Women at work: Discovering heterogeneous returns to tasks across genders.

Elizabeth J. Casabianca [*] Alessia Lo Turco [†] Claudia Pigini[‡]

PRELIMINARY VERSION. PLEASE DO NOT QUOTE.

Abstract

We explore the existence of gender differences in returns to motor, cognitive and social tasks by using individual level data for a sample of U.S. workers. We uncover in this data a wage premium for women performing highly manual or highly cognitive jobs, after finding no evidence of cross-occupation correlation between task returns nor of self-selection of workers according to their comparative advantages, that would otherwise threaten the identification of average task returns in the labour market. Furthermore, a wage penalty characterises women engaged in highly social intensive jobs. These results hold both across- and within-occupations as well as when we control for within-industry effects.

KEYWORDS: Task approach, Gender-specific task returns, motor-cognitive-social tasks.
JEL CLASSIFICATION : J24, J31

[*]Università Politecnica delle Marche (IT) - e.j.casabianca@univpm.it
[†]Università Politecnica delle Marche (IT) - a.loturco@univpm.it
[‡]Università Politecnica delle Marche (IT) - c.pigini@univpm.it

1 Introduction

Increasing returns to cognitive tasks in the U.S. labour market and the higher occurrence of women in cognitive intensive jobs have been interpreted as one of the main causes of the declining gender wage gap in recent decades (Bacolod and Blum, 2010). Recently, though, the diffusion of information and communication technologies has led employers to value increasingly more interpersonal - "people" - skills which have, therefore, gained importance in the labour market (Autor et al., 2003; Black and Spitz-Oener, 2010).[1] Declining gender wage disparities could, hence, further reflect the growing returns to people tasks and the higher occurrence of female workers in social intensive occupations (Borghans et al., 2006, 2014; Deming, 2015).

In this respect, previous studies have confirmed that men and women sort differently into occupations: while men show a higher propensity for “brawn" intensive activities, women are more inclined to perform “brain" intensive activities (Welch, 2000; Rendall, 2010). Also, among highly cognitive jobs, women show highest preference for jobs with better anticipated work-life balance and tend to be less identified with stereotypically masculine jobs (Barbulescu and Bidwell, 2013). In this line, research has shown that women differently assess, with respect to men, the importance of money and people in a job and these factors have different returns in the labour market (Fortin, 2008). Experimental evidence further stresses the higher pro-sociality of female individuals (Eckel and Grossman, 2008; Kamas and Preston, 2015; Khachatryana et al., 2015). Hence, women tend to sort relatively more into occupations that are more socially rather than monetarily oriented (Krueger and Schkade, 2008; Grove et al., 2011). All this evidence points towards defining occupations as the outcome of the combination of different job tasks to be performed with heterogeneous intensities and, as a consequence, individuals are likely to choose their job in order to achieve the best possible matching between their skills/or preferences and an occupation’s tasks requirement (Autor and Handel, 2013; Boehm, 2015). Then, if men and women have different relative abilities/preferences in performing labour tasks, the latter’s returns may differ across genders.

Against this background, we inspect the existence of gender differences in returns to motor, cognitive and social tasks in a sample of U.S. workers. We use the extended questionnaire from the Princeton Data Improvement Initiative (henceforth PDII), recently employed by Autor and Handel (2013), where tasks are measured at the individual level. Differently from them, though, we identify from the survey those questions that inform on the pure manual, cognitive and social content of a job. More specifically, we define a task as manual on the basis of the time spent in motor activities while we define it as cognitive on the ground of the proportion of time involved in problem solving and in the use of mathematics and of the typical length of the documents read on the job. As far as the definition of social task is concerned, we select information on the proportion of time spent managing or supervising other workers, on the proportion of time spent working in team and on the extent of face to face contacts with people other than co-workers or supervisors required by the job. The individual task level measures that we define from the PDII are highly similar to occupation level task measures based on the O*NET database and usually adopted by

[1] Recent contributions enrich this framework by showing that *directness* and *caring* are two relevant components of the social nature of a job (Borghans et al., 2006). Additionally, Weinberger (2014) provides evidence on the increasing complementarity between cognitive and social tasks. Despite acknowledging the potential existence of these effects in our data, we do not test them here as this is not the scope of our work.

extant literature, such as the motor and cognitive task indicators proposed by Peri and Sparber (2009) and Bacolod and Blum (2010) and the definition of social tasks by Deming (2015). Also, they are very close to an occupation level task definition based on a multivariate factor analysis of all items in the O*NET Abilities, Context and Activities survey without initial hand-picking of relevant characteristics (Ingram and Neumann, 2006). Once defined the three relevant labour tasks, we inspect whether, across and within occupations as well as within industries, task returns differ between men and women.

The adoption of a task approach (Autor and Handel, 2013) to the analysis of labour market returns,[2] though, demands to recognize that tasks cannot be treated as human capital characteristics, as they represent the demand and supply of skills respectively. By extending the Roy (1951) model, an occupation can be described by an indivisible bundle of differently rewarded job tasks workers have to perform jointly. Workers will then choose the occupation yielding the highest premium to the whole bundle of tasks. In this framework, cross-occupation heterogeneity of task returns represents a crucial difference with respect to Mincer's human capital framework, in which the marginal productivity of one year of education is invariant across jobs. Another issue concerns the self-selection of workers into occupations according to their comparative advantage in performing job tasks, which prevents the identification of average task returns in the typical Mincerian wage regression. To tackle these issues, we derive suitable model parametrizations that allow us to test for the absence of non-uniformly positive cross-occupation covariance of task returns, a necessary condition for the data to be consistent with self-selection, and of comparative advantages. If we fail to reject the above hypotheses with this specific data, we can confidently interpret the estimated task coefficients as average returns to tasks.

Anticipating our findings, we show that cognitive and social tasks are positively rewarded in the labour market while the opposite occurs to manual tasks, both across and within occupations. Individual level characteristics and occupational heterogeneity only partially drive the observed patterns. Also, when we account for heterogenous returns to tasks across genders, we find that, *ceteris paribus*, both across and within occupations as well as within industries, women performing highly cognitive or highly manual tasks enjoy a wage premium while those employed in highly social intensive jobs face a wage penalty. Although this last piece of evidence may seem the opposite of what one would expect, it may be consistent with our definition of social tasks, which include managerial and team work activities. Indeed, it is a consolidated fact that the gender wage gap widens as women advance in their career (Frank, 2015).

By adopting a job task perspective to gender differences in the labour market, our work is close to extant literature measuring tasks at the occupation level and applying the task approach to explain the evolution of the gender wage gap in the U.S. labour market. Yamaguchi (2015) applies a task framework and assumes that occupations are made up of motor and cognitive tasks measured from the Dictionary of Occupational Titles (DOT). Using labour data for the U.S. from the late 1970s to the mid-1990s, he finds that technological progress has contributed to the narrowing of

[2] The task-approach finds strong support in recent literature showing that the human capital exploited by individuals in the work place is task-specific (Gathmann and Schonberg, 2010). Task based models represent an evolution of the traditional human capital theory. The latter seems inappropriate to explain the gender pay gap observed on the U.S. labour market (England, 1982). During the last decade, women outperformed men in terms of educational attainment (Bailey and Dynarski, 2011). However, *ceteris paribus*, women's higher educational attainment is not reflected in higher wages.

the gender wage gap since it tended to substitute male workers concentrated in motor-task intensive occupations. The definition of cognitive task in this study includes a few relevant interpersonal abilities, hence the possibility of separate returns for cognitive and interpersonal activities on the job is neglected in the analysis. Bacolod and Blum (2010), instead, explore the contribution of changes in returns to people tasks beyond the contribution of changes in returns to cognitive and motor tasks to the wage gap. They apply principal component analysis to the DOT to recover scores for each of the four abilities and explore the evolution of returns to these four different sets of job tasks in the U.S. labour market between the early 70s and 1990. They conclude that social activities are only remunerated in conjunction with cognitive ones and that the narrowing of the gender wage gap was due to increases in the relative prices of cognitive tasks and to the concentration of female labour in cognitive intensive occupations. Nonetheless, as in the study by Yamaguchi (2015), they do not account for heterogeneous returns to tasks across genders. Finally, Deming (2015) theoretically addresses the importance of social activities, beyond the cognitive and the motor ones, by modeling social tasks as reducing coordination costs allowing workers to specialise and trade more efficiently. Empirically he tests his model's predictions on the National Longitudinal Survey of Youth (NLSY79) on the 1980-2012 period. He shows evidence of endogenous occupational sorting of more sociable individuals into social intensive occupations and of the more cognitive ones into more cognitive occupations. Also, he finds that workers highly intensive in cognitive and social tasks earn relatively higher wages in social intensive occupations. However, in Deming's (2015) study gender disparity issues are only descriptively addressed.

Compared to the above studies, we therefore move a step further by inspecting returns to manual, cognitive and social tasks in the U.S. labour market exploiting, for the first time to our knowledge, task measures at the individual level rather than at the occupational one. We can, therefore, answer the question on whether labour tasks constitute a relevant level of analysis when investigating returns to jobs in the labour market. In particular, we can highlight whether heterogenous tasks are differently rewarded even within narrowly defined occupations. In this respect, Autor and Handel (2013), on the same PDII data that we use, have already shown that individual task indicators have a relevant explanatory power of individual wages, both across and within occupations. However, they have defined an occupation as a composite of abstract, manual and routine tasks. In this work we modify their setting and from their abstract task we separate the strictly cognitive activities from the supervision and managerial ones and combine the latter with information on team work and face to face contacts to build a separate measure of social tasks. Furthermore, while retaining Autor and Handel's (2013) definition of manual task, we neglect the role of routine activities, which in their analysis proved to be non robust and less relevant. A second original contribution of our work, is the inspection of the possibility of labour market segmentation across genders: the existence of different wage returns across genders within narrowly defined occupations could play a role in shaping the evolution of the gender wage gap in recent decades. Additionally, we are able to interpret our estimates in this data as job tasks returns as we find no evidence of cross-occupation correlation between task returns nor of self-selection of workers according to their comparative advantages. Finally, we provide a comparison of our individual task based empirical setting to a similar setting based on the use of occupation level task measures and find that although the latter is able to account for the pattern of wage premia and penalties that characterise motor, cognitive and social tasks in the PDII sample, it is not able to account for heterogeneous returns across genders.

The rest of the paper is structured as follows. Section 2 describes the individual and occupational level data used throughout the paper. The theoretical framework and the empirical strategy are introduced in Section 3. Our main results are presented in Section 4 while Section 5 concludes.

2 Data

The main data source we use is a module of the Princeton Data Improvement Initiative (PDII) survey which provides, among other, information on job activities regularly performed by workers. The main advantage of the PDII is that it collects job tasks data at the individual level, thereby allowing to account for within occupation differences in job tasks returns. It also provides information on workers' hourly wages and characteristics, such as gender, age, education and self-reported ethnicity, which we use in our analysis as individual level controls together with a measure of potential work experience that we build as in Peri and Sparber (2009).

After some data cleaning,[3] we build three task measures that capture the cognitive, social and manual dimensions of job tasks. We exploit three items from the PDII that provide information on cognitive task demands: 1) frequency of problem solving tasks using mathematics; 2) frequency of problem-solving tasks requiring at least 30 minutes to find a solution; and 3) the length of a document typical read on the job (ranging from one page or less to more than 25 pages). We yield a standardized cognitive task measure by performing a principal component analysis (PCA) on the three items, which explains 50% of their variation. We measure the interpersonal skill tasks required by an occupation by exploiting three PDII items that capture the social character of a job: 1) amount of workday spent managing and supervising others; 2) how much of the work is performed in a unit/group; and 3) the extent of required face-to-face contact on the job with people other than co-workers or supervisors. We use these three items to build a broad measure of social task requirements by taking the first yielded component of a PCA, which explains 43% of their variation. The measure of manual task is built by standardizing the single PDII item on how much of the workday involves doing physical tasks.

Table A1 in the Appendix shows each PDII item used to build the three main task measures used in our analysis and describes the content of jobs in our sample. Regarding cognitive related tasks, worthy of note is that more than 45% of workers engage in extended problem solving tasks on a daily basis, while only 15% make use of mathematics at least once a day. Almost half of our sample spends most of their time supervising and/or managing others as well as having face-to-face contact with people other than colleagues and supervisors. Yet, most individuals work

[3]We begin by restricting the initial sample of 2,513 observations by retaining individuals who: are currently working (239 observations dropped); are between 18 and 64 years of age (215 observations dropped); and have non-missing education (7 observations dropped). Moreover, we apply some changes to the scale of the responses: we first replace refused and/or ambiguous responses (e.g. inapplicable, DK) with missing values and, for questions on job tasks regularly accomplished, we create consistent scales across responses (from an initial value of 1 when the job task is never performed, to a maximum value of 5 when the job tasks is regularly performed.). Next, we consider individuals with: non-missing task measures (25 observations dropped); non-missing wage data (505 observations dropped); non-missing, non-military and non-farm occupations (29 observations). Finally, we keep occupations with at least two workers (167 observations dropped) and individuals with a valid number of potential years of experience (10 observations dropped). Our final estimation sample comprises 1,316 observations.

by themselves. Finally, almost 40% of workers report spending all of their time at work doing physically engaging tasks, 53% of whom are women. At the same time, more than a third of wage and salary workers, and in particular women, don't spend any of their time doing physical tasks. It is worth highlighting that, compared to men, women spend more of their work time in performing physical and social related tasks especially involving managerial duties.

Table 1 presents descriptive statistics of the dependent and independent variables used in the empirical application. It emerges that female hourly wages are on average lower than male hourly wages. Concerning the individual characteristics there is no significant difference between women and men. With regard to the task indicators, it is interesting to note that males and females do not significantly differ in terms of their manual content of jobs.

For the sake of comparison, we match the individual level data with the O*NET 3.0 database released in 2000,[4] which represents an alternative source of information on job task content at the occupation level.[5] To construct task indicators comparable to those of the PDII survey, we hinge on extant literature and build Peri and Sparber's (2009) broad measure of motor abilities[6] and derive a cognitive job task measure similar to that of Bacolod and Blum (2010).[7] As we are interested in capturing the social dimension of a job too, we follow Deming (2015) and build an indicator measuring the social tasks required by an occupation and hand-pick the following O*NET Skills survey items: 1) Coordination; 2) Negotiation; 3) Persuasion; and 4) Social Perceptiveness. For each occupation, the O*NET assigns a score to each item reflecting how important it is for the performance of that job. The items are combined into a standardized scale by taking the first yielded component of a PCA. We end up with three task measures identifying the manual, cognitive and social character of jobs.[8] To complete the comparison of our task proxies to the literature based task measures, we further consider the O*NET based set of abstract, routine and non-routine manual indicators as in Acemoglu and Autor (2011) and the set of PDII based individual level task

[4]To this purpose, we use readily available crosswalks from the O*NET SOC2000 to the SOC2000 classification scheme.

[5]The O*NET, the Occupational Information Network of the U.S. Department of Labor, is a database of occupational information. The database is freely available on the following web-page: $https://www.onetcenter.org/db\_releases.html$. It is worth stressing that we use its 2000 version due to the time span covered in the IPUMS-CPS sample (see below for more details). As such, our analysis ignores changes in the task content of occupations over time, nonetheless we choose the 2000 O*NET survey as it represents an intermediate year in our IPUMS-CPS sample and we expect it to consistently depict the occupational content of our time span.

[6]From the O*NET Abilities survey we select the following set of abilities: Perceptual Speed, Spatial Orientation Importance, Visualization, Selective Attention, Time Sharing, Arm-Steadiness, Manual Dexterity, Finger Dexterity, Control Precision, Multi-Limb Coordination, Response Orientation, Rate Control, Reaction Time, Wrist-Finger Speed, Speed of Limb, Static Strength, Explosive Strength, Dynamic Strength, Trunk Strength, Stamina, Extent Flexibility Level, Dynamic Flexibility, Gross-Body Coordination, Gross-Body Equilibrium. Compared to Peri and Sparber (2009) we neglect visual and hearing perception to focus on strictly motor characteristics which do not necessarily entail interactive activities which in our framework are meant to be captured by the social task indicator.

[7]From the O*NET Abilities survey we select the following set of cognitive abilities: Fluency of Ideas, Originality, Problem Sensitivity, Deductive Reasoning, Inductive Reasoning, Information Ordering, Category Flexibility, Mathematical Reasoning, Number Facility, Memorization, Speed of Closure. Compared to Bacolod and Blum (2010) we neglect oral expression and comprehension abilities not to include in the definition of cognitive tasks any interactive activity which we aim at measuring by means of the social task indicator.

[8]The first component explains 51% of the variation of the items corresponding to the manual task measure, 54% of the variation of the items corresponding to the cognitive task indicator and 77% of the variation of the items concerning the social task measure.

Table 1: Summary statistics of selected variables from the PDII survey.

Variable	Female	No. Obs.	Mean	Min	Max	Std. Dev.	T-test
Log hourly wage	0	592	3.07	1.10	7.45	0.64	8.50
	1	724	2.81	0.41	6.55	0.63	
Age	0	592	45.31	18	64	11.57	-1.37
	1	724	46.17			10.96	
Less than high school	0	592	0.11	0	1	0.31	1.48
	1	724	0.08			0.27	
Some college	0	592	0.11	0	1	0.32	-0.79
	1	724	0.18			0.39	
College	0	592	0.30	0	1	0.46	-1.21
	1	724	0.36			0.48	
Postcollege	0	592	0.13	0	1	0.34	0.85
	1	724	0.14			0.35	
Experience	0	592	21.42	0	47	12.80	-1.41
	1	724	23.46			11.86	
Spanish language	0	592	0.04	0	1	0.21	**?**
	1	724	0.05			0.22	
Black	0	592	0.10	0	1	0.30	**?**
	1	724	0.12			0.32	
Hispanic	0	592	0.17	0	1	0.38	**?**
	1	724	0.15			0.35	
Asian	0	592	0.02	0	1	0.15	**?**
	1	724	0.01			0.11	
Manual	0	592	0.04	-1.16	1.11	1.01	1.89
	1	724	-0.20			0.99	
Cognitive	0	592	0.16	-2.33	1.98	1.00	6.08
	1	724	-0.17			0.96	
Social	0	592	0.05	-1.68	1.95	0.99	1.92
	1	724	-0.06			0.97	

Note: Total no. of obs.: 1,316. Dummies for educational attaintment are built on a single PDII question as follows: *Less than high school* is equal to 1 when the individual has less than a high school/GED diploma, and 0 otherwise; *Some college* is equal to 1 when the individual has some college education but no degree, and 0 otherwise; *College* is a dummy equal to 1 when the individual has a vocational/ tech school degree, associate college degree or a bachelor's college degree, and 0 otherwise; *Postcollege* is equal to 1 when the individual has a master's degree, a professional school degree or a doctoral degree, and 0 otherwise. *Experience* is measured in years and is built assuming that individuals without a high school diploma started working at the age of 17, while those with a diploma started working at the age of 19. *Spanish language* is a dummy for primary Spanish language speaker equal to 1 for individuals who required the Spanish language version of the PDII survey. *Black*, *Asian* and *Hispanic* are dummies built according to self-reported ethnicity.
Source: PDII survey.

measures as from Autor and Handel (2013).

We supplement these six hand-picked indicators with three other found by applying a methodology proposed by Ingram and Neumann (2006).[9] Specifically, we perform a multi-factor component analysis (MCA) on fourteen clusters grouping the 150 items included in the O*NET Abilities, Activities and Context surveys.[10] By doing so, we yield three factor scores - y_1, y_2 and y_3 - that are able to summarize the O*NET information in a more meaningful way.[11]

Table A2 in the Appendix shows correlation coefficients between the PDII and O*NET task measures. In the top panel we consider the PDII person level task indicators. It is worth stressing the similarity between the O*NET occupation level tasks and their corresponding person level PDII measures. Also, it is important to note how well y_1, y_2 and y_3 proxy the O*NET occupation level task measures. Specifically, y_1 provides the same information as *Manual* ($\rho = 0.96$), y_2 is strongly correlated with *Social* ($\rho = 0.87$) and y_3 with *Cognitive* ($\rho = 0.76$). This finding is supported when we look at the correlations between y_1, y_2 and y_3 and the person level PDII task indicators. In the bottom panel, we replace the PDII person level task measures with their means at the occupation level. The correlation coefficients are very similar to those observed in the top panel and further corroborate our previous findings.

Before delving into our main estimation results, we run a descriptive analysis on the extent to which self-reported tasks are related to worker and job characteristics. To this end we run an OLS regression which takes the following form (Autor and Handel, 2013):

$$T_{ij} = \alpha + \delta_1 H_i + \psi fem_i + \gamma' X_i + \epsilon_{ij} \tag{1}$$

where T_{ij} are the standardized PDII person level task measures of individual i in occupation j, H are human capital measures (education, potential experience, primary language), X is a set of demographic characteristics (ethnicity, sex), and γ is a vector of 240 occupation dummies (with one omitted) which proxy for technical job requirements. Here, the reference group is white male English-speaking high school graduates.

Table 2 presents the estimation results. Models (1), (3) and (5) use demographic and human capital variables as explanatory variables; models (2), (4) and (6) also include the 240 occupation dummies. As one would expect, individuals with at least some college education are least likely to use manual intensive job tasks and this effect is not attenuated when conditioning on occupational sorting in Column (2). A primary Spanish language speaker is more likely to employ manual intensive occupations, although this effect is no longer significant when accounting for within-occupation differences. With regard to *Cognitive* and *Social*, highly educated workers are more likely to make use of cognitive and social intensive job tasks, although this effect in part disappears when controlling for occupation specific requirements. Primary Spanish language speakers are

[9] Ingram and Neumann (2006) run a multi-factor component analysis (MCA) on the fifty-three skill characteristics provided in the DOT and find that four factors are able to synthesize all of the information provided therein.

[10] Worthy of note is that our classification in fourteen clusters is not random at all but are inspired by the aggregation provided on the O*NET web-page. In particular, we run the analysis on the original eleven O*NET clusters by reversing the scales of those items ranking occupations differently from the majority of the items in each cluster, or directly on the 150 items, which, however, did not alter the insights and outcomes of our analysis.

[11] It is worth stressing that here we apply an exploratory factor analysis, that is we have no a priori on the number of common factors to be selected. Since the factor analysis does not yield a unique solution, we used an orthogonal factor rotation that, however, did not produce a different interpretation of the obtained factor scores.

less likely to use highly cognitive and social intensive jobs. As for ethnicity, blacks are more likely to employ social intensive jobs, differently from asians. Finally, across occupations women tend to make less use of cognitive and social tasks, however within occupations it appears that there are no differences in job tasks use between genders. Our findings are substantially in line with those of Autor and Handel (2013). Like us, they find that the effect of demographic and human capital characteristics on the use of job tasks is in part mediated by occupational assignment. They also find that higher education and primary Spanish language are able to predict both between- and within-occupation job task differences.

Table 2: Regressions of standardized PDII person level task measures on demographics, human capital measures and occupation dummies.

	Manual		**Cognitive**		**Social**	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.00	0.09	-0.35***	-0.04	-0.13*	-0.09
	[0.06]	[0.07]	[0.06]	[0.08]	[0.07]	[0.07]
Less than high school	-0.04	0.01	-0.10	-0.08	0.45**	0.19
	[0.13]	[0.12]	[0.14]	[0.13]	[0.18]	[0.14]
Some college	-0.32***	0.06	0.44***	0.19*	0.03	-0.01
	[0.09]	[0.10]	[0.10]	[0.10]	[0.12]	[0.10]
College	-0.59***	-0.15**	0.54***	0.09	0.25***	0.06
	[0.08]	[0.08]	[0.09]	[0.09]	[0.09]	[0.08]
Postcollege	-0.99***	-0.37***	1.01***	0.36***	0.36***	0.21*
	[0.10]	[0.10]	[0.09]	[0.12]	[0.10]	[0.12]
Experience	-0.03***	-0.02***	0.03***	-0.00	0.02*	0.02**
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]
Experience^2	0.00***	0.00**	-0.00***	-0.00	-0.00***	-0.00***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Spanish language	0.31**	0.17	-0.47**	-0.52***	-0.54**	-0.42**
	[0.14]	[0.14]	[0.19]	[0.18]	[0.26]	[0.19]
Black	0.10	0.03	-0.22**	-0.07	0.27**	0.36***
	[0.09]	[0.09]	[0.11]	[0.11]	[0.12]	[0.12]
Hispanic	0.15	0.02	0.02	0.26**	0.04	0.20*
	[0.10]	[0.12]	[0.11]	[0.11]	[0.14]	[0.11]
Asian	-0.33	-0.08	-0.20	-0.23	-0.50**	-0.56**
	[0.26]	[0.20]	[0.32]	[0.26]	[0.19]	[0.23]
240 Occupation dummies	No	Yes	No	Yes	No	Yes
No. Obs.	1,316	1,316	1,316	1,316	1,316	1,316

Note: All models include an intercept term and are weighted by sampling weights. Standard errors in squared brackets. * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.
Source: PDII survey.

Last but not least, we complement our analysis by considering a more representative survey of U.S. workers than that provided by the PDII. The sample we use comes from the March Annual and Social Supplement (ASEC) of the IPUMS-CPS data (King et al., 2010). In particular, the data used in this work covers the period from 1992 to 2010. Each year a set of data is released, providing individual level information on wage and salary income, weeks and hours worked as well as other employment related information for the previous calendar year. The sample includes males and females located within the U.S. borders. To build a sample of individuals comparable to that of the PDII, we retain wage and salaried individuals between 18 and 64 years of age with full information on wages, weeks and weekly hours worked, age, education, self-reported ethnicity and task data. Occupations are classified according to the 1990 census classification scheme up to 2002 and according to the 2002 classification scheme after 2003. To create time-consistent occupation categories, we adopt the classification scheme proposed by Autor and Dorn (2013).[12] Similarly, the industry classification changed twice, in 2003 and 2009. To keep track of these changes, we interact two digit industry dummies with two period dummies, one taking value 1 from year 2003 onwards and 0 otherwise, and the other taking value 1 for years 2009 and 2010 and 0 otherwise.

To this sample, we attach the O*NET items after using readily available crosswalks to match the O*NET SOC2000 occupations to the time-consistent occupation categories (*occ1990dd*). We re-scale the O*NET raw scores to obtain a final score that ranges between 0 and 1: the higher the value is, the more representative the characteristic is of the nature of that job. This is done by calculating the percentile score of each item based on the 2003 sample (Peri and Sparber, 2009). Then, we build the O*NET task measures as described above. Furthermore, we apply the PDII occupation task means after appropriately matching the PDII occupation categories to the IPUMS-CPS time-consistent occupations.

3 Theoretical underpinnings and empirical framework

In the following we draw on Autor and Handel (2013) and extend their conceptual model to account for task returns heterogeneity across genders and to identify gender heterogeneity in comparative advantages.

The main peculiarity of the task approach is that tasks have structurally different characteristics from human capital, as they represent the demand and supply of skills, respectively. Extending the theoretical framework of the Roy (1951) model, occupations are described by an indivisible bundle of tasks that workers have to perform jointly. Workers choose the occupation returning the highest premium to the whole bundle of tasks, and not the highest premium to each of the tasks in that occupation. Moreover, this framework is characterized by self-selection, meaning that workers will choose the occupation yielding high returns to the tasks they are more able to perform given their endowment of human capital, that is tasks in which they have a comparative advantage. Task returns heterogeneity across occupations represents the crucial difference between the task

[12] Autor and Dorn (2013) design a unique classification scheme (*occ1990dd*) and develop a crosswalk with various census classification schemes, including the 1990 and 2005 ones. We use the latter and the readily available crosswalk between the 2002 and 2005 census occupation categories to translate into *occ1990dd* occupation codes from 2003 onwards. We end up with 324 occupations in our final sample of 985,372 workers.

approach and Mincer's human capital framework, in which the marginal productivity of, say, one year of education is invariant across jobs.

More formally, the skill endowment of worker i can be written as the vector $\Phi_i = [\phi_{i1}, \ldots, \phi_{iK}]$ where each element is positive, has continuous support, and represents the ability of worker i in performing task k, for $i = 1, \ldots, n$, $k = 1, \ldots, K$. In order to include gender heterogeneity, we assume that men and women are endowed with different task abilities, that is they are diversely efficient in performing the same task. Therefore, we further differentiate the skill endowments of male and female workers by writing $\Phi_i^G = [\phi_{i1}^G, \ldots, \phi_{iK}^G]$, where $G = m, f$ for male and female workers, respectively. Let the production function of the aggregate output produced by all workers in occupation j be defined as

$$Y_j = \exp\left[\sum_{i \in I_j}\left(\alpha_j + \sum_{G=m,f}\sum_{k=1}^{K}\lambda_{jk}^G\ \phi_{ik}^G + \mu_i\right)\right], \quad \text{for } j = 1, \ldots, J \tag{2}$$

where I_j is the set of workers in occupation j. The task returns $\lambda_{jk}^G \geq 0$ are occupation specific and we further assume that the productive value of tasks differs across genders (Rendall, 2010; Barbulescu and Bidwell, 2013; Grove et al., 2011; Yahmed, 2013). Therefore we allow for two sources of gender heterogeneity: i) male and female workers have different efficiencies at performing the same task, and ii) these task efficiencies are rewarded differently within the same occupation. In this setting, α_j is not constrained to be positive, that is workers, male or female, with an insufficient skill endowment may have a negative marginal product, μ_i is an error term, and the output price is normalized to unity. Workers are paid their marginal product and the log-wage of worker i in occupation j will differ according to the gender

$$w_i^G = \alpha_j + \sum_{k=1}^{K}\lambda_{jk}^G\ \phi_{ik}^G + \mu_i. \tag{3}$$

Workers will choose the occupation maximizing their earnings, that is the one returning the highest premium to the whole bundle of tasks, and not the highest premium to each of the tasks in that occupation. In their model, Autor and Handel (2013) assume that all occupations have non-zero employment. In extending their framework to heterogeneous task returns across genders, we indeed ensure that occupations have non-zero employment over all, however allowing for occupations to have either male or female zero employment. This amounts to assuming that for two occupations $j, j' = 1, \ldots, J$, $j \neq j'$, it is possibe to have $\alpha_{j'} > \alpha_j$ and $\lambda_{j'k}^G > \lambda_{jk}^G$, $k = 1, \ldots, K$ for either $G = m$ or $G = f$, but not both. In practice, there may be an occupation j' with a higher intercept and higher returns to every task with respect to occupation j for either male or female workers, so that they would always choose j' over j.

Expression (3) yields the following empirical model:

$$w_i = \alpha + \sum_{k=1}^{K}\beta_k^m\, T_{ik} + \sum_{k=1}^{K}\beta_k^f fem_i T_{ik} + \psi fem_i + \gamma' X_i + \varepsilon_i \tag{4}$$

where T_{ik} are individual measures of the tasks performed on the job, fem_i is a dummy variable equal to 1 if the worker is female and ψ captures the wage gap independent of gender-specific task rewards. The model is further enriched by individual level controls on demographic characteristics and education achievements, X which are education dummies, ethnicity and language (Autor and Handel, 2013). Finally, ε_i is a zero-mean error term.

The OLS estimator of the β coefficients in model (4) will not retrieve the average task returns in the labour market. Workers will self-select into the occupations according to their comparative advantage, meaning that workers with a higher ability in performing a certain task will choose occupations that have a higher reward for that task. A necessary but not sufficient condition for the data to be consistent with self-selection of workers, male and female, is for the cross-occupation covariance between task returns not to be uniformly positive: for each $j, j' = 1, \ldots, J$, $j \neq j'$, $\mathrm{E}(\lambda^G_{jk}, \lambda^G_{j'h}) \leq 0$, for some k, h, and/or $\mathrm{E}(\alpha_j, \lambda^G_{j'k}) \leq 0$, for some k, for $G = m, f$. The negative covariance between task returns implies that there will be some worker preferring j to j' because it has a higher return to the task k given his/her own task efficiency ϕ^G_{ik}, whereas there will be some other worker with a certain level of ability at performing task h for which the productive value is higher in j' than in j, and will therefore prefer occupation j'. Therefore, the sign of the estimated $\hat{\beta}_k$ could depend on the cross-occupation covariance between tasks returns which could not be uniformly positive, even though the occupation-specific task returns λ_{jk} are non-negative. In other words, if relevant cross occupation correlations among task rewards exists, the estimates of returns to task across occupations could be biased and lead to spurious inference.

In the following Section, we will test for the presence of cross-occupation covariances between tasks returns; were we not to reject the null hypotheses of absence of correlation, then the necessary condition for self-selection would not hold and the OLS estimates may confidently be interpreted as average task returns. We will also derive a suitable model parametrization that will allow us to test for the presence of comparative advantages.

4 Results

We analyse the relationship between wages and person level PDII task measures by regressing workers' hourly wages on their job tasks, human capital, demographic characteristics and occupation dummies as described in equation (4). As previously mentioned, this regression provides a descriptive analysis only and cannot inform on the average returns to jobs' tasks in presence of cross-occupations covariance between task returns. Nevertheless, it can provide interesting insights on the explanatory power of tasks and, for our specific purposes, whether there are any differences between the cross-occupation average returns to the task intensities performed on the job between men and women. We present results of this exercise in Table 3. In the baseline specification of Column (1), we find a negative association between manual task intensity and wages and a positive association between the latter with cognitive and social task intensity. This also holds when we account for demographic characteristics (Column (2)) and within occupation, that is when occupation dummies (Columns (3) and (4)) are included to the baseline specification. When we allow for differences between male and female individuals, we find that the latter, compared to men, enjoy a higher "return" to manual and cognitive tasks and a lower "return" to social tasks. This evidence persists both across and within occupation (Columns (6) and (7)), when demographic controls are accounted for and also when we add 2-digit industry dummies (Column (8)). Across all specifications, the overall effect on women's wages is significant only when considering the manual and cognitive content of jobs, as shown by the test at the bottom of the Table. It is worth highlighting that the manual, cognitive and social nature of a job has a high explanatory power *per se*. As a matter of fact, by looking at the R-squared of each regression, individual tasks alone explain almost one third of the total wage variation. On the contrary, accounting for heterogeneity in task coefficients between men and women only marginally improves the explanatory power of individual tasks. Finally, the coefficient on the female dummy is unaffected by the inclusion of its interaction with the task indicators.

In Table A3 in the Appendix we replicate estimates of model (4) by considering the IPUMS-CPS. It emerges that the PDII and O*NET occupation level measures are still able to map the pattern of wage premia and penalties that characterize manual, cognitive and social tasks. However, both sets of indicators are not able to account for differences between genders. Worthy of note, is that this evidence holds even when controlling for within occupation and industry effects (Columns (7) and (8)), though, in this case, the O*NET measures fail to have a significant explanatory power. It is worth stressing that the IPUMS-CPS sample that we use is substantially large and characterized by a considerable amount of heterogeneity that we are neglecting here as this is not the core of our work. Further research in this direction is warranted.

Table 3: OLS regressions of *Log hourly wages* on PDII person level task measures, demographic variables and occupation dummies.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manual	-0.23***	-0.14***	-0.16***	-0.13***	-0.25***	-0.17***	-0.17***	-0.16***
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.04]	[0.04]
Cognitive	0.19***	0.11***	0.06***	0.05**	0.15***	0.07***	0.02	0.02
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]	[0.03]
Social	0.07***	0.05***	0.04**	0.05***	0.10***	0.09***	0.09***	0.09***
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	[0.02]	[0.02]	[0.02]
Female $\times$ *Manual*					0.05	0.08**	0.07*	0.07*
					[0.04]	[0.04]	[0.04]	[0.04]
Female $\times$ *Cognitive*					0.090**	0.087**	0.060*	0.045
					[0.04]	[0.04]	[0.04]	[0.03]
Female $\times$ *Social*					-0.08**	-0.08**	-0.08**	-0.09***
					[0.04]	[0.03]	[0.03]	[0.03]
Female		-0.28***		-0.11***		-0.28***	-0.10***	-0.10***
		[0.04]		[0.04]		[0.03]	[0.04]	[0.04]
Less than high school		0.05		0.02		0.04	0.01	0.02
		[0.10]		[0.10]		[0.10]	[0.10]	[0.10]
Some college		0.01		0.03		0.00	0.03	0.05
		[0.05]		[0.05]		[0.05]	[0.05]	[0.05]
College		0.20***		0.09**		0.19***	0.09*	0.10**
		[0.05]		[0.05]		[0.05]	[0.05]	[0.05]
Postcollege		0.41***		0.25***		0.40***	0.25***	0.26***
		[0.06]		[0.06]		[0.06]	[0.06]	[0.06]
Experience		0.02***		0.01		0.03***	0.01	0.01
		[0.01]		[0.01]		[0.01]	[0.01]	[0.01]
Experience^2		-0.00***		-0.00		-0.00***	-0.00	-0.00
		[0.00]		[0.00]		[0.00]	[0.00]	[0.00]
Spanish language		-0.39***		-0.23		-0.39***	-0.22	-0.16
		[0.14]		[0.15]		[0.13]	[0.15]	[0.14]
Black		-0.11		-0.14***		-0.10	-0.13***	-0.14***
		[0.07]		[0.05]		[0.07]	[0.05]	[0.05]
Asian		0.11		0.09		0.12	0.10	0.05
		[0.15]		[0.11]		[0.14]	[0.11]	[0.10]
Hispanic		-0.09		-0.06		-0.09	-0.07	-0.08
		[0.07]		[0.07]		[0.07]	[0.07]	[0.07]
240 Occupation Dummies	No	No	Yes	Yes	No	No	Yes	Yes
2-digit Industry Dummies	No	No	No	No	No	No	No	Yes
No. Obs.	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316
R-squared	0.28	0.43	0.63	0.66	0.29	0.44	0.66	0.68
Test of joint significance:								
Manual + *Female* $\times$ *Manual*					-0.20***	-0.10***	-0.10***	-0.10***
					[0.03]	[0.03]	[0.02]	[0.03]
Cognitive + *Female* $\times$ *Cognitive*					0.24***	0.16***	0.08***	0.06**
					[0.03]	[0.03]	[0.03]	[0.03]
Social + *Female* $\times$ *Social*					0.02	0.01	0.01	-0.00
					[0.03]	[0.02]	[0.02]	[0.03]

Note: All models include an intercept term and are weighted by sampling weights. Standard errors in squared brackets. * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.
Source: PDII survey.

(htbp)

Table 4: Averages of task returns

Variable	No. Obs.	Mean	Min	Max	Std. Dev.
Manual	89	-0.18	-6.00	0.92	0.74
Cognitive	89	0.10	-1.82	3.62	0.70
Social	89	0.02	-6.79	2.44	0.89

Source: Authors' own elaborations based on PDII survey.

4.1 Testing for covariance between task returns across occupations

We test for the possible presence of significant cross-occupation correlations between task returns, that would prevent the OLS estimator of β_k in (4) from recovering the average task returns.

To this purpose, we check whether the sign and significance of baseline task returns estimates actually represent average task returns across occupations. Similarly to Autor and Handel (2013), we estimate task returns by occupation and check whether the averages of estimated coefficients actually mimic the ones we find by pooling occupations altogether. More specifically, to avoid that within occupations task returns may reflect particular education or demographic features of workers performing specific tasks, we first regress individual wages on demographic and education characteristics to purge heterogeneity in individual wages from any observable factor other than labour tasks. Then, to preserve sufficient degree of freedom, we regress the residual wages $\widehat{w}_i$ on task measures by occupation j for each of the 89 occupations with at least 5 observations according to the following specification:

$$\widehat{w}_{ij} = \alpha_j + \beta_{j1} Manual + \beta_{j2} Cognitive + \beta_{j3} Social + \varepsilon_{ij} \tag{5}$$

Corresponding averages of parameter estimates are reported in Table 4. Here the size of average task returns across occupations substantially reflect task returns estimates obtained from our baseline empirical model. This implies that by estimating model (4) we are retrieving a sensible estimate of task returns across occupations. However, to further inspect this issue, we display cross occupation variations in task returns for each task combination in Figure 1. From the picture no particular pattern emerges, that is task returns appear neither positively nor negatively correlated.

Finally, we use parameter estimates from model (5) and run separate bivariate regressions for the elements of $\alpha_j, \hat{\beta}_{1j}, \hat{\beta}_{2j}, \hat{\beta}_{3j}$ on one another, in all cases weighting by the sum of worker's weights within an occupation: $\hat{\beta}_{j1} = \alpha_1 + \lambda_1 \hat{\beta}_{j2} + e_{12}$, $\hat{\beta}_{j1} = \alpha_2 + \lambda_2 \hat{\beta}_{j3} + e_{13}$, $\hat{\beta}_{j2} = \alpha_3 + \lambda_3 \hat{\beta}_{j3} + e_{23}$, $\hat{\alpha}_j = \alpha_4 + \lambda_4 \hat{\beta}_{j1} + e_{01}$, $\hat{\alpha}_j = \alpha_5 + \lambda_5 \hat{\beta}_{j2} + e_{02}$ and $\hat{\alpha}_j = \alpha_6 + \lambda_6 \hat{\beta}_{j3} + e_{03}$. Corresponding estimates are presented in Table 5 where bivariate regressions of task returns on one another deliver no significant covariance between each couple of task returns. Also, from the last column of the Table, a significantly higher $\hat{\alpha}_j$ is recorded for highly social jobs only, while the higher the manual intensity the lower $\hat{\alpha}_j$. However, this coefficient is not significant at all. From the above evidence we conclude that, in this specific setting, we can confidently interpret our baseline model task coefficient estimates as returns to tasks.

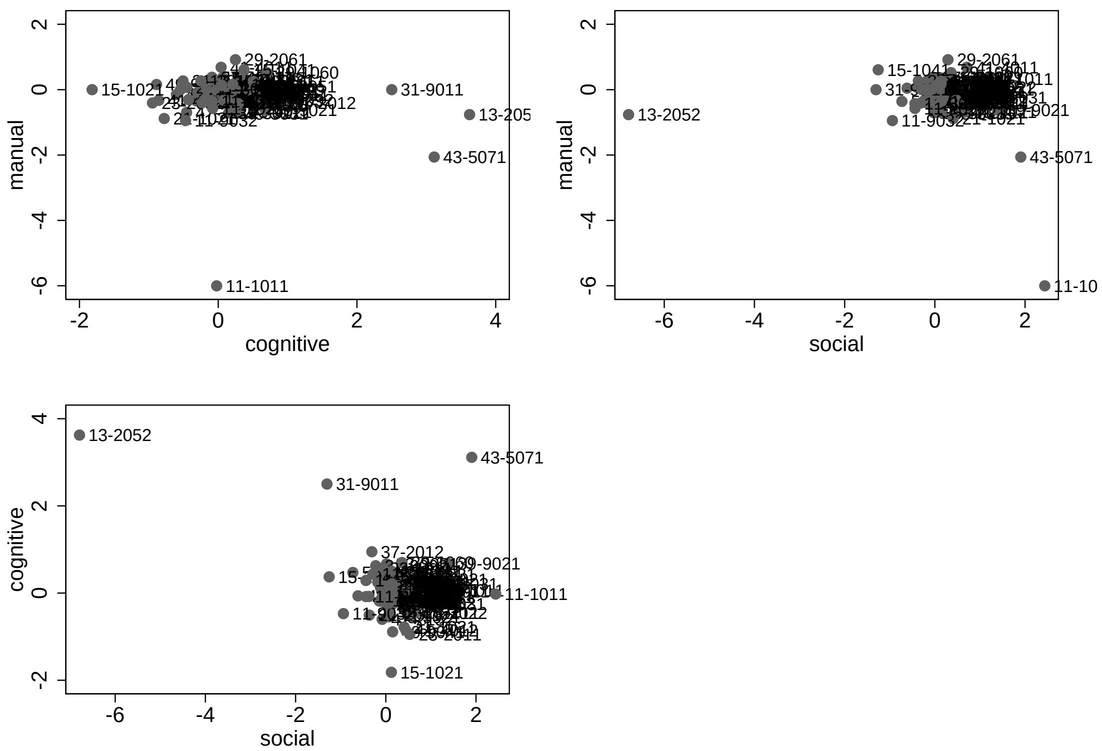
Table 5: Bivariate relationship among regression coefficients obtained from occupation level wage regression models: OLS estimates.

	(1) $\beta_{Cognitive}$	(2) β_{Social}	(3) $Intercept$
$\beta_{Cognitive}$			0.39
			[0.79]
β_{Social}	-0.16		0.95***
	[0.31]		[0.34]
β_{Manual}	-0.40	-0.25	-0.49
	[0.32]	[0.21]	[0.82]

Note: Standard errors in squared brackets. * p<0.10, ** p<0.05 and *** p<0.01. In the above estimates 89 occupations over 240 are considered.
Source: PDII survey.

(!htbp)

Figure 1: Cross occupation variations in task returns.



Source: Authors' own elaborations based on PDII survey.

4.2 Testing for comparative advantages

As argued in Section 3, the negative cross-occupation covariance among tasks is a necessary condition for the data to be consistent with workers' self-selection into occupations. As we fail to reject the absence of covariance bewteen tasks, these data, together with our particular choice of task measures, should not be reflecting the presence of workers' comparative advantages in perfoming certain job tasks.

In order to further corroborate our finding, we estimate an augmented wage regression where we include interaction of individual and occupation level task indicators according to the following model (Autor and Handel, 2013):

$$w_i = \alpha+\sum_{k=1}^{K} \beta_k^m T_{ik}+\sum_{k=1}^{K} \beta_k^f fem_i T_{ik}+\sum_{k=1}^{K} \delta_k \overline{T}_{jk}+\sum_{k=1}^{K} \gamma_k^m T_{ik}\overline{T}_{jk}+\sum_{k=1}^{K} \gamma_k^f fem_i T_{ik}\overline{T}_{jk}+\psi fem_i+\varepsilon_i \tag{6}$$

where $\overline{T}_{jk} = \frac{1}{|I_j|}\sum_{i\in I_j} T_{ik}$ is the average requirement of task k in occupation j.

Corresponding results are shown in Columns (1) and (2) of Table 6. In Column (1) interaction terms are never significant both individually and jointly. When heterogeneity across genders is accounted for in Column (2), the joint significance of interactions of individual times occupation level task measures for women is not rejected as from standard testing displayed at the bottom of the Table.

These findings corroborate the absence of self-selection and no evidence of workers comparative advantages in certain job tasks emerge.

Turning to task returns, the specification of model (6) offers a parametrisation in which individual tasks vary by the occupation level task content of occupations. Table 7 presents task returns for male and female workers at the lower decile, the median and the top decile of the three task distributions. Returns from manual tasks are always negative both for men and women and the latter are not differently rewarded with respect to the former, regardless the manual intensity level of a job. A different story is recorded for the evolution of task returns across the ladder of *Cognitive* intensity. Here women always enjoy a wage premium which is increasing with the cognitive intensity of the job. Returns from cognitive task for men, instead, are positive although appear to be declining. along the distribution of *Cognitive* Finally, as the social intensity of jobs increases women experience a slightly declining job penalty and men a slightly declining wage premium.

Table 6: Testing for comparative advantages: OLS regressions of *Log hourly wages* on PDII person and occupation level task measures and demographic variables.

	(1)	(2)
Manual	-0.164***	-0.164***
	[0.038]	[0.037]
Cognitive	0.009	0.008
	[0.031]	[0.031]
Social	0.089***	0.090***
	[0.027]	[0.027]
Female $\times$ *Manual*	0.082*	0.083*
	[0.043]	[0.043]
Female $\times$ *Cognitive*	0.089***	0.094***
	[0.033]	[0.034]
Female $\times$ *Social*	-0.078**	-0.079**
	[0.031]	[0.033]
$Manual_{occ}$	0.023	0.023
	[0.044]	[0.044]
$Cognitive_{occ}$	0.166***	0.168***
	[0.053]	[0.053]
$Social_{occ}$	-0.012	-0.013
	[0.045]	[0.046]
$Manual \times Manual_{occ}$	0.017	-0.000
	[0.039]	[0.049]
$Cognitive \times Cognitive_{occ}$	-0.023	-0.029
	[0.028]	[0.034]
$Social \times Social_{occ}$	0.018	0.014
	[0.028]	[0.039]
$Female \times Manual \times Manual_{occ}$		0.036
		[0.069]
$Female \times Cognitive \times Cognitive_{occ}$		0.015
		[0.046]
$Female \times Social \times Social_{occ}$		0.010
		[0.052]
No. Obs.	1,316	1,316
R-squared	0.452	0.453
F-test p-value:		
$Manual \times Manual_{occ} = Cognitive \times Cognitive_{occ} = Social \times Social_{occ} = 0$	0.75	0.83
$Female \times (Manual \times Manual_{occ} = Cognitive \times Cognitive_{occ} = Social \times Social_{occ}) = 0$		0.91

Note: All models include an intercept term, education and demographic controls and are weighted by sampling weights. Standard errors in squared brackets. * $p<0.10$, ** $p<0.05$ and *** $p<0.01$. occ = occupation level task measure.
Source: PDII survey and O*NET 3.0 Database.

Table 7: Testing task returns along the task intensity of occupations

Manual		Cognitive		Social	
		10^{th} percentile			
M: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	-0.17***	M: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.08*	M: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	0.102***
	[0.05]		[0.04]		[0.04]
F: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	-0.010	F: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.00*	F: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	-0.11**
	[0.01]		[0.00]		[0.05]
$M + F$	-0.18***	$M + F$	0.08*	$M + F$	-0.01
	[0.05]		[0.04]		[0.04]
		50^{th} percentile			
M: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	-0.16***	M: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.07**	M: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	0.07***
	[0.04]		[0.03]		[0.03]
F: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	0.02	F: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.05*	F: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	-0.10***
	[0.02]		[0.02]		[0.03]
$M + F$	-0.14***	$M + F$	0.11***	$M + F$	-0.02
	[0.04]		[0.04]		[0.02]
		90^{th} percentile			
M: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	-0.150***	M: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.05	M: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	0.05
	[0.05]		[0.05]		[0.03]
F: $\beta_{Manual} + \gamma_{Manual} \times Manual_{\text{O*NET}}$	0.10	F: $\beta_{Cognitive} + \gamma_{Cognitive} \times Cognitive_{\text{O*NET}}$	0.09*	F: $\beta_{Social} + \gamma_{Social} \times Social_{\text{O*NET}}$	-0.08*
	[0.06]		[0.05]		[0.04]
$M + F$	-0.06	$M + F$	0.14**	$M + F$	-0.03
	[0.07]		[0.06]		[0.04]

Note: M: Male sample; F: Female sample; $M + F$: Overall sample. *p<0.10, ** p<0.05 and *** p<0.01.
Source: PDII survey.

4.3 Robustness checks

We present several robustness checks in Table 8 and 9. First, we add to our baseline specification a set of additional explanatory variables capturing the individual's family characteristics. The PDII survey collects information on marital status and on household composition. Based on this data, we build the dummy variables *Married* and *Children*. As shown in Columns (1) and (2) of Table 8, our main results hold when including these additional individual controls. Our findings are confirmed even when we add the interaction terms *Female* $\times$ *Married* and *Female* $\times$ *Children*. Noteworthy is that married women receive a wage penalty compared to their male counterparts. Cohabiting children, instead, don't seem to influence individual wages neither for men nor for women.

Second, we split the sample according to the respondent's gender. Results in Column (5) and (6) of Table 8 are consistent with our baseline results in Column (8) of Table 3. The same pattern emerges when including family related individual characteristics and also when adding a measure of tenure (Columns (5) to (10)), that is number of months worked in the same job.

Last, we estimate model (4) for each task measure and then for each of their pairwise combinations to control for any correlation among them. Results of this exercise are shown in Table 9. The coefficients associated with each task measure are comparable in magnitude and significance to those of our corresponding baseline specifications (Column (7) of Table 3). This suggests that, even within occupations, task indicators are independent and capture different job dimensions.

Table 8: Robustness check: Family characteristics and sub-samples by gender.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	ALL	ALL	ALL	ALL	MALE	FEMALE	MALE	FEMALE	MALE	FEMALE
Manual	-0.13***	-0.13***	-0.16***	-0.15***	-0.20***	-0.08***	-0.20***	-0.08***	-0.20***	-0.10***
	[0.02]	[0.02]	[0.04]	[0.04]	[0.05]	[0.03]	[0.05]	[0.03]	[0.05]	[0.03]
Cognitive	0.11***	0.05**	0.02	0.02	0.03	0.07**	0.04	0.07**	0.07**	0.06*
	[0.02]	[0.02]	[0.03]	[0.03]	[0.04]	[0.03]	[0.04]	[0.03]	[0.03]	[0.03]
Social	0.05***	0.05***	0.08***	0.09***	0.09***	-0.02	0.09***	-0.02	0.05*	0.00
	[0.02]	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]
Female × Manual			0.07*	0.04						
			[0.04]	[0.04]						
Female × Cognitive			0.06*	0.07*						
			[0.03]	[0.04]						
Female × Social			-0.08**	-0.08**						
			[0.03]	[0.03]						
Female	-0.27***	-0.11***	-0.10***	0.02						
	[0.04]	[0.04]	[0.04]	[0.07]						
Less than high school	0.05	0.01	0.01	0.01	-0.29*	0.17	-0.27	0.17	-0.08	0.05
	[0.10]	[0.10]	[0.10]	[0.10]	[0.17]	[0.14]	[0.17]	[0.14]	[0.19]	[0.13]
Some college	0.01	0.03	0.03	0.02	0.12	0.01	0.11	0.01	0.10	0.01
	[0.05]	[0.05]	[0.05]	[0.05]	[0.09]	[0.08]	[0.09]	[0.08]	[0.09]	[0.08]
College	0.20**	0.10**	0.09**	0.10**	0.15**	0.04	0.15**	0.04	0.09	0.06
	[0.05]	[0.05]	[0.05]	[0.05]	[0.08]	[0.07]	[0.08]	[0.07]	[0.08]	[0.06]
Postcollege	0.41***	0.26***	0.27***	0.27***	0.21**	0.30***	0.23**	0.31***	0.21**	0.32***
	[0.06]	[0.06]	[0.06]	[0.06]	[0.10]	[0.09]	[0.10]	[0.09]	[0.10]	[0.09]
Experience	0.02***	0.00	0.00	0.00	0.01	-0.01	0.01	-0.01	0.01	-0.01
	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]
Experience^2	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Tenure									0.00**	0.00*
									[0.00]	[0.00]
Tenure^2									-0.00*	-0.00
									[0.00]	[0.00]
Spanish language	-0.39***	-0.25	-0.23	-0.24	-0.19	0.00	-0.22	-0.01	-0.23	-0.27
	[0.14]	[0.15]	[0.15]	[0.15]	[0.21]	[0.20]	[0.21]	[0.20]	[0.19]	[0.21]
Black	-0.10	-0.14***	-0.13**	-0.13**	-0.37***	-0.10	-0.36***	-0.10	-0.21*	-0.10
	[0.06]	[0.05]	[0.05]	[0.05]	[0.11]	[0.08]	[0.10]	[0.08]	[0.11]	[0.08]
Hispanic	-0.08	-0.05	-0.07	-0.07	0.01	-0.18**	0.02	-0.18**	0.14	-0.08
	[0.07]	[0.07]	[0.07]	[0.06]	[0.10]	[0.08]	[0.10]	[0.08]	[0.11]	[0.07]
Asian	0.10	0.09	0.10	0.10	0.32*	0.00	0.27	0.00	0.18	0.01
	[0.14]	[0.11]	[0.11]	[0.11]	[0.19]	[0.12]	[0.19]	[0.12]	[0.21]	[0.16]
Married	0.09**	0.03	0.04	0.15***			0.11	-0.02		
	[0.04]	[0.04]	[0.04]	[0.05]			[0.07]	[0.07]		
Children	0.07	0.08**	0.08**	0.07			0.05	0.03		
	[0.05]	[0.04]	[0.04]	[0.05]			[0.06]	[0.06]		
Female × Married				-0.19***						
				[0.07]						
Female × Children				0.00						
				[0.08]						
240 Occupation Dummies	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2-digit Industry Dummies	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	1,315	1,315	1,315	1,315	592	724	592	723	503	642
R-squared	0.44	0.66	0.66	0.67	0.77	0.67	0.77	0.67	0.82	0.74

Note: All models include an intercept term and are weighted by sampling weights. Standard errors in squared brackets. * $p<0.10$, ** $p<0.05$ and *** $p<0.01$. *Children* is a dummy variable taking value 1 if there are any children living in the household at the time of the interview.
Source: PDII survey.

Table 9: Robustness: Individual task measures and pairwise combinations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Manual	-0.13***				-0.13***	-0.12***	-0.16***				-0.17***	-0.16***
	[0.02]				[0.02]	[0.02]	[0.04]				[0.04]	[0.04]
Cognitive		0.07***		0.06***		0.06***		0.05*		0.05*		0.04
		[0.02]		[0.02]		[0.02]		[0.03]		[0.027]		[0.03]
Social			0.03	0.01	0.04**				0.06**	0.05*	0.07***	
			[0.02]	[0.02]	[0.02]				[0.03]	[0.03]	[0.03]	
Female × Manual							0.05				0.07*	0.06
							[0.04]				[0.04]	[0.04]
Female × Cognitive								0.03		0.03		0.04
								[0.03]		[0.03]		[0.03]
Female × Social									-0.06*	-0.06*	-0.07*	
									[0.03]	[0.03]	[0.03]	
Female	-0.12***	-0.12***	-0.12***	-0.12***	-0.11***	-0.11***	-0.12***	-0.12***	-0.12***	-0.12***	-0.11***	-0.12***
	[0.04]	[0.04]	[0.04]	[0.04]	[0.04]	[0.04]	[0.03]	[0.04]	[0.04]	[0.04]	[0.04]	[0.04]
Less than high school	0.02	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.01	0.02
	[0.10]	[0.10]	[0.10]	[0.10]	[0.098]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]	[0.10]
Some college	0.04	0.02	0.03	0.02	0.041	0.03	0.04	0.02	0.03	0.02	0.04	0.03
	[0.05]	[0.05]	[0.05]	[0.05]	[0.051]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]
College	0.10**	0.11**	0.12**	0.11**	0.095**	0.10**	0.10**	0.11**	0.11**	0.11**	0.09*	0.09*
	[0.05]	[0.05]	[0.05]	[0.05]	[0.047]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]
Postcollege	0.28***	0.30***	0.32***	0.30***	0.272***	0.26***	0.28***	0.30***	0.33***	0.30***	0.27***	0.25***
	[0.06]	[0.06]	[0.06]	[0.06]	[0.060]	[0.06]	[0.06]	[0.07]	[0.06]	[0.06]	[0.06]	[0.06]
Experience	0.01	0.01**	0.01*	0.01**	0.009	0.01	0.01	0.01**	0.01**	0.01**	0.01	0.01
	[0.01]	[0.00]	[0.00]	[0.00]	[0.006]	[0.00]	[0.01]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Experience^2	-0.00	-0.00	-0.00	-0.00	-0.000	-0.00	-0.00	-0.00	-0.000	-0.00	-0.00	-0.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.000]	[0.00]	[0.00]	[0.00]	[0.000]	[0.00]	[0.00]	[0.00]
Spanish language	-0.28*	-0.26*	-0.29*	-0.26*	-0.261*	-0.25	-0.27*	-0.26*	-0.29*	-0.26*	-0.26*	-0.24
	[0.16]	[0.16]	[0.16]	[0.16]	[0.155]	[0.15]	[0.16]	[0.16]	[0.16	[0.15]	[0.15]	[0.15]
Black	-0.13***	-0.13**	-0.14***	-0.13**	-0.140***	-0.13**	-0.13**	-0.13**	-0.14***	-0.13**	-0.14***	-0.12**
	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]
Hispanic	-0.04	-0.06	-0.04	-0.06	-0.04	-0.05	-0.04	-0.06	-0.05	-0.06	-0.05	-0.06
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]
Asian	0.05	0.08	0.07	0.08	0.067	0.07	0.04	0.08	0.09	0.10	0.07	0.06
	[0.11]	[0.11]	[0.11]	[0.11]	[0.107]	[0.11]	[0.10]	[0.11]	[0.11]	[0.11]	[0.10]	[0.11]
No. Obs.	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316	1,316
R-squared	0.65	0.64	0.64	0.64	0.65	0.66	0.65	0.64	0.64	0.65	0.66	0.66

Note: All models include an intercept term and 240 occupation dummies and are weighted by sampling weights. Standard errors in squared brackets. * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.
Source: PDII survey and O*NET 3.0 Database.

5 Conclusions

In this paper we explore the existence of gender differences in returns to motor, cognitive and social tasks. We have therefore adopted a task based analysis of heterogenous wage returns to gender. In this framework, we have extended previous contributions by considering the social character of a job beside the motor and cognitive ones. Moreover, we have acknowledged that tasks are rewarded differently across genders.

We confidently interpret our estimates as task returns, since we fail to reject the hypothesis of cross-occupation variation in task returns that would render the data consistent with workers' self-selection. Given this evidence, we identify a wage premium for women performing highly manual or highly cognitive jobs. However, a wage penalty characterises women engaged in highly social intensive jobs. These results hold both across- and within-occupations as well as when we control for within-industry effects.

Though our findings may seem at odds with the literature on the decreasing value of manual tasks on the one hand, and increasing value of social job tasks on the other, it is worth stressing that we find this pattern of task returns in our "gender-blind" estimations too. However, once we account for heterogenous returns across gender, we make an important discovery: we provide evidence of a lower return to women employed in highly social intensive jobs with respect to men engaged in the same kind of activities.

This result could well depend on our definition of social tasks which incorporates the managerial and team work content of a job. It is a well known fact that the gender wage gap widens along the job ladder as women in leadership positions are valued less than their male counterparts. As an example, using information on the five highest paid executives in each of a large number of U.S. firms, Bertrand and Hallock (2001) finds that women earn 45% less than men. They explain part of this wage gap by the fact that women manage smaller companies and that they are less likely to cover top managerial positions. In a similar vein, Smith et al. (2011) document that there still exists a large gender compensation gap among top executives in Denmark, even when controlling for observed individual and firm characteristics as well as unobserved individual heterogeneity.

Although we are not able to corroborate this evidence in the present study, we see a couple of avenues for future research. First, it would be interesting to breakdown the social component of tasks. Deming (2015) and Borghans et al. (2008) acknowledge the existence of an additional dimension of jobs tasks capturing the degree of service orientation of a job. In particular, Borghans et al. (2008) show that both *directness* and *caring* are two relevant aspects of social tasks and they are positively and independently rewarded from manual and cognitive tasks. This could be particularly informative in our context as women tend to be more concentrated in traditional service jobs, e.g. nurses and teachers, rather than in managerial positions. Second, it would be worth investigating the complementarity between cognitive and social tasks which has been found to be increasingly rewarded on the labour market, as recently documented by Weinberger (2014). Finally, an ideal extension of our work would be to apply our theoretical framework to a larger and more representative sample of workers with information on job tasks at the individual level.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics, Volume 4*, pages 1043–1171. Amsterdam: Elsevier- North Holland.

Autor, D., Levy, F., and Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the u.s. labor market. *American Economic Review*, 5(103):1533–1597.

Autor, D. H. and Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31(S1):S59 – S96.

Bacolod, M. P. and Blum, B. S. (2010). Two sides of the same coin: U.S. "residual" inequality and the gender gap. *Journal of Human Resources*, 45(1):197–242.

Bailey, M. J. and Dynarski, S. M. (2011). Gains and gaps: Changing inequality in u.s. college entry and completion. Working Paper 17633, National Bureau of Economic Research.

Barbulescu, R. and Bidwell, M. (2013). Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science*, 24(3):737–756.

Bertrand, M. and Hallock, K. F. (2001). The gender gap in top corporate jobs. *Industrial and Labor Relations Review*, 55(3):pp. 3–21.

Black, S. E. and Spitz-Oener, A. (2010). Explaining Women's Success: Technological Change and the Skill Content of Women's Work. *The Review of Economics and Statistics*, 92(1):187–194.

Boehm, M. (2015). The price of polarization: Estimating task prices under routine-biased technical change. Mimeo, University of Bonn.

Borghans, L., Duckworth, A. L., Heckman, J. J., and ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Borghans, L., ter Weel, B., and Weinberg, B. A. (2006). People people: Social capital and the labor-market outcomes of underrepresented groups. Working Paper 11985, National Bureau of Economic Research.

Borghans, L., ter Weel, B., and Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *Industrial & Labor Relations Review*, 67(2):287–334.

Deming, D. J. (2015). The growing importance of social skills in the labor market. Working Paper 21473, National Bureau of Economic Research.

Eckel, C. and Grossman, P. (2008). Differences in the economic decisions of men and women: experimental evidence. In Plott, C. and Smith, V., editors, *Handbook of Experimental Economics Results*, volume 1, pages 509–519. Amsterdam: Elsevier- North Holland.

England, P. (1982). The failure of human capital theory to explain occupational sex segregation. *Journal of Human Resources*, 17(3):358–70.

Fortin, N. M. (2008). The gender wage gap among young adults in the united states: The importance of money vs. people. *Journal of Human Resources*, 43:886–920.

Frank, L. (2015). How the gender pay gap widens as women get promoted. Harvard Business Review.

Gathmann, C. and Schonberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1).

Grove, W. A., Hussey, A., and Jetter, M. (2011). The gender pay gap beyond human capital: Heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, 46(4):827–874.

Ingram, B. F. and Neumann, G. (2006). The returns to skill. *Labour Economics*, 13(1):pp. 35–59.

Kamas, L. and Preston, A. (2015). Can social preferences explain gender differences ineconomic behavior? *Journal Economic Behavior & Organization*, 116:318–332.

Khachatryana, K., Dreberb, A., von Essenc, E., and Ranehilld, E. (2015). Gender and preferences at a young age: Evidence from armenia. *Journal Economic Behavior & Organization*, 118:318–332.

King, M., Steven, R., Alexander, J. T., Flood, S., Genadek, K., Schroeder, M. B., Trampe, B., and Rebecca, V. (2010). Integrated public use microdata series, current population survey: Version 3.0. [Machine-readable database] Minneapolis: University of Minnesota.

Krueger, A. B. and Schkade, D. (2008). Sorting in the labor market. do gregarious workers flock to interactive jobs? *Journal of Human Resources*, 43(4):859–883.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Rendall, M. (2010). Brain vs. brawn: The realization of women's comparative advantage. Technical report, Univerisity of Zurich.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Smith, N., Smith, V., and Verne, M. (2011). The gender pay gap in top corporate jobs in denmark: Glass ceilings, sticky floors or both? *International Journal of Manpower*, 32(2):156–177.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *The Review of Economics and Statistics*, 96(5):849–861.

Welch, F. (2000). Growth in women's relative wages and in inequality among men: One phenomenon or two? *The American Economic Review*, 90(2):444–449.

Yahmed, S. B. (2013). Gender gaps across skills and trade openess. Technical report, Aix-Marseille Schoof of Economics.

Yamaguchi, S. (2015). Changes in returns to task-specific skills and gender wage gap. *Journal of Human Resources*, forthcoming.

A Additional tables

Table A1: Definition of tasks from PDII.

Cognitive					**Social**				
How often solve problems taking more than 30 minutes to solve?					**Time spent supervising or managing others on the job**				
	Total	Percentage	Men	Women		Total	Percentage	Men	Women
never	71	5%	39%	61%	almost all the tim	665	51%	39%	61%
less than once a month	172	13%	34%	66%	more than half the	256	19%	51%	49%
at least once per month	112	9%	30%	70%	less than half the	136	10%	51%	49%
at least once a week	360	27%	46%	54%	almost none of the	259	20%	51%	49%
one or more times every day	601	46%	51%	49%					
How often solve problems using advanced mathematics					**Most of the work performed in a group (office staff, assembly line) or alone**				
	Total	Percentage	Men	Women		Total	Percentage	Men	Women
never	793	60%	37%	63%	mostly in a unit/g	505	38%	46%	54%
less than once a month	135	10%	50%	50%	varies/depends	89	7%	43%	57%
at least once per month	81	6%	68%	32%	mostly by myself	722	55%	45%	55%
at least once a week	115	9%	60%	40%					
one or more times every day	192	15%	55%	45%					
The longest document typically read as part of the job					**Face to Face contact with people other than colleagues or supervisors**				
never read at job	44	3%	41%	59%		Total	Percentage	Men	Women
one page or less	328	25%	46%	54%	none at all	145	11%	41%	59%
2 to 5 pages	401	30%	41%	59%	a little	242	18%	52%	48%
6 to 10 pages	185	14%	41%	59%	a moderate amount	209	16%	54%	46%
11 to 25 pages	121	9%	43%	57%	a lot	720	55%	41%	59%
more than 25 pages	237	18%	54%	46%					
					Manual				
					Time spent performing physical tasks on the job				
						Total	Percentage	Men	Women
					almost none of the time	470	36%	45%	55%
					less than half the time	166	13%	34%	66%
					more than half the time	166	13%	48%	52%
					almost all the time	514	39%	47%	53%
Total	1,316	100%	45%	55%	Total	1,316	100%	45%	55%

Source: Author's own calculations based on PDII Survey.

Table A2: Correlations between PDII and O*NET items, PDII and O*NET 3.0.

PDII person level tasks and O*NET occupation level task measures

	PDII *(person level)*								O*NET			AH	
	Social	*Cognitive*	*Manual*	*Abstract*	*Routine*	y_1	y_2	y_3	*Manual*	*Cognitive*	*Social*	*Abstract*	*Routine*
PDII *Social (person level)*	1												
PDII *Cognitive (person level)*	0.17	1											
PDII *Manual (person level)*	0.04	-0.33	1										
PDII *Abstract (person level)*	0.42	0.94	-0.31	1									
PDII *Routine (person level)*	-0.21	-0.12	-0.08	-0.16	1								
y_1	0.04	-0.17	**0.47**	-0.15	-0.04	1							
y_2	**0.32**	0.22	-0.28	0.31	-0.07	-0.11	1						
y_3	0.05	**0.40**	-0.36	0.41	0.07	0.07	0.25	1					
O*NET *Manual*	-0.04	-0.24	**0.49**	-0.25	-0.03	**0.96**	-0.28	-0.01	1				
O*NET *Cognitive*	0.17	**0.38**	-0.43	0.43	0.01	-0.09	0.58	**0.87**	-0.19	1			
O*NET *Social*	**0.21**	0.30	-0.46	0.36	-0.05	-0.45	**0.76**	0.24	-0.57	0.50	1		
AH Abstract	0.22	0.39	-0.48	**0.45**	0.04	-0.34	0.83	0.60	-0.49	0.80	0.74	1	
AH Routine	-0.21	-0.33	0.47	-0.39	**0.02**	0.56	-0.70	-0.16	0.68	-0.40	-0.80	-0.71	1
AH Manual	-0.09	-0.24	**0.48**	-0.25	-0.05	0.88	-0.41	-0.06	0.91	-0.28	-0.63	-0.58	0.73

PDII occupation level task means and O*NET occupation level task measures

	PDII *(occ mean)*								O*NET			AH	
	Social	*Cognitive*	*Manual*	*Abstract*	*Routine*	y_1	y_2	y_3	*Manual*	*Cognitive*	*Social*	*Abstract*	*Routine*
PDII *Social (occ mean)*	1												
PDII *Cognitive (occ mean)*	0.11	1											
PDII *Manual (occ mean)*	0.04	-0.49	1										
PDII *Abstract (occ mean)*	0.38	0.94	-0.45	1									
PDII *Routine (occ mean)*	-0.19	0.02	-0.15	-0.05	1								
y_1	0.13	-0.19	**0.57**	-0.12	-0.12	1							
y_2	**0.42**	0.21	-0.34	0.34	-0.01	-0.05	1						
y_3	0.11	**0.50**	-0.43	0.52	0.11	0.08	0.24	1					
O*NET *Manual*	0.03	-0.26	**0.59**	-0.23	-0.11	**0.96**	-0.20	0.01	1				
O*NET *Cognitive*	0.24	**0.47**	-0.53	0.53	0.10	-0.06	0.56	**0.87**	-0.15	1			
O*NET *Social*	**0.28**	0.32	-0.56	0.39	0.02	-0.40	**0.75**	0.22	-0.50	0.49	1		
AH Abstract	0.30	0.44	-0.57	**0.52**	0.10	-0.30	0.84	0.57	-0.44	0.77	0.72	1	
AH Routine	-0.23	-0.38	**0.61**	-0.43	-0.04	0.55	-0.68	-0.22	0.65	-0.46	-0.77	-0.73	1
AH Manual	-0.05	-0.26	0.61	-0.25	-0.15	0.89	-0.32	-0.04	0.92	-0.23	-0.54	-0.53	0.68

Note: Task indicators with the *occ mean* label are calculated as the occupation level mean of the person level PDII items.
Source: PDII survey and O*NET 3.0 Database. occ = occupation, AH = Acemoglu and Autor (2011) and Autor and Handel (2013).

Table A3: OLS regressions of *Log hourly wages* on PDII and O*NET occupation level task measures, demographic variables, industry and occupation dummies.

	(1) O*NET	(2) PDII	(3) O*NET	(4) PDII	(5) O*NET	(6) PDII	(7) O*NET	(8) PDII
Manual	-0.05**	-0.11***	-0.03*	-0.13***	-0.04*	-0.08**	-0.015	-0.038*
	[0.02]	[0.03]	[0.02]	[0.02]	[0.02]	[0.03]	[0.020]	[0.022]
Cognitive	0.12***	0.19***	0.13***	0.19***	-0.03	0.13***	-0.020	0.101***
	[0.02]	[0.03]	[0.02]	[0.03]	[0.02]	[0.04]	[0.019]	[0.031]
Social	0.05**	0.05**	0.05**	0.05*	0.12***	0.03	0.094***	0.061***
	[0.02]	[0.02]	[0.02]	[0.03]	[0.02]	[0.02]	[0.017]	[0.016]
Female × Manual			-0.03	0.03	0.01	0.04*	0.009	0.011
			[0.03]	[0.04]	[0.01]	[0.02]	[0.010]	[0.015]
Female × Cognitive			-0.00	0.02	-0.01	-0.01	-0.016	-0.023
			[0.03]	[0.05]	[0.01]	[0.02]	[0.011]	[0.016]
Female × Social			0.01	-0.01	-0.03*	-0.06***	-0.019	-0.046***
			[0.02]	[0.04]	[0.02]	[0.02]	[0.012]	[0.014]
Female	-0.32***	-0.27***	-0.32***	-0.27***	-0.22***	-0.24***	-0.196***	-0.211***
	[0.02]	[0.02]	[0.02]	[0.02]	[0.01]	[0.01]	[0.008]	[0.008]
Less than high school	-0.19***	-0.19***	-0.18***	-0.19***	-0.17***	-0.17***	-0.151***	-0.153***
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.008]	[0.008]
Some college	0.07***	0.07***	0.07***	0.06***	0.06***	0.07***	0.054***	0.056***
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.005]	[0.005]
College	0.27***	0.25***	0.27***	0.25***	0.22***	0.22***	0.186***	0.185***
	[0.02]	[0.03]	[0.02]	[0.02]	[0.02]	[0.02]	[0.014]	[0.014]
Post college	0.57***	0.52***	0.57***	0.51***	0.47***	0.46***	0.411***	0.407***
	[0.04]	[0.04]	[0.04]	[0.04]	[0.03]	[0.03]	[0.021]	[0.020]
Experience	0.04***	0.04***	0.04***	0.04***	0.03***	0.03***	0.029***	0.029***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.001]	[0.001]
Experience^2	-0.07***	-0.06***	-0.07***	-0.06***	-0.05***	-0.05***	-0.047***	-0.047***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.002]	[0.002]
Black	-0.10***	-0.08***	-0.10***	-0.08***	-0.07***	-0.08***	-0.070***	-0.070***
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.006]	[0.006]
Asian	-0.03*	-0.01	-0.03*	-0.01	-0.01	-0.02*	-0.008	-0.010
	[0.02]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.007]	[0.007]
Hispanic	-0.08***	-0.08***	-0.08***	-0.08***	-0.06***	-0.06***	-0.052***	-0.053***
	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.01]	[0.006]	[0.005]
Industry dummies (4-digit)	No	No	No	No	No	No	Yes	Yes
Occupation dummies (2-digit)	No	No	No	No	Yes	Yes	Yes	Yes
No. Obs.	980,676	980,676	980,676	980,676	980,676	980,676	980,676	980,676
R-squared	0.313	0.327	0.314	0.327	0.357	0.356	0.389	0.388

Note: All models include an intercept term. Standard errors clustered at the 3 digit occupation level in squared brackets. * p<0.10, ** p<0.05 and *** p<0.01.
Source: IPUMS-CPS survey.