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Appendix, Not for Publication

Grace Lordan and Jörn-Steffen Pischke
LSE

July 2016

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.[1] 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.[2] 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

[1.] Specifically, the ATUS uses Census occupation codes, so we crosswalk as described in the US data section.

[2.] 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

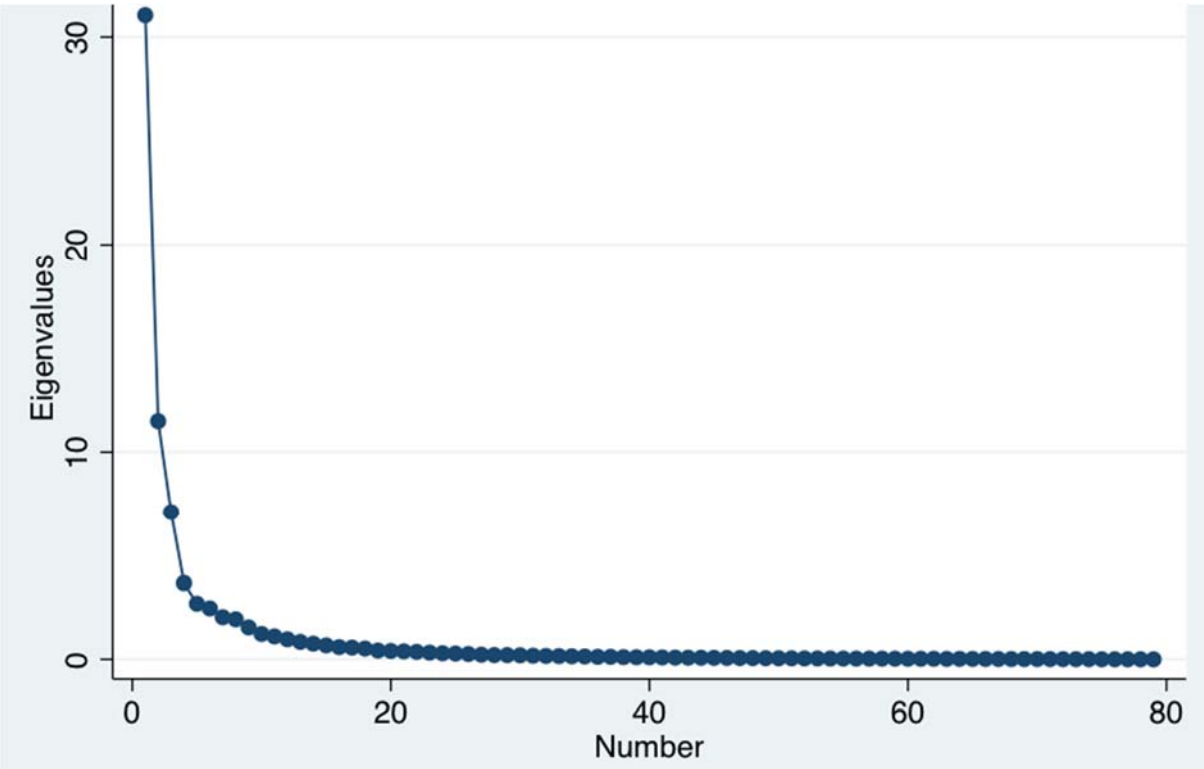

Table C.1 Rotated Factor Loadings of First Rotation (<0.40)

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 unweighted	-0.187	0.080	-0.296	0.118	-0.287	0.147
	(0.038)	(0.005)	(0.016)	(0.028)	(0.019)	(0.040)
Number of Observations	27302	24886	27302	24886	27302	24886
BHPS – with Longitudinal Weights	-0.166	0.122	-0.276	0.155	-0.284	0.166
	(0.054)	(0.011)	(0.064)	(0.047)	(0.055)	(0.045)
Number of Observations	16181	13371	16181	13371	16181	13371
RLMS unweighted	-0.368	0.175	-0.278	0.069	-0.235	0.095
	(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.