatly and independently, without affecting our conclusion about the importance of preferences. Again, our measures don't pinpoint this precisely but the idea that the firm level SOM and time variation in the SOM captures the work environment is a simple explanation which would be consistent with our results. Such environmental factors may be another important driving force of the occupation choices of women.

We have offered some descriptive evidence consistent with the idea that Rosie is able and willing to be a riveter if asked to do so but it is not her preferred line of work. Economists should explore the possibility that gender specific tastes matter for occupational choice because the policy implications of taste based sorting into occupations differ substantially from explanations based on constraints, or the influence of the work environment. Constraints suggest that policies should remove these constraints, for example through anti-discrimination legislation or mandates to employers to allow workers more flexibility when combining family responsibilities with work. One goal of these policies is to close gender gaps in occupational allocations, which in turn should help close the gender pay gap. On the other hand, if women do not want to do the same jobs as men there seems less sense in trying to tackle occupational segregation. In particular, policies like quotas seem to make little sense. Instead, governments may want to look towards increasing the demand in the types of jobs often done by women directly in order to close the gender pay gap. Tackling detrimental work environments may yet require completely different interventions. Understanding more about these rather underexplored sources of the behavior of women in the labor market is an important ingredient in designing more effective policies in the future.

References

Altonji, Joseph G. and Rebecca M. Blank (1999). "Race and Gender in the Labor Market." In: Orley Ashenfelter and David Card (eds.) *Handbook of Labor Economics*, volume 3C, Amsterdam: Elsevier, 3143-3259.

Autor, David H. and David Dorn (2013). "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553-1597.

Baker, Michael and Corneslon Kirsten (2016). "Gender Based Occupational Segregation and Sex Differences in Sensory, Motor and Spatial Aptitudes." *NBER Working Paper Series.* No. 22248.

Bayard, Kimberly, Judith Hellerstein, David Neumark, and Kenneth Troske (2003). "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data." *Journal of Labor Economics,* 21(4): 887-922.

Beaudry, Paul, and Ethan Lewis (2014). "Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence from 1980-2000." *American Economic Journal: Applied Economics*, 6(2): 178-194.

Belkin, Lisa (2003). The Opt-Out Revolution. *The New York Times,* October 26: 42.

Bertrand, Marianne. (2010). "New Perspectives on Gender." In: Orley Ashenfelter and David Card (eds), *Handbook of Labor Economics*, volume 4B, Amsterdam: Elsevier, 1545-1592.

Bielby, William and J. Baron (1984). "A Woman's Place Is with Other Women: Sex Segregation within Organizations." in Reskin, Barbara F. (ed.), *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, National Academies Press.

Blau, Francine D. and Lawrence M. Kahn (2000). "Gender Differences in Pay." *The Journal of Economic Perspectives,* 14(4): 75-99.

Blau, Francine D. and Lawrence M. Kahn (2006). "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence." *Industrial and Labor Relations Review* 60(1): 45-65.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2011). "Robust Inference with Multi-Way Clustering." *Journal of Business and Economic Statistics*, 29 (2): 238-249.

Clark, Andrew E. and Andrew Oswald (1996). "Satisfaction and Comparison Income." *Journal of Public Economics,* 61(3): 359-381.

Clark, Andrew E. (1996). "Job Satisfaction in Britain." *British Journal of Industrial Relations,* 34(2): 189-217.

Croson, Rachel and Uri Gneezy (2009). "Gender Differences in Preferences". *Journal of Economic Literature* Vol. 47, No. 2 (Jun., 2009), pp. 448-474.

Dorn David (2009). “Essays on Inequality, Spatial Interaction, and the Demand for Skills.” *Dissertation,* University of St. Gallen no. 3613.

Fortin, Nicole M. (2008). “The Gender Wage Gap Among Young Adults in the United States – The Importance of Money versus People”, *The Journal of Human Resources* 43(4): 884-918.

Goldin, Claudia (2014). “A Grand Gender Convergence: Its Last Chapter.” *American Economic Review*, 104(4): 1091-1119.

Goldin, Claudia and Lawrence F. Katz. (2008). “Transitions: Career and Family Life Cycles of the Educational Elite.” *American Economic Review*, 98(2): 363-369.

Goldin, Claudia and Lawrence F. Katz. (2011). “The Cost of Work Place Flexibility of High Powered Professionals.” *The Annals of the American Academy*, 638 (1): 45-64.

Goldin, Claudia and Lawrence F. Katz (2016). “The Most Egalitarian Profession: Pharmacy and the Evolution of a Family–Friendly Occupation.” *Journal of Labor Economics*, 34(3): 705-746.

Gorsuch, Richard L. (1983). *Factor Analysis*. Second edition, Hillsdale: Lawrence Erlbaum Associates.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini (2003). “Performance in Competitive Environments: Gender Differences.” *The Quarterly Journal of Economics*, 118 (3): 1049-1074.

Heckman, James, Rodrigo Pinto and Peter Savelyev (2013). “Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes.” *American Economic Review*, 103(6): 2052-86.

Hochschild, Arlie with Anne Machung (1989). *The Second Shift: Working Parents and the Revolution at Home*. New York: Viking.

Hunt, Jenny (2016) “Why Do Women Leave Science and Engineering?” *Industrial and Labor Relations Review*, 69(1), 199-226.

Macpherson, David A. and Barry. T. Hirsch (1995). “Wages and Gender Composition: Why Do Women's Jobs Pay Less?” *Journal of Labor Economics,* 13(3): 426-471.

Manning, Alan (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton: Princeton University Press.

Manning, Alan and Joanna Swaffield (2008). “The Gender Gap in Early-Career Wage Growth.” *Economic Journal,* 118(530): 987-1024.

Muriel Niederle and Lise Vesterlund (2007). “Do Women Shy Away from Competition? Do Men Compete Too Much?” *The Quarterly Journal of Economics*, 122(3): 1067-1101.

Pinker, Susan (2008). *The Sexual Paradox: Troubled Boys, Gifted Girls and the Real Difference Between the Sexes*. New York: Macmillan.

Stigler, George J. and Gary S. Becker (1977). "De Gustibus Non Est Disputandum." *American Economic Review*, 67(2): 76-90.

Taylor, Marcia F. (ed.) with Brice John, Nick Buck, and Elaine Prentice-Lane (2010). *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.

Thompson, Bruce. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Stone, Pamela (2007). *Opting Out: Why Women Really Quit Careers and Head Home*. Berkeley, CA: University of California Press.

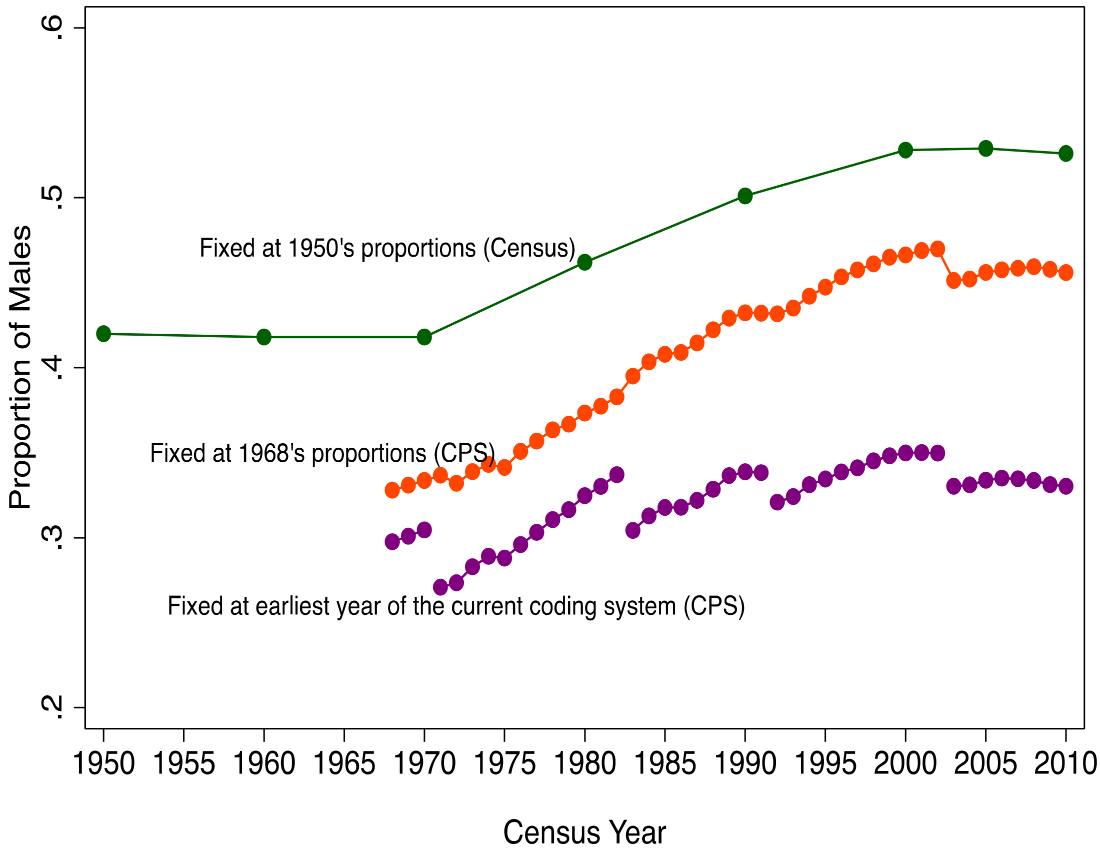
Usui, Emiko (2008). "Job Satisfaction and the Gender Composition of Jobs." *Economics Letters*, 99(1): 23-26.

van Praag, Bernard M.S. and Ada Ferrer-i-Carbonell (2008). *Happiness Quantified: A Satisfaction Calculus Approach*. Revised edition, Oxford: Oxford University Press.

Weinberg, Bruce (2000). "Computer Use and the Demand for Female Workers", *Industrial and Labor Relations Review*, 53(2): 290-308.

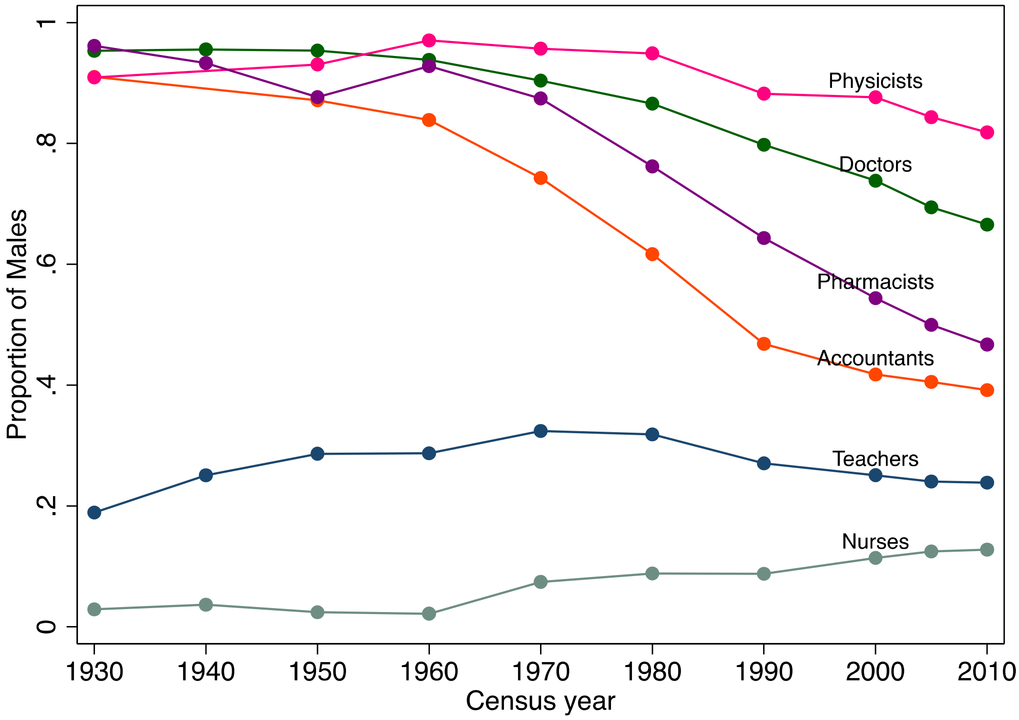
Wiswall, Matthew, and Basit Zafar (2016) "Preference for the Workplace, Human Capital, and Gender," *NBER Working Paper Series* No. 22173

Figure 1: The Share of Males in Female Jobs



Notes: The lines in this graph show the share of males (SOM) in the occupations in which females work in a particular year in the US. The top line uses Census data and is based on the SOM in each occupation in 1950 using the IPUMS 1950 consistent occupation code. The other lines use annual CPS data. In the second line, SOM in an occupation is calculated based on the 1968 data. The bottom line uses the current occupation codes and fixes the SOM in the year the current code was first introduced. The line is broken whenever a new set of occupation codes comes into use.

Figure 2: Trends in the Share of Males in Selected White Collar Jobs



Notes: This graph shows the share of males in selected white-collar occupations in the US Census.

Table 1: Basic Job Satisfaction Regressions

Occupation averages	Sample and Dependent Variable: US – NLSY Overall Job Satisfaction		Britain – BHPS Overall Job Satisfaction		Britain – BHPS Satisfaction with Work Itself		Russia – RLMS Overall Job Satisfaction	
	Females	Males	Females	Males	Females	Males	Females	Males
Share of Males	-0.209	-0.013	-0.124	-0.013	-0.283	-0.032	-0.130	-0.041
	(0.058)	(0.067)	(0.018)	(0.021)	(0.017)	(0.026)	(0.009)	(0.058)
Log of Wage	-0.003	-0.014	-0.062	0.058	-0.016	0.109	0.115	0.109
	(0.023)	(0.021)	(0.021)	(0.027)	(0.067)	(0.038)	(0.040)	(0.038)
Hours/100	0.294	0.686	-0.525	0.845	0.202	0.838	0.720	-0.003
	(0.446)	(0.423)	(0.077)	(0.085)	(0.227)	(0.086)	(0.795)	(0.387)
Degree holders	0.352	0.449	-0.113	0.067	0.012	0.083	0.628	0.447
	(0.066)	(0.080)	(0.049)	(0.050)	(0.058)	(0.059)	(0.117)	(0.057)
Age/100	1.000	1.178	1.249	1.178	0.760	1.088	-0.857	0.271
	(0.446)	(0.397)	(0.122)	(0.075)	(0.117)	(0.078)	(0.558)	(0.443)
Number of Observations	75672	80648	48141	43365	48141	43365	35443	27117

Notes: All regressions also include age and age squared of the individual, as well as time and area effects. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using ivreg2.

Table 2: Job Satisfaction and Mobility Regressions

	Samples					
	US – NLSY		Britain – BHPS		Russia – RLMS	
	Females	Males	Females	Males	Females	Males
	Dependent Variable: Overall Job Satisfaction					
Share of Males	-0.209	-0.013	-0.124	-0.013	-0.130	-0.041
	(0.058)	(0.067)	(0.018)	(0.021)	(0.049)	(0.058)
Number of Observations	75672	80648	48141	43365	35443	27117
	Dependent Variable: Satisfaction with Work Itself					
Share of Males			-0.283	-0.032		
			(0.017)	(0.026)		
Number of Observations			48141	43365		
	Dependent Variable: Stayers					
Share of Males	-0.228	0.082	-0.276	0.130	-0.369	0.141
	(0.052)	(0.035)	(0.070)	(0.013)	(0.061)	(0.056)
Number of Observations	76375	81144	35613	32100	23449	16792

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, as well as time and area effects. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using ivreg2.

Table 3: Individual Fixed Effects Regressions

	Samples					
	US – NLSY		Britain – BHPS		Russia – RLMS	
	Females	Males	Females	Males	Females	Males
	Dependent Variable: Overall Job Satisfaction					
Share of Males	-0.117	-0.027	-0.245	-0.076	-0.116	-0.020
	(0.029)	(0.028)	(0.031)	(0.007)	(0.053)	(0.030)
Number of Observations	75672	80648	48141	43365	35443	27117
	Dependent Variable: Satisfaction with Work Itself					
Share of Males			-0.325	-0.041		
			(0.007)	(0.013)		
Number of Observations			48141	43365		
	Dependent Variable: Stayers					
Share of Males	-0.188	0.069	-0.281	0.086	-0.281	0.072
	(0.013)	(0.013)	(0.027)	(0.042)	(0.042)	(0.038)
Number of Observations	76375	81144	35613	32100	23449	16792

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table 4: High and Low Ranked Occupations According to the Content of the Work (US)

Rank	People	Brains	Brawn
		10 Highest Ranked Occupations	
1	Occupational Therapist (99)	Chemical engineers (48)	Explosive Workers (615)
2	Sales Supervisors and Proprietors (243)	Aerospace engineers (44)	Miners (616)
3	Therapists (105)	Medical Scientists (83)	Fire Fighting & Inspection Occs (417)
4	Licensed practical nurses (207)	Geologists (75)	Structural Metal Workers (597)
5	Managers of medicine and health occs (15)	Financial Managers (7)	Other Mining Occupations (617)
6	Buyers wholesale and trade (29)	Physicists and astronomers (69)	Excavating & Loading Machine Operators (853)
7	Social Workers (174)	Chemists (73)	Heavy Equipment Mechanics (516)
8	Urban and Regional Planners (173)	Urban and Regional Planners (173)	Water & sewage treatment plant operators (694)
9	Child Care Workers (468)	Petrol mining & geological engineers (47)	Millwrights (544)
10	Business and Promotion Agents (34)	Mechanical Engineers (57)	Roofers and Slaters (595)
		10 Lowest Ranked Occupations	
1	Statistical Clerks (386)	Sales demonstrators and models (283)	Insurance Underwriters (23)
2	Physicists and astronomers (69)	Excavating & loading machine operators (853)	Interviewers, enumerators, and surveyors (316)
3	Motion picture projection (467)	Clothing and machine operators (747)	Lawyers and Judges (178)
4	Mathematicians and statisticians (68)	Garbage collectors (875)	Telephone operators (348)
5	Data Entry Keyers (385)	Personal service occupations, n.e.c (469)	Art/entertainment performers (194)
6	Biological Scientists (78)	Machine operators (779)	Payroll and timekeeping clerks (338)
7	Surveryors & mapping scientists (467)	Packers and Packagers by hand (888)	Clergy and religious workers (176)
8	Actuaries (66)	Athletes, sports instructors and officials (199)	Advertising and related sales jobs (256)
9	Office machine operators (347)	Messengers (357)	Economists, market and survey researchers (166)
10	Paper folding machine operators (765)	Helpers, constructions (865)	Other Financial Specialists (25)

Table 5: Factor Scores for Selected Occupations (US)

Occupation	1930 SOM	SOM	Factor Scores People	Factor Scores Brains	Factor Scores Brawn
Electricians (575)	0.993	0.979	0.464	0.382	1.476
Miners (686)	0.997	0.972	-1.031	-0.455	2.194
Chemical Engineers (48)	1.000	0.867	-0.413	3.023	0.926
Architects (43)	0.979	0.784	0.499	1.436	-0.099
Physicians (84)	0.944	0.722	0.682	1.956	-0.908
Butchers (686)	0.992	0.746	-0.039	-1.372	0.562
Mathematicians and Statisticians (66)	N/A	0.640	-1.966	1.087	-1.201
Financial managers (7)	N/A	0.572	-0.163	2.528	-0.973
Economists, market and survey researchers (166)	0.810	0.510	-1.379	1.177	-1.333
Bartenders (434)	0.960	0.434	1.503	-0.621	-0.052
Accountants and auditors (23)	0.912	0.423	-0.335	1.833	-1.119
Social Workers (174)	0.265	0.251	1.601	0.954	-0.848
Primary School Teachers (156)	0.188	0.165	0.606	0.530	-0.862
Registered Nurses (95)	0.025	0.068	1.311	1.234	0.787

Notes: SOM are the share of males in an occupation based on CPS data from 1983-2012 to match the NLSY data. 1930 SOMs calculated using the 1930 census.

Table 6: The Relationship Between the Share of Males and People, Brains, and Brawn

	Samples		
	US – CPS	Britain – LFS	Russia – RLMS
People	-0.025	-0.035	-0.084
	(0.020)	(0.012)	(0.031)
Brains	-0.009	0.000	0.001
	(0.020)	(0.015)	(0.031)
Brawn	0.152	0.144	0.186
	(0.025)	(0.012)	(0.032)
Number of Observations	5,848, 990	4,023, 894	268, 635

Notes: All regressions also include the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, as well as time and area effects. Standard errors are clustered by occupation.

Table 7: Job Satisfaction Regressions Including People, Brains, and Brawn

	Sample and Dependent Variable							
	US – NLSY Overall Job Satisfaction		Britain – BHPS Overall Job Satisfaction		Britain – BHPS Satisf. with Work Itself		Russia – RLMS Overall Job Satisfaction	
	Females	Males	Females	Males	Females	Males	Female	Males
Baseline Regressions								
Share of Males	-0.117	-0.027	-0.245	-0.076	-0.325	-0.041	-0.116	-0.020
	(0.029)	(0.028)	(0.031)	(0.007)	(0.007)	(0.013)	(0.053)	(0.030)
Number of Observations	75672	80648	48141	43365	48141	43365	35443	27117
Including People, Brains and Brawn								
Share of Males	-0.029	0.011	-0.195	-0.041	-0.235	-0.009	-0.050	-0.023
	(0.033)	(0.034)	(0.015)	(0.013)	(0.006)	(0.015)	(0.060)	(0.030)
People	0.027	0.001	0.011	0.013	0.065	0.018	0.012	0.003
	(0.006)	(0.007)	(0.008)	(0.012)	(0.008)	(0.002)	(0.014)	(0.015)
Brains	0.052	0.005	0.016	-0.004	0.035	-0.012	0.035	0.024
	(0.008)	(0.007)	(0.009)	(0.002)	(0.011)	(0.015)	(0.013)	(0.018)
Brawn	-0.021	-0.004	-0.034	-0.014	-0.039	-0.000	-0.053	-0.007
	(0.010)	(0.009)	(0.012)	(0.002)	(0.006)	(0.006)	(0.017)	(0.019)
Number of Observations	75672	80648	48141	43365	48141	43365	35443	27117
Hausman Test	9.91	3.35	6.64	2.66	7.75	2.07	7.73	0.07
P-value	0.002	0.068	0.010	0.103	0.005	0.151	0.005	0.795

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table 8: Mobility Regressions Including People, Brains, and Brawn

	Samples					
	US – NLSY		Britain – BHPS		Russia – RLMS	
	Females	Males	Females	Males	Females	Males
	Baseline Regressions					
Share of Males	-0.188	0.069	-0.281	0.086	-0.281	0.072
	(0.013)	(0.013)	(0.027)	(0.042)	(0.042)	(0.038)
Number of Observations	76375	81144	35613	32100	23449	16792
	Including People, Brains, and Brawn					
Share of Males	-0.193	0.117	-0.263	0.120	-0.261	0.053
	(0.016)	(0.017)	(0.035)	(0.048)	(0.050)	(0.045)
People	0.018	0.016	0.016	0.001	0.027	0.016
	(0.003)	(0.003)	(0.001)	(0.004)	(0.013)	(0.016)
Brains	0.024	0.015	-0.014	0.006	0.015	0.010
	(0.004)	(0.004)	(0.001)	(0.007)	(0.018)	(0.017)
Brawn	-0.023	-0.008	-0.012	-0.013	-0.012	0.023
	(0.005)	(0.005)	(0.003)	(0.002)	(0.016)	(0.015)
Number of Observations	76375	81144	35613	32100	23449	16762
Hausman Test	0.62	5.44	0.22	1.77	0.20	1.23
P-value	0.431	0.012	0.640	0.184	0.654	0.268

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2

Table 9: Job Satisfaction Regressions in the WERS

	Samples			
	Females	Males	Females	Males
	Baseline			
Share of Males	-0.275	-0.059	-0.172	-0.079
	(0.043)	(0.047)	(0.061)	(0.059)
People			0.144	0.074
			(0.011)	(0.009)
Brains			0.006	-0.021
			(0.012)	(0.010)
Brawn			-0.039	-0.021
			(0.012)	(0.010)
	Adding Firm Share of Males			
Share of Males (occupation)	-0.160	0.086	-0.117	-0.020
	(0.042)	(0.051)	(0.059)	(0.062)
Share of Males (firm)	-0.305	-0.154	-0.182	-0.119
	(0.034)	(0.039)	(0.035)	(0.041)
People			0.126	0.072
			(0.011)	(0.009)
Brains			0.010	-0.007
			(0.012)	(0.010)
Brawn			-0.034	-0.020
			(0.012)	(0.010)
	Adding Firm Fixed Effects			
Share of Males (occupation)	-0.151	-0.093	-0.104	0.007
	(0.046)	(0.066)	(0.062)	(0.074)
People			0.095	0.049
			(0.012)	(0.011)
Brains			0.037	0.026
			(0.012)	(0.012)
Brawn			-0.010	0.032
			(0.013)	(0.013)
Number of Observations	18299	16655	18299	16655

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, along with time effects. Standard errors are two-way clustered by firm and worker's occupation and shown in parentheses. Models are estimated using ivreg2.

Table 10: Job Satisfaction Regressions with Time Varying Share of Males

	Sample and Dependent Variable			
	Females	Males	Females	Males
NLSY, Dependent Variable: Overall Job Satisfaction				
Average Share of Males	-0.117	-0.027	-0.110	0.006
	(0.029)	(0.028)	(0.030)	(0.037)
Time Varying Share of Males			-0.429	0.209
			(0.108)	(0.111)
Number of Observations	75672	80648	75672	80648
BHPS, Dependent Variable: Overall Job Satisfaction				
Average Share of Males	-0.245	-0.076	-0.240	-0.090
	(0.031)	(0.007)	(0.030)	(0.021)
Time Varying Share of Males			-0.392	-0.119
			(0.161)	(0.187)
Number of Observations	48141	43365	48141	43365
BHPS , Dependent Variable: Satisfaction with Work Itself				
Average Share of Males	-0.325	-0.041	-0.304	-0.050
	(0.007)	(0.013)	(0.020)	(0.009)
Time Varying Share of Males			-0.102	0.007
			(0.116)	(0.333)
Number of Observations	48141	43365	48141	43365

Notes: Time Varying Share of Males is the difference between the share of males in the current year and its average within a sub-period with consistent occupation codes. The NLSY analysis relates to 1983-2012. The BHPS analysis pertains to 1991-2010. All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, the deviation of the same variables from the mean over time, and time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Does Rosie Like Riveting?
Male and Female Occupational Choices-
Appendix, Not for Publication

Grace Lordan and Jörn-Steffen Pischke
LSE

Appendix A: Robustness to the inclusion of personal characteristics and flexibility proxies

This appendix considers the addition of personal characteristics to the baseline fixed effects regressions in Table 3. In particular, we add the log of the own wage, own hours, number of children, and a dummy indicating whether a person is married to the regressions.

We also add measures of hours flexibility to the US and British regressions. Ideally, we would like a measure which reflects how easy it is to combine family and career in a particular occupation. We approximate this in various ways. First, we consider data from the American Time Use Survey (ATUS) from 2003-2013. This survey documents the time a person spends doing various activities in a particular day and includes three digit occupation codes.[21] We construct a measure of whether women are caring for their children during a standard male workday specific to their own occupation. We define a standard male day as the interval between: 1. the mode of the male start time for a particular occupation and 2. the mode of the male end time given the start time in 1.[22] We then calculate the average hours that a woman provides care as defined by the ATUS codes 30101 to 49999 during these hours defined by a standard male work day. This measure of flexibility captures whether an occupation accommodates females combining work and home. Additionally, the ATUS provides codes that indicate if an individual was participating in social events at work during their diary day. Such events may impinge on females trying to juggle family and career, and cultivate an ‘old boys’ club. To capture this, we calculate the average hours spent by males in an occupation in activities given by ATUS codes 50201 to 50289, excluding 50204 but including 59999. We then match these variables into the NLSY data.

For the UK, the LFS began asking specific question on flexibility in the second quarter of 2004. The same questions have since been asked every second quarter. Based on data from 2004-2013, we calculate the occupation specific proportion of

21. Specifically, the ATUS uses Census occupation codes, so we crosswalk as described in the US data section.
22. Sensitivity analysis highlights that fixing the start time given the end time does not change our results.

individuals who currently work 1. a flexible time schedule and 2. annualized hours. In an annual hours system, an employee works a given number of hours in a year, but with a certain degree of flexibility about when those hours are worked. Normally, a period of regular hours forms the core of the arrangement, with the remaining time left unallocated. We then match these variables into the BHPS data.

We do not have any variables capturing flexibility for Russia, so we only add the other covariates to the RLMS regressions. Results for these regressions are displayed in Tables A.1 – A.7.

Table A.1 US Regressions for Overall Job Satisfaction with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.127	-0.034	-0.135	-0.112
	(0.031)	(0.031)	(0.031)	(0.041)
Log of Wage	0.008	-0.014	0.009	0.317
	(0.010)	(0.010)	(0.010)	(0.051)
Hours/100	0.396	0.677	0.571	0.010
	(0.167)	(0.185)	(0.167)	(0.309)
College Graduates	0.293	0.407	0.278	0.156
	(0.042)	(0.060)	(0.042)	(0.077)
Age/100	0.717	0.310	0.680	-0.468
	(0.182)	(0.210)	(0.182)	(0.322)
Log Own Wage			0.007	0.036
			(0.006)	(0.007)
Own Hours			-0.003	0.002
			(0.001)	(0.001)
College Graduate			0.080	-0.065
			(0.040)	(0.031)
No of Children			0.023	0.007
			(0.009)	(0.007)
Married			0.007	-0.026
			(0.014)	(0.014)
Time spent Caring			0.060	0.012
			(0.016)	(0.016)
Time spent Socializing			0.151	0.165
			(0.137)	(0.131)
Number of Observations	67852	54980	67852	54980

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.2 British Regression for Overall Job Satisfaction with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.244	-0.081	-0.257	-0.076
	(0.031)	(0.009)	(0.037)	(0.007)
Log of Wage	0.138	0.122	0.132	0.083
	(0.000)	(0.055)	(0.000)	(0.049)
Hours/100	0.107	0.152	0.190	0.158
	(0.156)	(0.040)	(0.191)	(0.066)
College Graduates	0.205	0.032	0.184	0.003
	(0.094)	(0.059)	(0.084)	(0.063)
Age/100	0.517	0.427	0.581	0.509
	(0.089)	(0.241)	(0.101)	(0.206)
Log of own Wage			0.061	0.169
			(0.003)	(0.017)
Own Hours			-0.005	-0.005
			(0.001)	(0.000)
College Graduate			-0.019	0.114
			(0.031)	(0.055)
Married			0.032	0.009
			(0.003)	(0.023)
No. of Children			0.022	0.007
			(0.011)	(0.009)
Flexi Time			-0.009	0.220
			(0.000)	(0.031)
Annualized Hours			-0.949	0.402
			(0.106)	(0.113)
Number of Observations	42297	38420	42297	38420

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.3 British Regression for Satisfaction with Work Itself with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.320	-0.043	-0.328	-0.044
	(0.010)	(0.016)	(0.011)	(0.012)
Log of Wage	0.125	0.106	0.125	0.074
	(0.023)	(0.055)	(0.020)	(0.047)
Hours/100	0.563	0.205	0.597	0.209
	(0.157)	(0.019)	(0.200)	(0.016)
College Graduates	0.276	0.218	0.256	0.158
	(0.077)	(0.054)	(0.080)	(0.051)
Age/100	0.141	0.498	0.200	0.676
	(0.195)	(0.323)	(0.229)	(0.316)
Log of Own Income			0.019	0.077
			(0.020)	(0.002)
Own Hours			-0.002	-0.003
			(0.001)	(0.001)
College Graduate			0.016	0.061
			(0.004)	(0.022)
Married			-0.007	-0.011
			(0.018)	(0.016)
No. of Children			0.022	0.007
			(0.001)	(0.009)
Flexi-Time			-0.029	0.264
			(0.042)	(0.015)
Annualized Hours			-0.840	-1.326
			(0.416)	(0.053)
Number of Observations	42297	38420	42297	38420

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.4 Russian Regression for Overall Job Satisfaction with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.145	-0.066	-0.159	-0.090
	(0.060)	(0.051)	(0.059)	(0.039)
Log of Wage	0.069	0.034	0.057	0.020
	(0.026)	(0.020)	(0.026)	(0.020)
Hours/100	0.739	0.188	-0.029	0.155
	(0.380)	(0.246)	(0.033)	(0.252)
College Graduates	0.270	0.341	0.067	0.327
	(0.071)	(0.066)	(0.046)	(0.066)
Age/100	-0.546	0.194	-0.455	0.188
	(0.324)	(0.319)	(0.314)	(0.315)
Log Own Wages			0.187	0.208
			(0.013)	(0.017)
Own Hours			-0.029	-0.023
			(0.033)	(0.033)
College Graduate			0.067	0.070
			(0.046)	(0.064)
Married			0.024	0.028
			(0.030)	(0.028)
Number of Kids			0.010	-0.008
			(0.029)	(0.021)
Number of Observations	28282	22897	28282	22897

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.5 US Regression for Stayers with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.234	0.075	-0.227	0.078
	(0.016)	(0.017)	(0.016)	(0.025)
Log of Wage	0.000	-0.003	0.006	0.014
	(0.005)	(0.006)	(0.005)	(0.027)
Hours/100	0.929	0.664	0.805	0.240
	(0.090)	(0.111)	(0.090)	(0.169)
College Graduates	0.138	0.238	0.136	0.161
	(0.027)	(0.034)	(0.027)	(0.048)
Age/100	-0.093	-0.126	-0.012	0.008
	(0.099)	(0.112)	(0.098)	(0.189)
Log of Own Wage			0.032	0.020
			(0.003)	(0.003)
Own Hours			0.000	0.000
			(0.000)	(0.000)
College Graduate			0.037	-0.009
			(0.028)	(0.032)
Married			0.003	0.020
			(0.008)	(0.009)
Number of Children			0.005	-0.002
			(0.006)	(0.005)
Time Spent Caring			0.010	0.003
			(0.009)	(0.008)
Time Spent Socializing			-0.380	0.026
			(0.082)	(0.069)
Number of Observations	56247	47620	56247	47620

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.6 British Regression for Stayers with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.191	0.048	-0.189	0.050
	(0.028)	(0.023)	(0.027)	(0.023)
Log of Wage	0.082	-0.050	0.085	-0.049
	(0.002)	(0.031)	(0.006)	(0.031)
Hours/100	0.201	0.185	0.172	0.171
	(0.026)	(0.028)	(0.011)	(0.020)
College Graduates	0.021	0.113	0.029	0.130
	(0.012)	(0.036)	(0.003)	(0.043)
Age/100	-0.088	0.352	-0.104	0.304
	(0.142)	(0.088)	(0.190)	(0.096)
Log of Own Income			0.008	0.035
			(0.002)	(0.008)
Own Hours			0.001	0.000
			(0.000)	(0.000)
College Graduate			0.046	-0.004
			(0.013)	(0.019)
Married			-0.057	0.000
			(0.003)	(0.003)
No. of Children			0.002	-0.007
			(0.006)	(0.002)
Flexi Time			-0.041	-0.055
			(0.066)	(0.008)
Annualized Hours			0.318	0.412
			(0.526)	(0.168)
Number of Observations	31011	27936	31011	27936

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table A.7 Russian Regression for Stayers with Own Characteristics and Flexibility Conditions

	Samples			
	Females	Males	Females	Males
Share of Males	-0.343	0.028	-0.350	0.019
	(0.072)	(0.051)	(0.071)	(0.054)
Log of Wage	0.129	0.072	0.125	0.070
	(0.034)	(0.032)	(0.034)	(0.033)
Hours/100	-0.357	0.138	-0.545	0.083
	(0.430)	(0.310)	(0.415)	(0.327)
College Graduates	-0.332	-0.068	-0.343	-0.076
	(0.067)	(0.066)	(0.067)	(0.070)
Age/100	0.073	-0.000	0.067	0.183
	(0.413)	(0.315)	(0.397)	(0.329)
Own Hours			0.047	0.024
			(0.030)	(0.026)
College Graduate			0.055	0.009
			(0.040)	(0.064)
Married			0.008	0.062
			(0.024)	(0.031)
Number of Kids			-0.035	0.005
			(0.022)	(0.019)
Number of Observations	10546	8354	10546	8354

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Appendix B: WERS analysis including aspects of the work environment

The WERS data include questions to the firm representative that gauge the type of grievances that have been raised in the past year in the firm. Of interest are the questions that ask whether any sexual harassment or sexual discrimination grievances were raised. These questions were only asked in 2004. From the responses we create a dummy variable that is equal to 1 if grievances in either of these categories were raised and zero otherwise. We then replicate Table 9 from the main text including these two variables. We are interested in whether the inclusion of these indicators changes significantly the coefficients on the occupation SOM or firm SOM. Given these variables are defined at the level of the firm, the model which includes firm fixed effects is not useful here. The results from this robustness analysis are documented in Tables B.1 and B.2 below.

Table B.1: Basic Job Satisfaction Regressions in the WERS with Sexual Harassment Variables

	Samples							
	Females	Males	Females	Males	Females	Males	Females	Males
Share of Males (occupation)	-0.229	-0.065	-0.141	-0.010	-0.224	-0.064	-0.142	-0.009
	(0.070)	(0.067)	(0.095)	(0.084)	(0.070)	(0.067)	(0.095)	(0.084)
People			0.146	0.075			0.144	0.075
			(0.016)	(0.013)			(0.016)	(0.013)
Brains			0.009	-0.025			0.011	-0.024
			(0.020)	(0.015)			(0.020)	(0.015)
Brawn			-0.030	-0.005			-0.031	-0.005
			(0.018)	(0.015)			(0.018)	(0.015)
Sexual Harassment					-0.099	-0.058	-0.090	-0.061
					(0.056)	(0.059)	(0.057)	(0.059)
Sexual Discrimination					-0.045	-0.046	-0.040	-0.037
					(0.036)	(0.039)	(0.036)	(0.037)
Number of Observations	11800	10265	11800	10265	11800	10265	11800	10265

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, along with time effects. Standard errors are two-way clustered by firm and their occupation and shown in parentheses. Models are estimated using ivreg2

Table B.2: Job Satisfaction Regressions in the WERS with Firm Share of Males and Sexual Harassment Variables

	Samples							
	Females	Males	Females	Males	Females	Males	Females	Males
Share of Males (occupation)	-0.150	0.095	-0.081	-0.034	-0.155	0.112	-0.083	-0.033
	(0.048)	(0.055)	(0.050)	(0.055)	(0.008)	(0.068)	(0.050)	(0.055)
Share of Males (firm)	-0.292	-0.095	-0.182	-0.119	-0.285	-0.099	-0.173	-0.111
	(0.048)	(0.055)	(0.035)	(0.041)	(0.047)	(0.055)	(0.050)	(0.047)
People			0.129	0.074			0.129	0.073
			(0.017)	(0.013)			(0.017)	(0.013)
Brains			0.009	-0.025			0.011	-0.024
			(0.020)	(0.015)			(0.020)	(0.015)
Brawn			-0.028	-0.015			-0.029	-0.003
			(0.018)	(0.015)			(0.018)	(0.015)
Sexual Harassment					-0.093	-0.062	-0.088	-0.063
					(0.057)	(0.059)	(0.057)	(0.059)
Sexual Discrimination					-0.026	-0.047	-0.029	-0.037
					(0.036)	(0.039)	(0.036)	(0.037)
Number of Observations	11800	10265	11800	10265	11800	10265	11800	10265

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, along with time effects. Standard errors are two-way clustered by firm and their occupation and shown in parentheses. Models are estimated using ivreg2

Appendix C: Construction of Latent Factors from ONET

In order to create latent measures to capture job content we use the O*NET database version 5. ONET provides a diverse set of information on occupational attributes, requirements, and characteristics of the workers in an occupation in the US. We focus on the 79 items describing work activities and context. For each individual item, a level from 1 to 7 is reported by an incumbent. We standardize each of these variables to have a mean of 0 and a standard deviation of 1.

We follow the psychometric literature (Gorsuch, 1983, 2003; Thomson, 2004) and use exploratory factor analysis to reduce the dimensionality of the ONET data. To extract the underlying latent factors, we first determine the number of factors to retain based on a scree plot from an orthogonal exploratory analysis and the eigenvalue of each individual factor. For example, in the scree plot depicted for the US in Figure D.1 the point where the slope of the curve levels off is just after the third factor (in Figure D.1 the eigenvalues are on the y-axis and the number of factors on the x-axis). This is similar for Britain and Russia. For all three countries the first three factors can explain between 65% and 70% of the variability in the data.

Using orthogonal rotation, we next perform Confirmatory Factor Analysis (CFA) to extract three latent variables. Details of how the items load onto each factor is documented in Table C.1. Utilizing this version of ‘people’ ‘brains’ and ‘brawn’ (PBB) in our job satisfaction regressions does not change the conclusions drawn in the main text (see Table C.3).

The results in the main text follow an approach recommended by Heckman et al. (2012). Specifically, once the first confirmatory analysis is performed, to identify three latent uncorrelated factors we review how every item loads on each factor with the view to dropping items that are weakly associated with all three factors or those that are associated with two or more factors. That is, we remove items that are either weal loaders or cross loaders. Specifically, we remove items with a loading of 0.4 or less on all factors. We remove items that have a loading that is greater than 0.4 on more than one factor. We then repeat the factor analysis using the remaining ONET items and extract the final latent variables which have no items that are weakly loaded or cross loaded and are freely correlated. These latent factors are used in the main analysis. Table C.2 documents how each item loads on these final factors.

The reason we use ONET Version 5 is that it is the only version of ONET where we can match directly to the British data. There are however many versions of this database, with the most recent version being version 14. Tables C.4 and D.5 repeat the US analysis using version 14. The three latent factors are created following the method described in the previous paragraph.

Figure C.1 Scree Plot of the US Exploratory Analysis

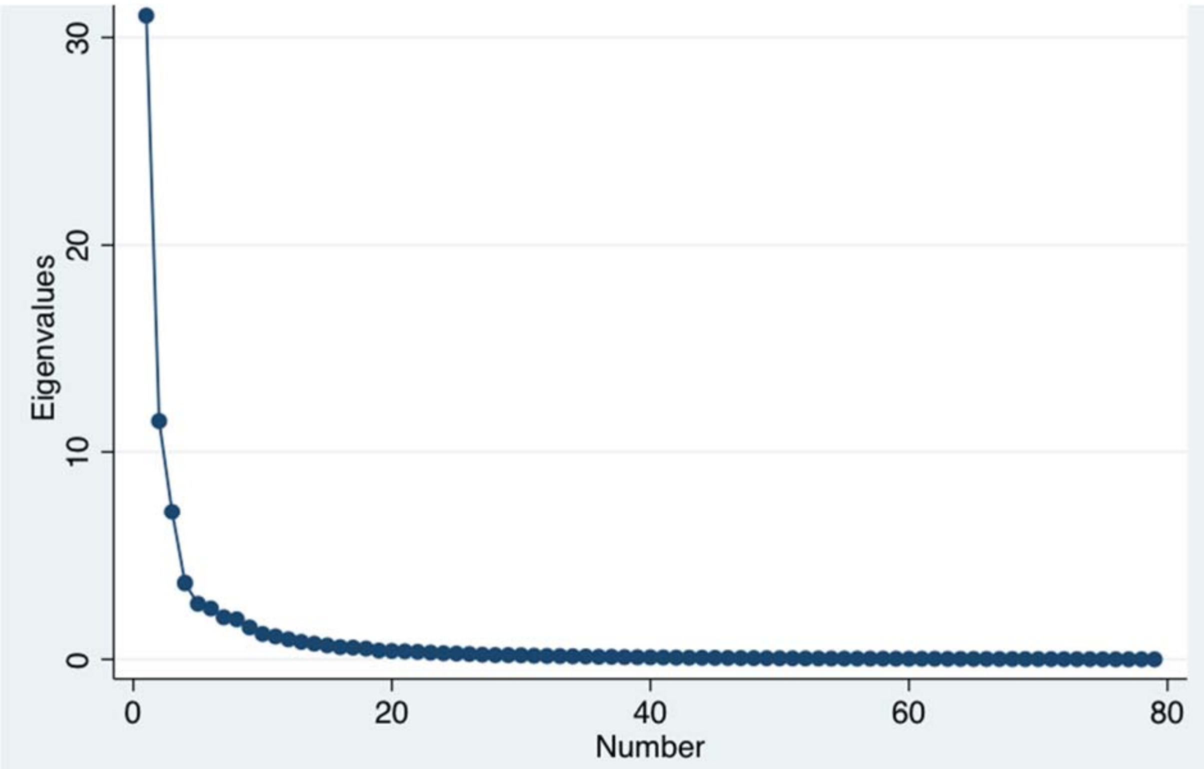


Table C.1 Rotated Factor Loadings of First Rotation (<0.40)

ONET Item	People	Brains	Brawn
Monitoring and Controlling Resources (A)		0.734	
Staffing Organizational Units (A)	0.570	0.488	
Performing Administrative Activities		0.712	
Provide Consultation and Advice to Others (A)	0.796		
Coaching and Developing Others (A)	0.745	0.512	
Getting Information (A)		0.875	
Monitor Processes, Materials, or Surroundings (A)		0.645	
Identifying Objects, Actions, and Events (A)		0.817	
Inspecting Equipment, Structures, or Material (A)		0.402	0.617
Estimating the Quantifiable Characteristics of Products, Events, or Information (A)		0.853	
Judging the Qualities of Things, Services, or People (A)		0.770	
Processing Information (A)		0.833	-0.432
Evaluating Information to Determine Compliance with Standards (A)		0.807	
Analyzing Data or Information (A)		0.870	
Making Decisions and Solving Problems (A)		0.861	
Thinking Creatively (A)		0.701	
Updating and Using Relevant Knowledge (A)		0.739	
Developing Objectives and Strategies		0.819	
Scheduling Work and Activities (A)	0.701	0.554	
Organizing, Planning, and Prioritizing Work (A)	0.443	0.723	
Performing General Physical Activities (A)			0.750
Handling and Moving Objects (A)	-0.442		0.634
Controlling Machines and Processes (A)			0.656
Operating Vehicles, Mechanized Devices, or Equipment (A)			0.512
Interacting With Computers (A)		0.662	
Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment (A)		0.551	
Repairing and Maintaining Mechanical Equipment (A)			0.680
Repairing and Maintaining Electronic Equipment (A)			
Documenting/Recording Information (A)		0.738	
Interpreting the Meaning of Information for Others (A)		0.768	
Communicating with Supervisors, Peers, or Subordinates (A)		0.812	
Communicating with Persons Outside Organization (A)	0.481	0.544	-0.439
Establishing and Maintaining Interpersonal Relationships (A)	0.710	0.411	
Assisting and Caring for Others (A)	0.6051		
Selling or Influencing Others (A)	0.583	0.401	
Resolving Conflicts and Negotiating with Others (A)	0.690	0.483	
Performing for or Working Directly with the Public (A)	0.710		-0.434
Coordinating the Work and Activities of Others (A)	0.538	0.657	
Developing and Building Teams (A)		0.838	
Training and Teaching Others (A)	0.637	0.575	
Guiding, Directing, and Motivating Subordinates A)			
Contact With Others (C)	0.825		

Table C.1 (Continued) Rotated Factor Loadings of First Rotation (<0.40)

ONET Item	People	Brains	Brawn
Deal With External Customers (C)	0.738		
Coordinate or Lead Others (C)	0.723		
Responsible for Others' Health and Safety (C)			0.656
Responsibility for Outcomes and Results (C)			0.701
Frequency of Conflict Situations (C)	0.727		
Deal With Unpleasant or Angry People (C)	0.811		
Deal With Physically Aggressive People (C)	0.704		
Indoors, Environmentally Controlled (C)			-0.539
Outdoors, Exposed to Weather (C)			0.598
Sounds, Noise Levels Are Distracting or Uncomfortable (C)			0.724
Very Hot or Cold Temperatures (C)			0.767
Extremely Bright or Inadequate Lighting (C)			0.707
Exposed to Contaminants (C)			0.789
Cramped Work Space, Awkward Positions (C)			0.840
Exposed to Whole Body Vibration (C)			0.625
Exposed to Radiation (C)			
Disease			
Exposed to High Places (C)			0.623
Exposed to Hazardous Conditions (C)			0.795
Exposed to Hazardous Equipment (C)			0.810
Exposed to Minor Burns, Cuts, Bites, or Stings (C)			0.778
Spend Time Sitting (C)		0.437	-0.531
Spend Time Standing (C)		-0.432	0.499
Spend Time Climbing Ladders, Scaffolds, or Poles (C)			0.690
Spend Time Walking and Running (C)	0.544		
Spend Time Kneeling, Crouching, Stooping, or Crawling? (C)			0.745
Spend Time Keeping or Regaining Balance (C)			0.691
Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls (C)			0.696
Spend Time Bending or Twisting the Body (C)			0.776
Spend Time Making Repetitive Motions (C)	-0.414		
Wear Common Protective or Safety Equipment such as Safety Shoes, Glasses, Gloves, Hearing Protection, Hard Hats, or Life Jackets (C)			0.879
Wear Specialized Protective or Safety Equipment such as Breathing Apparatus, Safety Harness, Full Protection Suits, or Radiation Protection (C)			0.667
Consequence of Error (C)		0.622	
Degree of Automation (C)			
Importance of Being Exact or Accurate (C)		0.588	
Importance of Repeating Same Tasks (C)			
Pace Determined by Speed of Equipment (C)			

Notes: Blanks indicate an item has loaded <0.40 on that factor.

Table C.2 Rotated Factor Loadings of Final Latent Factors (<0.40 is blank)

ONET Item	People	Brains	Brawn
Monitoring and Controlling Resources (A)		0.795	
Staffing Organizational Units (A)			
Performing Administrative Activities		0.682	
Provide Consultation and Advice to Others (A)		0.813	
Coaching and Developing Others (A)			
Getting Information (A)		0.903	
Monitor Processes, Materials, or Surroundings (A)			
Identifying Objects, Actions, and Events (A)		0.857	
Inspecting Equipment, Structures, or Material (A)		0.634	
Estimating the Quantifiable Characteristics of Products, Events, or Information (A)		0.907	
Judging the Qualities of Things, Services, or People (A)		0.813	
Processing Information (A)		0.824	
Evaluating Information to Determine Compliance with Standards (A)		0.842	
Analyzing Data or Information (A)		0.907	
Making Decisions and Solving Problems (A)		0.884	
Thinking Creatively (A)		0.701	
Updating and Using Relevant Knowledge (A)		0.770	
Developing Objectives and Strategies			
Scheduling Work and Activities (A)			
Organizing, Planning, and Prioritizing Work (A)		0.727	
Performing General Physical Activities (A)			0.730
Handling and Moving Objects (A)			0.529
Controlling Machines and Processes (A)			0.657
Operating Vehicles, Mechanized Devices, or Equipment (A)			0.561
Interacting With Computers (A)			0.619
Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment (A)			0.597
Repairing and Maintaining Mechanical Equipment (A)			0.687
Repairing and Maintaining Electronic Equipment (A)			
Documenting/Recording Information (A)		0.789	
Interpreting the Meaning of Information for Others (A)		0.779	
Communicating with Supervisors, Peers, or Subordinates (A)		0.807	
Communicating with Persons Outside Organization (A)		0.521	
Establishing and Maintaining Interpersonal Relationships (A)	0.610		
Assisting and Caring for Others (A)	0.550		
Selling or Influencing Others (A)			
Resolving Conflicts and Negotiating with Others (A)			
Performing for or Working Directly with the Public (A)	0.700		
Coordinating the Work and Activities of Others (A)			
Developing and Building Teams (A)		0.839	
Training and Teaching Others (A)			
Guiding, Directing, and Motivating Subordinates A)			
Contact With Others (C)			0.820
Deal With External Customers (C)	0.771		

Table C.2 (Continued) Rotated Factor Loadings of First Rotation (<0.40)

ONET Item	People	Brains	Brawn
Coordinate or Lead Others (C)	0.634		
Responsible for Others' Health and Safety (C)	0.797		0.460
Responsibility for Outcomes and Results (C)	0.593		
Frequency of Conflict Situations (C)	0.726		
Deal With Unpleasant or Angry People (C)	0.845		
Deal With Physically Aggressive People (C)	0.702		
Indoors, Environmentally Controlled (C)			-0.576
Outdoors, Exposed to Weather (C)			0.639
Sounds, Noise Levels Are Distracting or Uncomfortable (C)			0.742
Very Hot or Cold Temperatures (C)			0.774
Extremely Bright or Inadequate Lighting (C)			0.737
Exposed to Contaminants (C)			0.791
Cramped Work Space, Awkward Positions (C)			0.864
Exposed to Whole Body Vibration (C)			0.646
Exposed to Radiation (C)			
Exposed to Disease or Infections (C)			
Exposed to High Places (C)			0.655
Exposed to Hazardous Conditions (C)			0.804
Exposed to Hazardous Equipment (C)			0.786
Exposed to Minor Burns, Cuts, Bites, or Stings (C)			0.740
Spend Time Sitting (C)			
Spend Time Standing (C)			
Spend Time Climbing Ladders, Scaffolds, or Poles (C)			0.701
Spend Time Walking and Running (C)	0.589		
Spend Time Kneeling, Crouching, Stooping, or Crawling? (C)			0.726
Spend Time Keeping or Regaining Balance (C)			0.710
Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls (C)			0.550
Spend Time Bending or Twisting the Body (C)			0.756
Spend Time Making Repetitive Motions (C)			
Wear Common Protective or Safety Equipment such as Safety Shoes, Glasses, Gloves, Hearing Protection, Hard Hats, or Life Jackets (C)			0.888
Wear Specialized Protective or Safety Equipment such as Breathing Apparatus, Safety Harness, Full Protection Suits, or Radiation Protection (C)			0.731
Consequence of Error (C)		0.677	
Degree of Automation (C)			
Importance of Being Exact or Accurate (C)		0.601	
Importance of Repeating Same Tasks (C)			
Pace Determined by Speed of Equipment (C)			

Notes: Italics indicates that an item has been dropped either because it loaded weakly on all factors (<0.40) or it cross-loaded on more than one factor (>0.40 on more than one factor). Blanks indicate an item has loaded <0.40 on that factor.

Table C.3 Job Satisfaction Regressions
Uncorrelated PBB Factors/All ONET Items

	Samples					
	US – NLSY		Britain – BHPS		Britain – BHPS	
	Overall Job Satisfaction		Overall Job Satisfaction		Satisfaction with Work Itself	
	Females	Males	Females	Males	Females	Males
Share of Males	-0.021	0.034	-0.196	-0.054	-0.271	-0.012
	(0.033)	(0.035)	(0.019)	(0.008)	(0.008)	(0.009)
People	0.035	0.012	0.007	0.005	0.056	0.014
	(0.006)	(0.007)	(0.012)	(0.011)	(0.010)	(0.001)
Brains	0.057	0.046	0.004	-0.005	0.034	-0.018
	(0.008)	(0.007)	(0.012)	(0.002)	(0.014)	(0.015)
Brawn	-0.033	-0.012	-0.033	-0.015	-0.028	0.004
	(0.010)	(0.009)	(0.011)	(0.005)	(0.006)	(0.004)
Number of Observations	75672	80648	48141	43365	48141	43365

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as individual fixed effects. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2. PBB factors are created by confirmatory factor analysis on the first rotation of the exploratory factor analysis described in the methods. The extracted factors are uncorrelated. For all countries three factors are extracted that can loosely be labelled ‘people’ ‘brains’ and ‘brawn’

Table C.4 Regressions for Overall Job Satisfaction using 2014 version of ONET US only

	Samples			
	Uncorrelated PBB Factors All ONET Items		Correlated PBB Factors Weak Items Dropped	
	Females	Males	Females	Males
Share of Males	-0.022	0.007	-0.013	0.011
	(0.034)	(0.034)	(0.034)	(0.034)
People	0.028	0.026	0.026	0.024
	(0.006)	(0.006)	(0.006)	(0.006)
Brains	0.081	0.032	0.072	0.029
	(0.008)	(0.008)	(0.008)	(0.008)
Brawn	-0.047	0.001	-0.044	0.002
	(0.010)	(0.008)	(0.010)	(0.008)
Number of Observations	75004	79743	75004	79743

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as time, area and individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg

Table C.5 Regressions for Stayers using 2014 version of ONET
US only

	Sample and Methods			
	Uncorrelated PBB Factors All ONET Items		Correlated PBB Factors Weak Items Dropped	
	Females	Males	Females	Males
Share of Males	-0.170	0.088	-0.170	0.090
	(0.016)	(0.016)	(0.016)	(0.016)
People	0.017	0.022	0.018	0.025
	(0.003)	(0.003)	(0.003)	(0.003)
Brains	-0.005	-0.005	-0.006	-0.008
	(0.004)	(0.004)	(0.004)	(0.004)
Brawn	0.002	-0.001	0.003	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Observations	75729	80239	75729	80239

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, time and area as well as time, area and individual fixed effects interacted with the sub-periods with consistent occupation codes. Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg

Appendix D: Robustness to Chosen Weights

Our US analysis of the NLSY utilizes sampling weights that reflect that the NLSY79 oversampled Blacks, Hispanics, and the economically disadvantaged. In this appendix we show the corresponding unweighted results.

Our British analysis uses all 18 waves of the original sample of the British Household Panel Survey (BHPS), a longitudinal study of around 5,050 households and approximately 10,000 individuals that began in 1991. This sample was nationally representative of the Great British population. We combine this with the Welsh extension from 1999 (about 1500 households), a Scottish extension from 1999 and a Northern Ireland extension from 2001 (about 1900 households). We make this decision to preserve as many data points as possible, however we document in this appendix results which are based on responses from the original nationally representative sample only. Additionally, we documented results from weighted regressions of the main BHPS sample, where the weights are the longitudinal weights described in Taylor et al (2010). These are the weights recommended for use in longitudinal analysis, however we lose a significant amount of our sample owing to these weights only being provided when an individual was present in all waves.

Our RLMS regressions use weights that allow for the complex design of the RLMS where many observations are derived from following the housing unit rather than the person, as well as having oversamples from the first wave to allow for forecasted attrition. In this appendix we document unweighted regressions.

Table D.1 Overall Job Satisfaction Regressions

	Model and Sample					
	OLS		Fixed Effects		Fixed Effects with ONET	
Share of Males	Females	Males	Females	Males	Females	Males
NLSY –Unweighted	-0.266	0.006	-0.132	0.002	-0.029	0.068
	(0.058)	(0.060)	(0.024)	(0.023)	(0.027)	(0.029)
Number of Observations	75672	80648	75672	80648	75672	80648
BHPS – Original Sample	-0.289	0.005	-0.233	-0.100	-0.182	-0.059
unweighted	(0.029)	(0.025)	(0.040)	(0.025)	(0.028)	(0.007)
Number of Observations	35525	32266	35525	32266	35525	32266
BHPS – with	-0.234	-0.091	-0.218	-0.010	-0.129	-0.057
Longitudinal Weights	(0.030)	(0.036)	(0.003)	(0.028)	(0.016)	(0.008)
Number of Observations	19793	16064	19793	16064	19793	16064
RLMS	-0.156	-0.008	-0.131	-0.079	-0.101	-0.082
Unweighted	(0.095)	(0.059)	(0.056)	(0.050)	(0.057)	(0.063)
Number of Observations	35443	27117	35443	27117	35443	27117

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, and time and area fixed effects. Individual fixed effects interacted with the sub-periods with consistent occupation codes are included in columns (2) to (6). Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table D.2 Regressions for Job Satisfaction with Work Itself

	Model and Sample					
	OLS		Fixed Effects		Fixed Effects with ONET	
Share of Males	Females	Males	Females	Males	Females	Males
BHPS – Original Sample unweighted	-0.295 (0.020)	-0.052 (0.032)	-0.285 (0.037)	-0.050 (0.031)	-0.198 (0.040)	-0.010 (0.034)
Number of Observations	35525	32266	35525	32266	35525	32266
BHPS – with Longitudinal Weights	-0.300 (0.030)	-0.058 (0.046)	-0.255 (0.064)	-0.119 (0.043)	-0.139 (0.083)	-0.081 (0.039)
Number of Observations	19793	16064	19793	16064	19793	16064

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, and time and area fixed effects. Individual fixed effects interacted with the sub-periods with consistent occupation codes are included in columns (2) to (6). Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Table D.3 Stayers

	Model and Sample					
	OLS		Fixed Effects		Fixed Effects with ONET	
Share of Males	Females	Males	Females	Males	Females	Males
NLSY Unweighted	-0.228	0.067	-0.188	0.072	-0.192	0.114
	(0.048)	(0.033)	(0.011)	(0.010)	(0.012)	(0.013)
Number of Observations	76375	81144	76375	81144	76375	81144
BHPS – Original Sample	-0.187	0.080	-0.296	0.118	-0.287	0.147
unweighted	(0.038)	(0.005)	(0.016)	(0.028)	(0.019)	(0.040)
Number of Observations	27302	24886	27302	24886	27302	24886
BHPS – with Longitudinal	-0.166	0.122	-0.276	0.155	-0.284	0.166
Weights	(0.054)	(0.011)	(0.064)	(0.047)	(0.055)	(0.045)
Number of Observations	16181	13371	16181	13371	16181	13371
RLMS	-0.368	0.175	-0.278	0.069	-0.235	0.095
unweighted	(0.073)	(0.065)	(0.063)	(0.051)	(0.068)	(0.059)
Number of Observations	23449	16792	23449	16792	23449	16792

Notes: All regressions also include age and age squared of the individual, the averages of the log hourly wage, hours, fraction college graduates, and age in the occupation, and time and area fixed effects. Individual fixed effects interacted with the sub-periods with consistent occupation codes are included in columns (2) to (6). Standard errors are two-way clustered (by individual and their occupation) and shown in parentheses. Models are estimated using xtivreg2.

Appendix E: Other Data Sets

We consider the association between the SOM and overall job satisfaction in three other data sets. These are the (US based) General Social Survey (GSS), the European Social Survey (ESS) and the International Social Survey Programme (ISSP). For these three surveys we only have cross sectional data so cannot consider stayers. Thus we estimate the following equation:

$$Y_{ijct} = \delta SOM_j + X_j\beta + X_{ijct}\gamma + \mu_t + \varpi_c + \varepsilon_{ijct}$$

where Y_{ijct} is job satisfaction of individual i in occupation j, residing in country/area c in year t, SOM_j is the proportion of males in a particular occupation, X_j is a vector of other occupational averages, X_{ijt} is a vector of individual-level control variables, μ_t are wave effects, and ϖ_c are country/area fixed effects. In the baseline specification, we follow the specification as described for equation 1 in the main text. We calculate standard errors using two-way clustering by country and occupation. The second specification adds the 'people', 'brains' and 'brawn' factors. For the GSS data, the factors and occupational averages are calculated using the CPS data following the same procedure as described for the NLSY analysis. For the ISSP and ESS data PBB are calculated by matching the ONET 5.0 data to merged ISSP and ESS data and calculating the factors in the same manner we describe in the main text. Occupation averages are calculated using the same data. Table G.1 documents the results. We note that estimating separate regressions for the UK using the ESS yields far larger negative coefficients for the SOM in the baseline female regressions. The same is true if we run separate regressions for the USA and the UK using the ISSP data.

Table E.1 Overall Job Satisfaction Regressions in Various Datasets

	Data and Sample											
	Europe – ESS				International – ISSP				US – GSS			
	Females	Males	Females	Males	Female	Males	Female	Males	Females	Males	Females	Males
Share of Males	-0.079	0.059	-0.024	0.068	-0.080	-0.019	0.018	0.014	-0.235	0.078	-0.128	0.007
	(0.032)	(0.024)	(0.035)	(0.029)	(0.033)	(0.039)	(0.033)	(0.044)	(0.036)	(0.041)	(0.044)	(0.053)
People			0.036	0.017			0.051	0.031			0.008	0.030
			(0.008)	(0.007)			(0.011)	(0.012)			(0.007)	(0.008)
Brains			0.041	0.072			0.036	0.047			0.091	0.063
			(0.009)	(0.009)			(0.011)	(0.014)			(0.012)	(0.011)
Brawn			-0.057	0.008			-0.052	0.011			-0.028	0.038
			(0.008)	(0.008)			(0.012)	(0.009)			(0.011)	(0.012)
Number of Observations	27703	28038	27703	28038	22959	23427	22959	23427	18608	15100	18608	15100

Notes: All regressions also include age and age squared of the individual, the averages of the, hours, fraction college graduates, and age in the occupation, time and country individual fixed effects. Standard errors are two-way clustered (by country and their occupation) and shown in parentheses. Models are estimated using xtivreg2

Appendix F: Cross-Walking Across Samples

US Analysis

We use pooled monthly CPS samples from 1983-1991 and 2003-2010 to calculate the share of males (SOM) and occupational averages for the 1980 and 2000 three-digit census occupation codes respectively. We match the CPS averages derived from the 1980 occupation codes directly to the 1982-2000 NLSY data and the averages derived from the 2000 occupation codes to the 2002-2012 NLSY data. There is then a single average for all the years within the sub-periods when occupation codes are unchanged. In order to allow for breaks in the occupation coding, we allow for individual times sub-period specific fixed effects in some of our regressions.

Our main analysis uses ONET version 5, whose items on activities and context are linked to Standard Occupation Codes (SOC) 2000. We start by using a Bureau of Labor Statistics (BLS) cross walk to assign a three-digit Census 2000 occupation code to each ONET item. We then use a further crosswalk created by Autor and Dorn (2013) and Dorn (2009) that matches three-digit Census 2000 occupation codes to earlier Census codes. Using these two crosswalks, we create a consistent set of occupations matching the 1980 and 2000 Census codes and SOC 2000. Call this consistent code occ1990dd. Since the CPS and NLSY use 1980 and 2000 Census codes, we can now match ONET variables to a single, consistent occupation for these two data sets throughout the entire sample period.

Merging the ONET items to the CPS file, we calculate three latent factors ‘people’ ‘brains’ and ‘brawn’ (PBB). Subsequently, we match the PBB variables to the NLSY data. As a result, there is a single PBB variable for the entire sample period in the NLSY. This is in contrast to the SOM and other occupation averages, which we have created by the sub-periods when Census 1980 and 2000 codes were in use. Note that we are only using a single version of the ONET data together with all our other data, hence the creation of the new, single occ1990dd code here.

British Analysis

We calculate the SOM and other occupation averages in a three-digit occupation using the 1993-2012 Quarterly Labor Force Survey (QLFS). The QLFS uses SOC90 codes from 1993 through 2000 and UK SOC00 from 2001. We calculate the occupation averages for each sub-period when the SOC90 and SOC00 were in use. We then match the occupation averages to the BHPS data for the relevant sub period.

Our main analysis uses ONET version 5, whose items on activities and context are linked to US Standard Occupation Codes (SOC) 2000. We match the US SOC00 codes in the ONET data directly to the British SOC00 using a crosswalk provided by Anna Salomons. We therefore need to assign a British SOC00 to every occupation in the QLFS, but no official cross-walk exists. The BHPS does provide a British SOC90 code for every wave and post 2000 this code appears alongside the British SOC00. We utilize this implicit crosswalk in the BHPS to assign a British SOC00 code to each British SOC90 code in the QLFS from 1993-2000. We match the ONET items to the QLFS using the British SOC00 codes. The three latent factors ‘people’ ‘brains’ and ‘brawn’ (PBB) are calculated using this data.

In addition, we use the same implicit crosswalk to assign a British SOC00 code to each British SOC90 in the BHPS from 1991-2000. We match the PBB factors for each occupation to the BHPS data using the British SOC00 codes.

Russian Analysis

Pooling the ISSP 1995-2011, the ESS 2002-2012 and the RLMS 1994-2012, we calculate the SOM in each occupation, along with the other occupation averages based on the three digit ISCO 2000 codes. We match the items from ONET version 5 to ISCO 2000 utilizing a crosswalk provided by the BLS between SOC 2000 and ISCO 2000.