

Occupational Social Value and Returns to Long Hours

By: [Dora Gicheva](#)

This is the peer reviewed version of the following article:

Gicheva, D. (2020). Occupational social value and returns to long hours. *Economica*, 87(347), 682-712.

which has been published in final form at <http://dx.doi.org/10.1111/ecca.12322>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

Abstract:

This paper examines the phenomenon of uncompensated long hours in jobs with pro-social characteristics and presents evidence that long-hour wage premiums and occupational social value are substitutes in compensating salaried workers who supply hours exceeding the standard working week. I show that the social value of an occupation—in particular the degree to which jobs involve helping or providing service to others—is inversely related to long-hour pay. Allowing for heterogeneity in the degree to which workers value their job's helping orientation lets me explore how gender differences in employees' attitudes toward pro-social behaviour can explain some of the observed occupational sorting trends and gender differences in long-hour compensation. Women tend to be more strongly drawn to 'helping' occupations and at the same time receive lower long-hour premiums in these jobs relative to men. I offer a theoretical framework to rationalize the empirical trends.

Keywords: long hour compensation | workplace | pro-social behavior

Article:

Introduction

Close to 40% of full-time psychologists work more than 40 hours per week, yet the median hourly wage of those who do is over 10% lower than the hourly wage of workers in the same occupation who put in between 35 and 40 hours per week. A comparable fraction of insurance sales agents work more than 40 hours per week, but the median hourly wage in this occupation is about 5% higher for the longer-hour group.¹ This paper aims to further our understanding of the occupational selection and earnings of salaried employees by addressing the

questions of what drives observed returns to long hours and what other than wages can incentivize employees to work more than the standard working week.

This study focuses on uncompensated long hours in jobs with pro-social characteristics. I estimate and analyse the occupation-specific effective long-hour premiums of salaried workers based on the observed weekly earnings of employees who work standard hours and those who work over 40 hours per week. I introduce a conceptual framework to rationalize several novel empirical facts presented in the paper regarding the association between long-hour compensation and occupational preferences among salaried workers. Using data on hours and earnings from the 2000 Census and 2000–2016 American Community Surveys, and occupational characteristics from the Occupational Information Network (O*NET) database, I present evidence that long-hour wage premiums tend to be decreasing in an occupation's non-monetary social value, suggesting that the two are substitutes in incentivizing some workers to supply hours exceeding the standard working week. Social value does not need to enter the production function for the relationship to hold; the driving mechanism is hypothesized to be job selection based on worker preferences for occupational characteristics and disutility of hours.

Further, I show that job selection based on preferences for occupational social value likely accounts for a portion of the gender wage gap among long-hour employees. Allowing for heterogeneity in the degree to which workers value their job's helping orientation lets me explore how gender differences in employees' attitudes toward pro-social behaviour can explain some of the observed occupational sorting trends and gender differences in long-hour compensation. I use the 1979 cohort of the National Longitudinal Survey of Youth to explore how job satisfaction varies with pro-social orientation, showing evidence that women value pro-social work more than men do. The difference in preferences leads to women earning relatively lower long-hour premiums compared to men in occupations with higher social value. Even though they earn relatively lower premiums, females are more likely than males to work in pro-social occupations.

I measure occupational social value empirically as the tendency of jobs to involve providing help or service to others. Examples of occupations that score highly along this dimension include psychologists, therapists, nurses and social workers. Because the focus of the paper is on salaried employees, I restrict the analysis to occupations that require relatively high levels of education, experience and training, as defined by their O*NET job zone, because workers in such occupations are less likely to be paid by the hour.

This paper adds to the literature on long-hour compensation and labour supply for salaried employees. Unpaid overtime is driven by career concerns in the literature on rat-race equilibria in the labour market (Landers et al. 1996; Bell and Freeman 2001). Long working hours are associated with future wage increases, especially in certain professional occupations (Gicheva 2013). Goldin (2014) argues that in occupations with low levels of flexibility, where workers are relatively poor substitutes for one another, productivity is convex in hours. On the other hand, productivity has been found to be concave in hours due to fatigue in a wide range of occupations (Collewet and Sauermann 2017; Pencavel 2015). Implicit wage-hour contracts, non-divisible tasks, or uncertainty about completion times may also account for variations in salaried employees' propensity to supply long working hours.

The result in this paper that long-hour premiums differ for men and women working in the same occupation because of selection, and that this difference contributes to the gender gap in earnings, complements several recent studies relating long-hour compensation to the gender wage gap. Goldin (2014) argues that preferences for flexibility drive much of the gender wage gap in occupations in which workers are imperfect substitutes for one another. Erosa et al. (2017)

explore a similar idea—that across-occupation variations in the returns to long hours combined with women spending more time on home production drive women into occupations where the production function is less convex in hours, which in turn accounts for a large portion of the gender wage gap. Cortes and Pan (2019) also relate occupational differences in the returns to long hours to the gender pay gap through gender differences in home production. My paper contributes to this literature by showing that even in the absence of productivity differences at long hours, selection driven by preferences for other occupational characteristics can exacerbate the earnings differences between men and women within and across occupations. I do not link long hours explicitly to higher productivity; it is possible that pro-social motivation induces workers to supply inefficiently long hours, for example in the presence of incomplete information as in principal–agent models such as the one in Landers et al. (1996).

The focus of my study on within-occupation earnings differentials between standard-hour and long-hour employees differentiates it from papers that identify wage penalties for working in pro-social occupations at any level of hours worked, such as England et al. (2002) and Hirsch and Manzella (2015). My results are in agreement with a strand of the literature, reviewed in Bertrand (2011), that finds women to be more socially minded than men. Furthermore, women tend to act in a more socially oriented manner in laboratory experiments (Eckel and Grossman 2008) and are more likely to favour public good and public health spending (Aidt and Dallal 2008; Miller 2008). Grove et al. (2011) show that among Graduate Management Admission Test registrants, women are more likely to attribute very high importance to their job contributing to society (19% compared to 11%), and such preferences are associated with lower salary. Including the variable in a wage regression reduces the gender earnings gap. Fortin (2008) shows similar trends using the National Longitudinal Study of 1972 (NLS72) and the National Education Longitudinal Study of 1988 (NELS88), and a preference index based in part on the attributed importance in selecting a career of ‘opportunities to be helpful to others or useful to society’. In the survey of post-secondary graduates used by Chevalier (2007), men are slightly more likely to say that ‘status and respect’ are very important job characteristics, while women attribute much more importance to the job being ‘socially useful’.

The paper also adds to the literature on non-monetary incentives—the consumption value of one's job—for salaried employees to donate hours or effort, while being the first to relate this idea both empirically and theoretically to observed occupation and gender differences in long-hour compensation. The findings of Gregg et al. (2011) are most similar to mine; they use data from the British Household Panel Survey to show that workers in non-profit caring industries (health, education and social care) are more likely to report unpaid overtime. Carpenter and Myers (2010) show that their experimental measure of altruism is positively correlated with the labour supply of volunteer fire fighters. In the experimental study of Fehrler and Kosfeld (2014), a subset of the participants are motivated by pro-social behaviour to exert a level of effort that exceeds what is optimal for a purely self-interested agent. In their theoretical work, Besley and Ghatak (2005) present a model in which worker intrinsic motivation and incentive pay are substitutes in eliciting effort. Proposed applications are non-profit organizations, education providers and public administration. Carpenter and Gong (2016) present experimental evidence in support of the Besley and Ghatak (2005) model, showing that workers are more productive when performing tasks for an organization whose mission is aligned with the worker's own views. In the model of Brekke and Nyborg (2010), workers differ in their preference for being important to others, and this desire can elicit effort in occupations—such as nursing—in which it

is hard or impossible to reward effort monetarily. DeVaro and Brookshire (2007) argue that intrinsic motivation in the public sector is a substitute for promotions in eliciting effort.

Unlike several of the studies mentioned above, mine also incorporates worker selection into occupations, which is in line with the early work of Filer (1986), where he finds a link between occupational choice and preferences for job characteristics such as prestige or contribution of one's job to society. Heyes (2005) and Taylor (2007) develop models that are relevant for markets in which effort is not contractible and applied to the labour market for nurses, in which lower pay can attract higher-quality workers, who are more dedicated to the job and regard it as a 'vocation'. Their models suggest that such workers are underpaid from a social efficiency perspective in monopsony markets. Aside from worker preferences, Hotchkiss and Rupasingha (2018) show that part of the selection of workers into pro-social occupations is driven by comparative advantage.

In the next section I introduce the theoretical framework that describes how preferences for pro-social job characteristics drive worker selection into occupations and into working long hours within an occupation. The three datasets used in the empirical analysis are described in Section II, while Section III outlines the approach for estimating an occupation-specific long-hour premium. Predictions from the theory regarding the relationships among an occupation's social value, the size of the long-hour premium, and observed gender differences in long-hour compensation are tested in Section IV. Finally, Section V discusses some alternative explanations for the observed trends, and Section VI concludes.

1. Theoretical Framework

This section offers a conceptual model that shows how worker sorting based on heterogeneity in the preferences for a job's pro-social characteristics generates a negative relationship between an occupation's social value and its average long-hour premium. Consider a labour market with two occupations that differ in the cost of workers' effort. Occupation 1 is the high social value occupation in which the disutility of long hours is lower than it is in occupation 2.

Jobs within an occupation differ in the long-hour premium $1 + p$ depending on the production technology. Differences in the marginal returns to long hours can arise for various reasons. Productivity or the rate of learning-by-doing may decrease at high levels of weekly hours if fatigue is an important factor (e.g. Pencavel 2015), or display convexities if workers first perform routine tasks with low potential for learning and move on to more challenging tasks only if enough hours are left to devote to such activities. The latter is more likely to describe certain professional white-collar occupations. In Landers et al. (1996), long hours are a signal of high unobserved productivity; in the theory in Gicheva (2013), earnings are convex in hours because workers who have lower disutility of labour are more likely to select into jobs with higher returns to skill accumulation. Goldin (2014) links high returns to long hours to low flexibility. She discusses how the returns to long hours can be particularly high in professional occupations such as law and business, where employees are fairly poor substitutes for one another.

Workers select into occupations based on their preferences and then choose whether to supply standard hours \bar{h} or long hours $h > \bar{h}$. The utility of worker i from working in occupation j is given by

$$U_{ij} = \begin{cases} w_j \bar{h} - c_j + \varepsilon_{ij} & \text{if hours} = \bar{h}, \\ w_j \bar{h} - c_j + \varepsilon_{ij} + w_j(1 + p)(h - \bar{h}) - a_{ij}(h - \bar{h}) & \text{if hours} = h. \end{cases}$$

Here ε_{ij} has distribution $F_j(\varepsilon)$ and is a random shock that determines the taste of worker i for working in occupation j but does not vary with hours. It is assumed that ε_{i1} and ε_{i2} are independent, and that workers see their realizations before choosing an occupation. Without loss of generality, I assume that $F_2(\varepsilon)$ is a degenerate distribution such that $\varepsilon_{i2}=0$. Since the choice of hours in this model is dichotomous, functional form assumptions are going to be less important than when workers optimize over hours.

Assuming that $0 < a_{i1} < a_{i2}$ and $0 < c_1 < c_2$ allows for the disutility of effort to be lower in pro-social occupations. The theory of compensating wage differentials would imply that $w_2 > w_1$ when the expected value of ε_1 is 0 or positive. It is assumed that p is drawn from a known distribution, which is the same for occupations 1 and 2. The latter assumption is important because it allows the model to predict how observed premiums will differ in the two occupations when productivity at high hours is independent of an occupation's non-pecuniary value. For simplicity, I assume that $1 + p$ is uniformly distributed according to $1 + p \sim U[0,2]$. I also assume that $a_{ij} < 2w_j$: in both occupations, there are at least some values of p for which a worker would choose long hours. Further, I assume that $w_2/w_1 < a_{i2}/a_{i1}$: the disutility of long hours increases faster than the wage from occupation 1 to occupation 2.

There are two types of workers, A and B, where type B's disutility of long hours is less dependent on the job's pro-social value. Specifically, I assume that $0 < a_{A1} < a_{B1} < a_{B2} < a_{A2}$. This assumption is somewhat stronger than needed but makes the exposition more tractable.

The timing in the model is as follows:

1. A worker observes her realization of ε_{i1} and compares her expected utility in occupations 1 and 2.
2. After selecting an occupation based on expected utility, the worker is offered a job characterized by the long-hour premium p . The worker cannot change occupations after observing p .
3. The worker decides whether to supply standard or long hours. Worker i chooses to work long hours in occupation j when $1 + p \geq a_{ij}/w_j$.

Proposition 1. Conditional on worker type i , the average observed long-hour premium is lower in pro-social occupations.

Proof. The average premium in occupation j among workers of type i who choose to supply long hours is

$$1 + \bar{p}_{ij} = 1 + \frac{a_{ij}}{2w_j}.$$

The assumption that $w_2/w_1 < a_{i2}/a_{i1}$ guarantees that $\bar{p}_{i1} < \bar{p}_{i2}$. This prediction of the model stems from the inherent trade-off between earnings and disutility of hours, where the latter is

lower in occupations with desirable characteristics. There is a wide range of other theoretical setups that can yield a similar result.

Proposition 2. Workers of type A, who value pro-social characteristics more than type B workers do, are more likely to work in the pro-social occupation but earn relatively lower average premiums in this occupation.

Proof. The difference in the average observed premium in occupation j between workers of types A and B is

$$\bar{p}_{Aj} - \bar{p}_{Bj} = \frac{a_{Aj} - a_{Bj}}{2w_j}.$$

It holds that $\bar{p}_{A1} - \bar{p}_{B1} < \bar{p}_{A2} - \bar{p}_{B2}$ under the assumption that $a_{A1} < a_{B1} < a_{B2} < a_{A2}$. Expected utility for worker i in occupation j is

$$\begin{aligned} E(U_{ij}) &= w_j \bar{h} - c_j + \varepsilon_{ij} + \left(1 - \frac{a_{ij}}{2w_j}\right) \left[w_j \left(1 + \frac{a_{ij}}{2w_j}\right) (h - \bar{h}) - a_{ij}(h - \bar{h}) \right] \\ &= w_j \bar{h} - c_j + \varepsilon_{ij} + w_j (h - \bar{h}) \left(1 - \frac{a_{ij}}{2w_j}\right)^2. \end{aligned}$$

The probability that worker i chooses the pro-social occupation is given by

$$\begin{aligned} &P(E(U_{i1}) > E(U_{i2})) \\ &= 1 - F \left((w_2 - w_1) \bar{h} - (c_2 - c_1) + (h - \bar{h}) \left[w_2 \left(1 - \frac{a_{i2}}{2w_2}\right)^2 - w_1 \left(1 - \frac{a_{i1}}{2w_1}\right)^2 \right] \right), \end{aligned}$$

which is higher for workers of type A given the assumptions on the parameters.

This model can be extended in several ways to accommodate other job characteristics. For example, certain occupational characteristics may be related to the underlying production technology, making workers more productive at long hours. This concept is discussed in Goldin (2014) in the context of non-flexible occupations having a production function that is convex in hours. In the framework of the model in this section, occupations can be allowed to differ on the distribution of the long-hour premium $1 + p$. Then the model would predict that the observed average premium is lower in flexible occupations, but there will not necessarily be a difference in the premium gap by worker type, unlike the predictions regarding pro-social value.

The model presented in this section is very simple and does not take into account many aspects of the worker's labour supply decision. The focus of the paper is on Propositions 1 and 1, and Section IV examines how well they align with the data. Prior to that, I describe the data used in the analysis in the next section, and discuss the methods for estimating the long-hour premium p in Section III.

2 Data

Wages and long-hour premiums

Long-hour premiums averaged at the occupation level are estimated using data from the 5% sample of the 2000 Census and the 2000–2016 American Community Surveys (ACS) provided through the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al. 2018). The main advantage of these data is the large sample size, which allows me to derive occupation-level measures using reliable cell sizes. Since I do not model changes in the premiums over time, I restrict the time period to exclude pre-2000 Census data to reduce the probability of differential time trends.

The estimation sample is comprised of all employed workers between the ages of 18 and 65 who report usual weekly hours higher than 34 and who worked at least 48 weeks during the year preceding the interview. Individuals in the armed forces are excluded from the data. Weekly and hourly earnings are inferred from annual labour income and reported weeks and hours worked.² All dollar amounts are converted to 1999 values using the Consumer Price Index for All Urban Consumers. Workers are dropped from the sample if their inflation-adjusted hourly earnings are less than \$4 or more than \$160. This sample restriction applies to 2.7% of the sample and is imposed to minimize the influence of miscoded observations or outliers.

Other information used in the empirical work includes the respondent's gender, age, marital status, age of youngest child in the household, race and ethnicity. I use the information on educational attainment to construct a set of indicators for the highest level of completed schooling: grade 8, grade 12 but lacking a high school diploma, high school diploma (including GED), some college but no Bachelor's degree, completed Bachelor's degree, and any graduate degree.

All occupation codes are converted, if necessary, to the 2010 Standard Occupational Classification (SOC) system for compatibility with the O*NET data, described below; the conversion procedure is described in more detail in the Appendix. The final Census/ACS sample used to estimate occupational long-hour premiums consists of 7,872,111 male and 5,868,789 female workers in, respectively, 435 and 329 occupations.

Occupational characteristics

Occupational characteristics are inferred from the Department of Labor/Employment and Training Administration's O*NET database (version 19.0). I use the information provided on worker characteristics and worker, experience and occupational requirements. The Helping score of an occupation combines the 'Assisting and caring for others' work activity score, the 'Social' occupational interest score, the occupation's 'Service orientation' skill requirement, the 'Concern for others' work style, and the 'Relationships' work value.³ I re-scale each measure before calculating the average score because the 'Social' and 'Relationships' components are measured on a 7-point scale, while all others are measured on a 5-point scale. A natural way to combine several related measures into a single index is by using principal component analysis. In this case, principal component analysis indicates that each component of the social value index should enter with approximately the same weight, so I simply average the scores for each factor. This and all other constructed occupational characteristics are standardized to have mean 0 and standard deviation 1.

In addition to the Helping score, which is the occupational characteristic of main interest, data from O*NET are used to measure several other characteristics that enter some empirical

specifications as control variables. This first is an occupation's Flexibility score that Goldin (2014) uses. The five occupational characteristics that she selects to measure lack of flexibility are the 'Time pressure', 'Structured versus unstructured work', 'Freedom to make decisions' and 'Contact with others' work contexts, along with the 'Establishing and maintaining interpersonal relationships' work activity. For consistency with Goldin (2014), I take an average of the flexibility measure components and use the negative of the resulting index in order to create a variable that measures flexibility, not the lack of it.

A preliminary examination of the estimated occupational premiums reveals that they appear to be consistently lower in many fields that are considered to be analytical, so I construct an occupation's Analytical score to use as an additional control variable in some specifications. This score is comprised of the following: 'Deductive reasoning', 'Inductive reasoning' and 'Mathematical ability' among the available ability measures; 'Mathematics' knowledge score; 'Mathematics' and 'Complex problem solving' skill scores; 'Analytical thinking' work style score; and the score on the 'Investigative' occupational interest. Where applicable, the importance rather than the level scale is used, but the difference between the two is not substantial. All but the 'Investigative' score are measured on a 5-point scale, while the latter is measured on a 7-point scale, which I account for when combining the scores. Similarly to the Helping measure, principal component analysis suggests that each factor enters with approximately equal weight, so the index that I use is an average of all components. Finally, the O*NET database allows me to measure an occupation's Competitiveness using the 'Level of competition' work context score.

Occupations in the O*NET are categorized into job zones based on how much education, experience and training they require. The scores range from 1 (little or no preparation needed) to 5 (extensive preparation needed). Occupations in job zones 1 and 2 are excluded from the analysis because workers in these occupations are more likely to be paid an hourly rate and to be subject to the legally mandated overtime premium.⁴ The job zone score is missing for 12 occupations in the data, which are also dropped. For male workers, there are 256 occupations in the final data with job scores of 3, 4 or 5; the corresponding number of occupations for female workers is 217. The total number of distinct occupations in the data is 270.

As a robustness check in some specifications, I include a proxy measure of the difficulty of measuring performance or output in an occupation. Specifically, I use data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79); the 1988–2000 instalments of the survey include questions asking whether respondents receive compensation based on job performance, including piece rates, commissions, bonuses, stock options and tips. Ewing (1996) uses a similar measure derived from the NLSY79 to examine the link between pay and the degree to which it is possible to observe performance accurately. I construct a standardized occupation-level performance pay variable based on a linear probability regression model that includes year indicators and controls for gender and quadratics in age, experience and tenure, as well as occupation-specific intercepts. The performance pay measure is available for 182 of the 270 occupations in the data.

The coefficients of correlation for the occupational characteristics that I consider are shown in Table 1. The helping and flexibility scores are highly negatively correlated (correlation coefficient -0.497). The helping score is also negatively correlated with the analytical and competitiveness measures. Jobs that are more competitive also tend to be less flexible and more analytical. The NLSY79-based performance pay measure is not correlated with the helping score but, not surprisingly, is fairly highly correlated with an occupation's level of competitiveness.

Table 1. Occupational Characteristics Correlation Matrix

	Helping	Flexibility	Analytical	Competitiveness	Performance pay
Helping	1				
Flexibility	-0.497	1			
Analytical	-0.232	0.053	1		
Competitiveness	0.162	-0.289	0.262	1	
Performance pay	-0.001	-0.311	-0.105	0.388	1

Notes

Data for 270 occupations with at least 1000 male or 1000 female workers. The performance pay measure is available for 182 occupations.

Table 2. Occupations with the highest helping scores

1	Recreational Therapists
2	Physical Therapists
3	Educational, Guidance, School and Vocational Counsellors
4	Licensed Practical and Vocational Nurses
5	Clergy
6	Child, Family and School Social Workers
7	Clinical, Counselling and School Psychologists
8	Registered Nurses
9	Occupational Therapists
10	Special Education Teachers

Table 2 shows the ten occupations in the data with the highest helping scores. Five of these occupations are in the healthcare sector (SOC codes between 29-1000 and 29-2090). The other five include guidance and vocation counsellors, clergy, psychologists, social workers and special education teachers. My ranking of occupations based on their pro-social orientation is similar to those in Hirsch and Manzella (2015), also derived from O*NET data, and Hotchkiss and Rupasingha (2018), based on a career counselling website.

Job satisfaction

To provide evidence that women value job satisfaction and flexibility more than men do, I use the NLSY79. The NLSY79 sample size is much smaller than the Census/ACS data, so most occupation cells are not populated enough to aggregate data at the occupation level, but the NLSY79 has several advantages that can be exploited once occupation-level variables are derived from other data. Each instalment of the survey asks about employed respondents' level of job satisfaction. In addition, the panel nature of the data allows me to difference out any time-invariant individual heterogeneity.

The occupation-level O*NET characteristics described in the previous subsection are merged with the NLSY79 data after the NLSY79 occupation codes are converted to the 2010 SOC system following the procedure described in the Appendix.⁵ Observations for which

respondents are not employed or hold occupations in job zones 1 or 2 are dropped, as are those observations for which the relevant gender–occupation cell has fewer than 1000 observations in the Census/ACS data. I retain in the sample only observations for which the respondent is older than 19. Observations for which individuals are employed in a family business or working without pay constitute less than 1% of the sample and are dropped. I also drop employed respondents working fewer than 35 hours per week (14% of the remaining sample) and 48 weeks per year (18% of the sample). To avoid the influence of outliers and potential coding errors, as well as for consistency with the Census/ACS data, I drop 3% of the sample whose inflation-adjusted hourly wage is less than \$4 or more than \$160.

Each instalment of the survey asks respondents to rate their job satisfaction on a 4-point scale.⁶ In the empirical analysis I use either a 4-point Likert scale that assigns values between 1 for the lowest satisfaction category and 4 corresponding to the highest satisfaction response, or a binary measure equal to 1 if the highest satisfaction category is selected, and 0 otherwise. The final sample consists of 25,553 observations for 3688 men and 24,584 observations for 3805 women for the period 1982–2014. Summary statistics for the NLSY79 sample are shown in Table A1 of the Appendix.

3 Calculation of Long-hour Premiums and Descriptive Evidence

Long-hour premiums

There are different ways to conceptualize and estimate empirically an occupation-level measure of compensation for long hours. For the main specifications, I rely on an approach that is relatively simple but verify the robustness of the results to using alternative methods. I assume that in the case of standard-hour employment, the observed weekly earnings of worker i in occupation j are given by $y_{ij}=h_s w_{ij}$, where w_{ij} is an hourly wage, and h_s denotes standard hours. It is likely that h_s varies somewhat across occupations, but rather than estimating the occupation-specific standard level of hours as a parameter in the empirical model, I set the level of h_s at 40 for all jobs. The results are robust to using median hours among full-time employees in an occupation as the standard-hour cut-off. I also estimate the models using an alternative measure of the long-hour premium, discussed later in this section, that makes no assumptions about the length of the standard working week.

If the occupation-specific long-hour premium is denoted by p_j and an individual works $(h_{ij}-h_s)$ hours beyond the standard level, then observed weekly earnings are

$$y_{ij} = h_s w_{ij} + (h_{ij} - h_s)(1 + p_j)w_{ij}.$$

To estimate p_j empirically from observed weekly earnings y_{ij} and hours h_{ij} , I first use the sample of workers in the 2000 Census and 2000–2016 ACS whose hours are between 35 and 40 to estimate a wage equation separately by gender:

$$\ln(w_{ij}) = w_{0j} + X_i' \beta + v_t + e_{ij},$$

where t stands for year (2000–2016). The variables in X are: age and age squared; an indicator for married and living with spouse; indicators for the age of the youngest child in the household falling in the intervals 0–4, 5–9 and 10–17 years; indicators for race and Hispanic ethnicity; the binary variables measuring educational attainment; and indicators for self-employment, whether the worker is in the private for-profit sector, and whether she is currently enrolled in school.

Next, for workers who report usual hours higher than 40, I construct the predicted wage

$$\widehat{w}_{ij} = \exp(\widehat{w}_{0j} + X_i' \widehat{\beta} + \widehat{v}_i) \exp(\widehat{\sigma}^2/2)$$

using the sample standard deviation of the regression residuals $\widehat{\sigma}$.⁷ Then the estimated premium for worker i is

$$1 + \widehat{p}_{ij} = \frac{y_{ij} - 40\widehat{w}_{ij}}{(h_{ij} - 40)\widehat{w}_{ij}}.$$

(2)

I average the estimated premium for each occupation separately for men and women:

$$\bar{p}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \widehat{p}_{ij},$$

given n_j workers in occupation j who work more than 40 hours per week. It is assumed that occupation premiums did not change during the sample period, so the individual estimates of p_{ij} are averaged across all years.⁸ Table A2 in the Appendix shows, separately by gender, the mean and standard deviation, along with the 25th, 50th, 75th and 90th percentiles of the distribution of $1 + p$ for the occupations used in the analysis.

Measurement error in hours is a common concern in the literature (e.g. Bound et al. 2001). In the setting here, systematic over-reporting of hours by workers in the 35–40-hour group will lead to underestimation of w_{ij} and thus overestimation of p_{ij} . On the other hand, if it is mostly long-hour workers who tend to over-report their hours, then \widehat{p}_{ij} will be biased down in equation 2. Then one would incorrectly uncover a negative correlation between the estimated premiums and occupational social value if workers in helping occupations tend to under-report their hours when hours are around 40 and over-report their hours at higher levels of hours. There is no particular reason to believe that this behaviour is common in pro-social occupations but not in other occupations.

One other potential concern is that the wage in equation 1 is observed only if worker i works between 35 and 40 hours per week, and this decision may be correlated with unobservable characteristics related to one's earning potential. As another robustness check, I estimate equation 1 as part of a selection model; the variables included in the selection equation but excluded from the wage equation are: the indicator for married and living with spouse; indicators for the age of the youngest child in the household falling in the intervals 0–4, 5–9 and 10–17 years; an

indicator for whether anyone else in the respondent's immediate family has positive income and its interaction with the natural log of this income; and an interaction for whether the respondent has non-labour income and the interaction with the natural log of non-labour income. All other variables from the model above, including the occupation indicators, enter both the selection equation and the wage equation. Although the levels of the estimated premiums change with this approach, Table A3 in the Appendix shows that the correlation between the selection-based premiums and the premiums that I use in the main specifications in the paper is higher than 0.9 for both men and women. Not surprisingly given the high correlations, the results change very little if the selection-based premium is used; these results are available in an Online Appendix.

Another approach to quantifying the occupation-specific return to long hours is to estimate for workers in each occupation the elasticity ε_{yh} of annual earnings y with respect to weekly hours h . This is the method used in Goldin (2014) to examine the degree to which certain occupations reward workers for long hours. An elasticity greater than 1 means that the relationship between hours and earnings is convex, similarly to the premium p being greater than 0.

For this alternative definition of the premium, I use reported annual earnings y_{ij} to estimate

$$\ln(y_{ij}) = \tilde{X}'_i \alpha + \varepsilon_{yh,j} \ln(h_{ij}) + v_t + u_{ij}.$$

The controls in \tilde{X}_i are chosen to correspond to those in Goldin (2014) and differ slightly from the variables included in equation 1. Specifically, \tilde{X} includes a quadratic in age, indicators for education, race and Hispanic ethnicity, and the natural log of weeks worked during the previous year. The estimated occupation-specific coefficients $\hat{\varepsilon}_{yh,j}$ are an alternative measure of the return to working long hours. As Table A3 of the Appendix shows, the coefficient of correlation between the estimated premiums as defined by $1 + p$ and ε_{yh} is very high and similar for men (0.826) and women (0.824). I show in the Online Appendix that the main empirical results are robust to the choice between $1 + p$ and ε_{yh} as a measure of the long-hour premium. For the rest of the paper, I use the first definition of long-hour premiums discussed in this section.

Descriptive evidence

Some examples of occupations with long-hour premiums below the median are shown in Table 3. The table shows the percentile into which a given occupation's premium falls relative to the rest of the gender-specific distribution. The table also lists the share of female workers and the helping score for each of the occupations. These occupations are chosen as examples of the types of jobs that tend to be associated with low premiums. It is evident from the table that the helping scores in these occupations vary, but many low-premium occupations can be characterized as predominantly female and with high pro-social value.

Table 3 Selected low-premium occupations

Occupation	Percentile (male)	Percentile (female)	Share Female	Helping score
Aerospace Engineers	4.4	10.0	0.10	-0.72
Mechanical Engineers	7.4	11.9	0.06	-0.95
Industrial Engineers	5.6	22.2	0.18	-0.58
Electrical Engineers	11.5	9.6	0.08	-0.82
Conservation Scientists	5.2	12.2	0.17	0.06
Physicists	5.9	5.2	0.15	-1.14
Biologists	14.1	12.6	0.43	-0.96
Clinical, Counselling and School Psychologists	48.2	14.1	0.62	1.99
Child, Family and School Social Workers	8.1	4.8	0.79	2.03
Educational, Guidance, School and Vocational Counsellors	27.4	13.0	0.68	2.16
Clergy	17.4	14.4	0.13	2.11
Elementary and Middle School Teachers	0.7	1.5	0.77	1.55
Secondary School Teachers	1.1	1.9	0.55	1.51
Special Education Teachers	1.5	4.1	0.85	1.85
Occupational Therapists	4.1	4.4	0.86	1.95
Physical Therapists	45.2	8.5	0.63	2.16
Radiation Therapists	23.3	11.5	0.69	1.61
Recreational Therapists	0.0	6.7	0.75	2.18
Respiratory Therapists	7.8	3.7	0.60	1.82
Speech-Language Pathologists	4.8	2.6	0.95	1.78
Registered Nurses	29.3	5.9	0.90	1.96
Licensed Practical and Vocational Nurses	35.2	16.3	0.91	2.14
Veterinarians	20.4	8.9	0.48	1.02

Notes The percentiles refer to gender-specific occupation rankings based on the estimated long-hour premiums.

Engineers comprise the first group of occupations in Table 3. The table shows four engineering occupations; their premiums are in the 12th percentile or lower for men and the 22nd percentile or lower for women. All engineering occupations (SOC codes in the 17-2000 group) with more than 1000 workers have premiums below the median for both men and women. Engineering occupations tend to be less pro-social than average (helping scores in the sample range between -0.58 and -1.29), and have low shares of female workers. With the exception of medical scientists and male environmental scientists, all occupations in the natural sciences (SOC codes in the 19-1000 and 19-2000 groups) also have premiums below the median. Three of these occupations are listed in Table 3. Conservation scientists, physicists and biologists earn premiums in the 14th percentile or lower among both men and women. To account for the concentration of engineering and science occupations among those with low premiums, some of the empirical specifications in the next section include an analytical score that I construct from the O*NET data.

Further, Table 3 shows that workers in social service occupations, such as psychologists, social workers, guidance counsellors and clergy, also earn low premiums. These occupations are overwhelmingly female (with the exception of clergy) and score high on the helping measure. Elementary, middle, secondary school and special education teachers earn particularly low premiums (all below the 5th percentile) and are also predominantly female and pro-social. Low premiums (12th percentile or lower) are also observed in all therapist occupations in the sample (SOC group 29-1120) with the exception of male physical and radiation therapists. These occupations also have high concentrations of female workers and have very high helping scores. Nurses and veterinarians also earn low premiums, women much more so than men. Overall, all of the most helping occupations listed in Table 2 pay relatively low premiums, as shown in Table 3.

As seen in Smith (2007), there is considerable overlap between Table 3 and the occupations with the highest job satisfaction rates in the 1972–2006 waves of the General Social Survey. Teachers, for example, earn very low premiums but report high job satisfaction. Four other occupations (clergy, physical therapists, psychologists and special education teachers) rank at the top in both helping score and job satisfaction. The above observations offer preliminary descriptive evidence of links between long-hour premiums and the social value of occupations, which is in line with the theoretical framework in Section I.

Table 4 shows descriptive statistics for several variables of interest aggregated at the occupation–gender level. Occupations are divided into two groups, based on whether their estimated long-hour premium p is above or below the 25th percentile of the gender-specific distribution. The table shows that without conditioning on other observables, average weekly hours are slightly lower in low-premium occupations. Average hours in low-premium occupations equal 44.5 for men and 42.2 for women, with 40% of men and 26% of women working more than 40 hours per week. Median wages for workers who put in between 35 and 40 hours are higher in low-premium occupations, which may reflect compensating wage differentials, a trade-off between regular and overtime wages similar to the one for employees covered by the Fair Labor Standards Act as demonstrated in Trejo (1991), measurement error, or a combination of factors.⁹ The median hourly wage in longer-hour jobs is similar in high- and low-premium occupations. It appears that higher-premium occupations may be more competitive, since wages in these jobs are more dispersed, and this is supported by the higher competitiveness scores of these occupations.

The last section of Table 4 compares the O*NET-derived characteristics and the performance pay score of high- and low-premium occupations in the data. The difference in the average helping score for men is about 0.1 standard deviations and is not statistically significant. For females, the average helping score is about half of a standard deviation higher in low-premium occupations. In addition, women tend to work in occupations with higher helping scores compared to men. Low-premium jobs also tend to be characterized by higher analytical and flexibility scores, especially for men, and a lower share of workers receiving performance pay.

4. Empirical Results

Occupational characteristics and long-hour premiums

I investigate further the unconditional relationships observed in Table 4 by estimating a regression of the gender-specific occupation premium $1 + p$ on the set of O*NET occupational characteristics discussed in the previous sections.¹⁰ The coefficient estimates are presented in Table 5, where all models include job zone dummies, whose coefficients are assumed not to differ between male and female workers, and an indicator for gender. The reported standard errors are heteroscedasticity-robust.

Table 4. Summary Statistics

Variable	High premium	Male low premium	<i>t</i> -statistic	High premium	Female low premium	<i>t</i> -statistic
Average hours	45.2 (2.58)	44.5 (2.1)	1.84	43.1 (2.05)	42.2 (1.49)	3.052**
Percentage hours 41-54	0.296 (0.076)	0.286 (0.074)	0.90	0.237 (0.082)	0.205 (0.072)	2.55*
Percentage hours 55+	0.137 (0.082)	0.114 (0.071)	2.07*	0.079 (0.06)	0.055 (0.04)	2.79**
Median wage (35-40 hours)	19.8 (6.92)	22.2 (7.02)	-2.40*	16.3 (5.74)	19 (5.31)	-3.05**
Median wage (40+ hours)	19.7 (6.82)	19.3 (5.82)	0.37	16.4 (5.35)	16.9 (4.71)	-0.60
S.D. of wage (35-40 hours)	14.5 (6.68)	11.4 (4.23)	3.42***	10.6 (4.43)	8.71 (2.53)	3.00**
S.D. of wage (40+ hours)	17 (7.66)	11.3 (4.29)	5.66***	12.7 (5.75)	8.61 (2.75)	5.04***
Helping	-0.038 (0.944)	-0.141 (1.06)	0.73	-0.008 (0.89)	0.548 (1.22)	-3.62***
Flexibility	-0.103 (1.02)	0.282 (0.878)	-2.71**	-0.153 (1.01)	-0.013 (0.865)	-0.92
Analytical	-0.07 (0.908)	0.346 (1.12)	-2.99**	-0.047 (0.942)	0.242 (1.07)	-1.90
Competitiveness	0.115 (1.01)	-0.21 (0.855)	2.32*	0.161 (1.06)	-0.361 (0.985)	3.22**
Performance pay	0.112 (0.995)	-0.351 (1.04)	2.75**	0.216 (0.972)	-0.3 (0.493)	3.03**
<i>N</i>	192	64		162	55	

Notes

Standard deviations are in parentheses. The helping, flexibility, analytical and competitiveness scores are derived from the O*NET database. The performance pay measure is calculated using NLSY79 data. All other statistics are based on the Census and ACS data. Only occupation–gender cells with 1000 or more workers are included in the calculations. High-premium occupations are those that rank among the highest 75% in terms of the average long-hour premium $1 + p$. Low-premium occupations are in the 25th percentile or lower. The ranking is gender-specific.

*, **, *** indicate $p < 0.10$, $p < 0.05$, $p < 0.01$, respectively.

Table 5. Long-hour Premiums and Occupational Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Helping	-0.053** (0.026)	-0.165*** (0.031)	-0.155*** (0.031)	-0.163*** (0.031)	-0.126*** (0.037)	
Flexibility		-0.199*** (0.031)	-0.157*** (0.035)	-0.193*** (0.032)	-0.197*** (0.031)	
Analytical		-0.141*** (0.032)	-0.190*** (0.037)	-0.156*** (0.034)	-0.122*** (0.034)	
Competitiveness		0.138*** (0.030)	0.069** (0.029)	0.133*** (0.031)	0.150*** (0.032)	
Performance pay			0.212*** (0.040)			
\bar{W}_{40}				0.005 (0.005)		
Helping x Job zone 3						-0.095** (0.039)
Helping x Job zone 4						-0.161** (0.065)
Helping x Job zone 5						-0.291*** (0.074)
Flexibility x Job zone 3						-0.155*** (0.046)
Flexibility x Job zone 4						-0.236*** (0.048)
Flexibility x Job zone 5						-0.278*** (0.083)
Analytical x Job zone 3						-0.117** (0.050)
Analytical x Job zone 4						-0.128** (0.053)
Analytical x Job zone 5						-0.297*** (0.087)
Competitiveness x Job zone 3						0.031 (0.041)
Competitiveness x Job zone 4						0.243*** (0.056)
Competitiveness x Job zone 5						0.241*** (0.071)
Female	0.096* (0.057)	0.093* (0.052)	0.085 (0.056)	0.114** (0.054)	0.128** (0.055)	0.093* (0.051)
\bar{R}^2	0.016	0.204	0.370	0.204	0.208	0.2342
<i>N</i>	473	473	322	473	427	473

Notes

Standard errors are in parentheses. The dependent variable in all models is $1 + p$; see Section III for details about the construction of this variable. The variable \bar{W}_{40} measures the median wage

for workers in the occupation reporting between 35 and 40 hours. The results in column (5) exclude occupations in the healthcare sector with SOC codes between 29-1000 and 29-2090. All specifications include job zone indicators. The reported errors are robust.

*, **, *** indicate $p < 0.10$, $p < 0.05$, $p < 0.01$, respectively.

Consistent with the descriptive statistics in Table 4 and with Proposition 1 in Section I, premiums tend to be lower in ‘helping’ jobs. Column (1) of Table 5 shows results when only the helping score of an occupation is included; there is a negative and statistically significant relationship between this variable and the occupation premium, suggesting that premiums tend to be lower in pro-social occupations. When the flexibility, analytical and competitiveness scores are added to the model (column (2) of Table 5), the coefficient on the helping score remains negative and significant but more than triples in size, and the explanatory power of the regression increases substantially. Occupations with higher levels of flexibility also tend to pay lower overtime premiums, which is consistent with the argument in Goldin (2014). Long-hour premiums tend to be lower in analytical occupations as well, consistent with the observation in Table 3 that engineering and science occupations pay low long-hour premiums. More competitive occupations pay higher premiums on average, which is in line with existing theories of career concerns.

The model in column (3) of Table 5 includes the characteristics above along with the performance pay measure derived from the NLSY79; since this variable is not available for all occupations in the data, the sample for this regression is smaller. The probability of receiving performance pay is positively correlated with the long-hour premium; since performance pay is a proxy for the degree to which output in a given occupation is measurable, this would suggest that workers are more likely to be rewarded for long hours when their performance is observable. Because the probability of receiving performance pay is fairly highly correlated with how competitive an occupation is, as shown in Table 1, the inclusion of the performance pay score leads to a decrease in the estimated coefficient on the competitiveness measure, but it remains positive and significant. All other coefficients decrease in magnitude as well, but only slightly.

As a robustness check, the model in column (4) of Table 5 includes the covariates from column (2) but adds the gender-specific median of the wage in an occupation for workers who report between 35 and 40 hours per week. Adding this variable has very little impact on the other coefficient estimates, and there appears to be no relationship between standard-hour wages and occupation premiums. The model in column (5) differs from that in column (2) in that it excludes occupations in the healthcare sector with SOC codes between 29-1000 and 29-2090. This restriction is made as a robustness check because Table 3 suggests that many health-related occupations pay low premiums—the bottom nine rows of Table 3 fall in this group—while five of the ten occupations in Table 2 with very high helping scores are also in the healthcare sector. Since the negative relationship between helping scores and long-hour premiums appears to be very strong for this group, it is important to check whether all or most of the results are driven by the healthcare sector. Restricting the sample to exclude occupations in healthcare reduces the magnitude of the coefficient estimate on the helping score, going from -0.165 to -0.126 , but the estimate remains negative and highly statistically significant.

Finally, the specification in column (6) of Table 5 shows the relationship between each occupational characteristic and the long-hour premium separately by job zone. These coefficient estimates offer suggestive evidence that the results for each characteristic are stronger for occupations that require higher levels of preparation. Plausibly, this is the case because the share

of salaried workers increases with job zone, and the conceptual framework in this study does not apply to hourly employees who are subject to legally mandated overtime compensation.

The magnitude of the estimated relationship between premiums and the helping score could account for a considerable part of the gender wage gap even in the absence of differences in the number of hours worked between male and female workers. Women in both the NLSY79 and Census/ACS data are observed in occupations whose helping score is on average slightly over 0.6 standard deviations higher than the average helping score of the occupations in which men work; this relationship is later explored in more detail in Figure 2 and Table 6. Then the results in Table 5 would suggest that, everything else constant, the gender difference in average premiums due to women working in more pro-social occupations is around 0.1. Comparing an average male and female worker who otherwise earn similar wages and both work 50 hours per week, but the female is employed in an occupation whose helping score is 0.6 standard deviations higher, the female–male difference in the natural logarithms of weekly earnings will be

$$\ln(40w + 10(1 + p - 0.1)w) - \ln(40w + 10(1 + p)w) \approx -\frac{1}{40 + 10(1 + p)},$$

which equals -0.02 when $1 + p$ is 1. Of course, the gender wage gap will be higher if the female worker in the above example works fewer hours and earns a lower baseline wage, for example because of compensating wage differentials associated with helping jobs or differences in accumulated labour market experience. As a comparison, Goldin (2014) finds a log earnings difference of -0.247 for full-time workers when she controls for worker demographics and education, and a difference of -0.192 after adding occupation controls. Thus occupation absorbs 0.055 log points of the gender earnings differential, of which the calculation above suggests 0.02 log points is due to women's differential selection into helping occupations when looking at workers who report around 50 hours per week.

Table 6 Within-Occupation Difference in the Long-Hour Premiums for Male and Female Workers

Dep. Variable	$p^F - p^M$				Share female		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Helping	-0.115*** (0.029)	-0.096** (0.042)	-0.097**	0.158*** (0.013)	0.121*** (0.018)	0.101*** (0.017)	0.101*** (0.017)
Flexibility		0.024 (0.038)	0.017 (0.041)		-0.027 (0.018)	-0.021 (0.016)	-0.021 (0.016)
Analytical		0.004 (0.032)	0.008 (0.033)		-0.047*** (0.017)	-0.046*** (0.015)	-0.046*** (0.015)
Competitiveness		0.036 (0.030)	0.043 (0.030)		-0.053*** (0.017)	-0.046*** (0.015)	-0.046*** (0.015)
$\overline{W}_{40}^F - \overline{W}_{40}^M$			0.011 (0.013)				-0.0003 (0.005)
$p^F - p^M$						-0.214*** (0.032)	-0.214*** (0.032)
\overline{R}^2	0.061	0.055	0.054	0.309	0.354	0.485	0.483

<i>N</i>	269	269	269	269	269	269	269
----------	-----	-----	-----	-----	-----	-----	-----

Notes

Standard errors are in parentheses. The sample consists of occupations with at least 1000 male or 1000 female workers, but excludes one occupation with fewer than 50 females. All specifications include job zone indicators. The reported errors are robust.

*, **, *** indicate $p < 0.10$, $p < 0.05$, $p < 0.01$, respectively.

The coefficient estimates in Table 5 also imply that females tend to earn higher long-hour premiums than males in a given occupation. Unlike other results reported in this section, this is highly contingent on the choice of empirical methodology used to estimate long-hour premiums. The estimated coefficients on the female indicator are not statistically different from 0 in five of the six specifications when using the Goldin (2014) definition of the premium; the coefficients are negative and highly significant when premiums are estimated with an adjustment for selection into standard or long hours (results available in the Online Appendix). Because the standard-hour wage is estimated separately by gender, different methodologies result in different average wages, thus average premiums differ as well. However, this moves the whole gender-specific distribution of premiums without changing the relationship between premiums and occupational characteristics.

Occupational selection and gender differences in long-hour premiums

The statistics presented in Table 3 show some evidence that occupations with large concentrations of female workers and high helping scores—such as registered nurses, psychologists, or therapists—often pay higher long-hour premiums to men. At the same time, occupations with lower helping scores and higher concentrations of male workers—such as many engineering jobs—tend to rank higher in terms of their estimated premium for the underrepresented gender, women in this case. This is consistent with women's preferences for pro-social value and disutility of long hours aligning with those of workers of type A in the theory in Section I, and men being represented by type B workers, who have lower disutility of long hours in non-helping jobs but also derive less utility from a job's pro-social characteristics.

Figure 1 plots the difference within an occupation in the estimated premiums for female and male workers ($p_F - p_M$) against the share of females in the occupation, based on the Census and ACS data. The sample consists of the 269 occupations in the data that have at least 50 observations for each gender. The relationship is shown separately for job zones 3 (panel (a)), 4 (panel (b)) and 5 (panel (c)). A negative relationship can be inferred from all three graphs: the larger the share of female workers in an occupation, the lower the relative premiums tend to be for female workers. Furthermore, panels (a), (b) and (c) of Figure 2 show a positive relationship between an occupation's helping score and the share of female workers in that occupation for all job zones. The trends in Figures 1 and 2 are consistent with Proposition 1 in Section I; these relationships are quantified in Table 6.

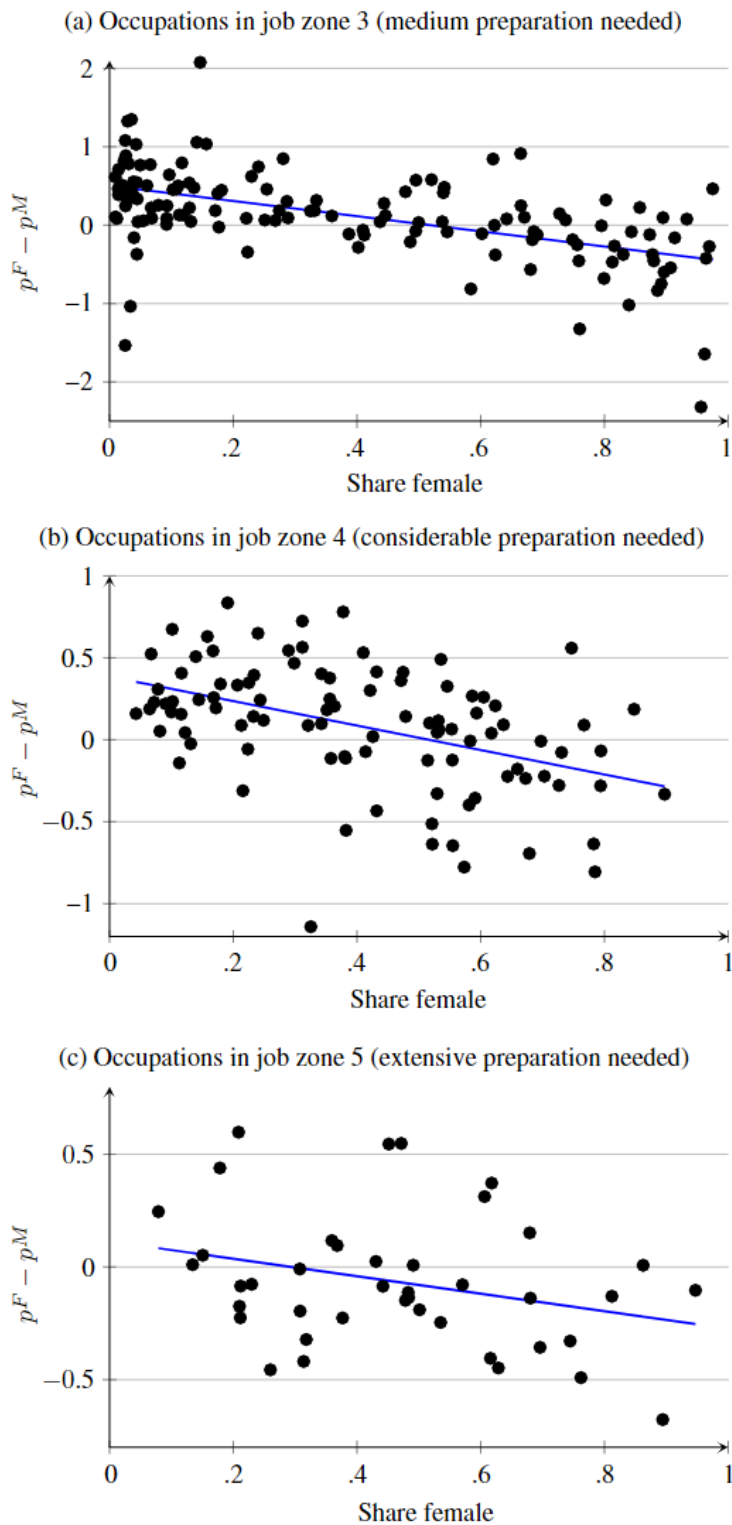


FIGURE 1. Share of females in an occupation and gender differences in the long-hour premium. Notes: The figures are based on the 269 occupations in the main data that have at least 50 male and 50 female workers. [Colour figure can be viewed at wileyonlinelibrary.com]

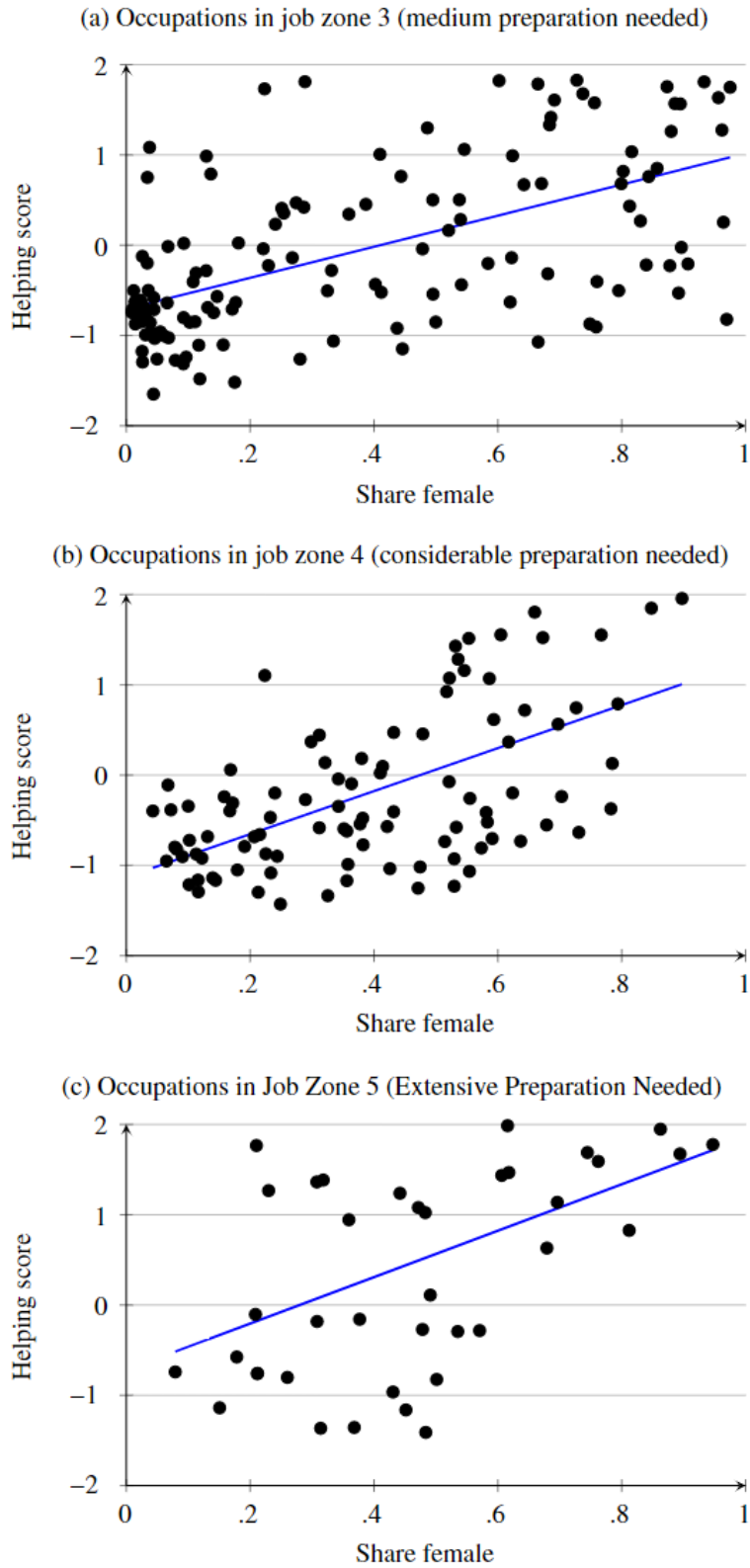


Figure 2 Share of females in an occupation and helping score.

Notes: The figures are based on the 269 occupations in the main data that have at least 50 male and 50 female workers. [Colour figure can be viewed at wileyonlinelibrary.com]

I estimate a regression model in which the difference between an occupation's female and male premiums is modelled as a function of occupational characteristics; the results are reported in columns (1)–(3) of Table 6. The regression in column (1) includes only job zone indicators and the helping score; other occupational characteristics are added to the specification in column (2), and the gender difference in median wages for workers in the 35–40-hour range is an additional regressor in column (3). An occupation's helping score has a negative and statistically significant relationship with the gender difference in premiums. This result suggests that women earn lower long-hour premiums relative to men in occupations with higher social value, which is in line with the theoretical framework presented in Section I. At the same time, the estimated coefficients on the flexibility and competitiveness scores are positive but not statistically significant, suggesting that women may earn slightly higher premiums relative to men in more flexible and more competitive jobs. The premium difference is not related to the occupation's analytical score or the difference in standard-hour wages.

Finally, the models in columns (4)–(7) of Table 6 show how an occupation's gender distribution is correlated with job characteristics by estimating regressions of the fraction of female workers in an occupation. In addition to job zone indicators, the estimated models include only the helping score (column (4)), all characteristics considered above (column (5)), all characteristics and the gender difference in premiums ($p^F - p^M$, column (6)), and all variables plus the gender difference in median standard-hour wages (column (7)). Consistent with the plots in Figure 2, the coefficient on the helping measure is positive and highly significant. One standard deviation increase in an occupation's helping score is associated on average with a 10–12 percentage point increase in the share of females in that occupation.

The estimated coefficients on the analytical and competitiveness measures suggest that a standard deviation increase in these characteristics is associated with a decrease of about 5 percentage points in the share of females in an occupation, while the coefficient on the flexibility measure is not statistically significant. While a few female-dominated occupations score high on the flexibility measure (e.g. file clerks or library technicians), many high-flexibility occupations, such as those in engineering or production fields, have an overwhelmingly large share of male workers. Occupations that pay relatively lower long-hour premiums to females also tend to be more predominantly female, as implied by the negative and highly significant coefficients on the $p^F - p^M$ variable. Furthermore, the regression's explanatory power increases noticeably when this variable is included. This is again consistent with the theoretical framework in Section I. The gender difference in the median wage for standard-hour workers, on the other hand, is not correlated with the share of females in an occupation.

The results in the two subsections of this section are consistent with Propositions 1 and 1 in Section I. Further, Tables 5 and 6 show examples of other characteristics correlated with the long-hour premium, but unlike the helping score, these characteristics are not related to the within-occupation gender difference in long-hour premiums. One potential explanation for this result is that other occupational characteristics may be correlated with productivity at high hours but preferences for these characteristics may not differ by gender. In the case of flexibility, a consistent gender-related pattern may fail to emerge because even if women derive higher utility from this characteristic, as Goldin (2014) argues, many male-dominated occupations such as those in engineering are also characterized by high levels of flexibility.

As further evidence consistent the conceptual framework in this paper, Table A5 of the Appendix shows the relationship between occupational social value and selection into long hours. The dependent variable for the linear regression models in columns (1)–(3) is an indicator

for the share of workers in an occupation who work between 41 and 54 hours per week; the results in columns (4)–(6) show the correlates of the share of workers in an occupation who report very long hours: 55 or more per week. Conditional on the long-hour premium, workers in helping occupations are more likely to supply 55 or more hours per week. Furthermore, when health-related occupations are excluded from the sample, the results suggest that women in helping occupations are more likely than men in these occupations to work over 40 but less than 55 hours per week. This is consistent with the idea that women have lower disutility of long hours in pro-social jobs, although the gender difference is not statistically significant for the share of workers reporting 55 or more hours per week. Women being relatively more likely to work long hours in jobs where long hours are rewarded less is another contributor to gender wage inequality.

As pointed out in Section I, the theoretical framework presented in this paper is chosen for being straightforward but this comes at the cost of ignoring multiple potentially important differences across occupations. The next section discusses some alternative explanations for the observed trends.

5 Alternative Explanations

In the absence of data on productivity, we cannot be certain that the observed long-hour premium is rewarding effort and output. This section discusses some alternative explanations for the observed results, several of which stem from the possibility of labour market frictions caused by incomplete information.

A common characteristic of many of the low-premium occupations shown in Table 3 is that it is relatively difficult to measure output in these jobs. At the same time, premiums tend to be high in sales occupations, for example, where it is often easier to assess individual performance. It is reasonable to expect employers to reward long hours more when these additional hours lead to higher levels of measurable output. If the difficulty of measuring performance is correlated with occupational social value, then the observed negative relationship between premiums and the helping score may be due to omitted variable bias. One proxy proposed in the literature for the difficulty of measuring performance is the incidence of performance pay (Ewing 1996; Lemieux et al. 2009). As discussed in Section II, the NLSY79 data allow me to construct such a measure for most occupations in the data. The correlation matrix in Table A3 of the Appendix reveals essentially no relationship between the incidence of performance pay and an occupation's pro-social value, and the results in the previous section show that the performance pay measure is indeed positively correlated with the long-hour premium but adding it to the regression equations does not have an impact on the 'Helping' coefficient estimate.

A related mechanism can lead to negative relationship between pro-social motivation and observed long-hour premiums if there are two main types of tasks in which long-hour employees engage in helping occupations: more easily measurable bureaucratic tasks (e.g. reports, checks with supervisors) and more client-oriented tasks whose output is more difficult to observe (e.g. visits, teaching, counselling).¹¹ The salience of the social impact generated is greater in the latter group of activities, but it is reasonable to expect compensation to be more closely tied to the former, where monitoring is less costly. Thus workers with higher pro-social motivation would be more likely than their career-oriented colleagues to engage in work activities for which they are not compensated, and the more opportunities there are for such activities, the lower the long-

hour premium would be. This mechanism could be explored further in a dataset containing information on the types of tasks that employees do, especially during extra hours of work.

Career concerns on their own are also likely to be a driving factor for long-hour premiums, as implied by Landers et al. (1996) and related studies. Working long hours may be a way for workers in helping occupations to signal to the employer their pro-social motivation, and if women are more likely to use this signalling mechanism, then this could lead to lower observed premiums for females in such occupations. As far as the measure of competitiveness of an occupation controls for career concerns, the results in the previous section suggest that the relationship between premiums and pro-social characteristics is robust to the inclusion of this variable. We also observe a negative correlation between competitiveness and the helping score of an occupation in Table A3 of the Appendix, suggesting that career concerns may be less of an issue in pro-social occupations. With richer data, one could further examine the relationship between hours and the probability of promotion in helping and non-helping occupations.

The model presented in Section I assumes that at the time when they choose their occupation, workers have perfect knowledge of the number of hours that they will be putting in. The theory also assumes that the premium $1 + p$ is observed before workers choose their hours. However, it is possible that at the time of job selection, the amount of hours that a salaried employee will need to put in is not known with certainty. If potential employees who value pro-social job characteristics tend to underestimate the amount of time needed to do their job in helping occupations, then this would lead to a negative relationship between long-hour premiums and the occupation's helping score.

To address to some extent this and other information-related issues, I use the NLSY79 data on job satisfaction to show that satisfaction is higher in helping occupations, especially for females. While these results do not rule out uncertainty about hours, it is plausible that employees would be dissatisfied if their expectations about the job did not align with actual outcomes. Researchers such as Hamermesh (2001) have pointed out that survey reports of job satisfaction are far from perfect measures of worker utility, but they are nonetheless predictive of objective labour market outcomes such as turnover and investments in firm-specific human capital, and are thus useful in understanding what drives worker welfare. Table 3 and Smith (2007) show that a large fraction of occupations in which job satisfaction tends to be very high also have high helping scores. Pro-social value is assumed to increase utility in the theory in Section I. Holding other factors constant, in particular wages and hours worked, a positive correlation between an occupation's helping score and its job satisfaction is consistent with the assumption of this characteristic being utility-increasing in addition to addressing the concerns related to uncertainty about the returns to long hours.

Job satisfaction in the NLSY79 is reported on a 4-point scale, and I estimate several versions of an empirical model relating occupational characteristics and satisfaction, taking advantage of the panel aspect of the data. The fixed effects models rely on occupational switching for identification; 69% of respondents in the sample are observed in more than one occupation, with most (51% of the sample) observed in two, three or four occupations over the course of the survey period. The estimation results are reported in Table 7. The top panel of the table presents results for males in the NLSY79 sample, while the results in the bottom panel are for women. All specifications include quadratics in actual experience, tenure and age, along with job zone and survey year indicators.

Table 7. Occupational Characteristics and Job Satisfaction: NLSY79 Results

	(1)	(2)	(3)	(4)
<i>Males</i>				
Helping	0.087** (0.038)	0.019** (0.008)	0.127*** (0.030)	0.081* (0.042)
Flexibility	0.007 (0.032)	0.003 (0.007)	0.010 (0.027)	0.009 (0.036)
Analytical	0.008 (0.038)	0.008 (0.009)	-0.016 (0.031)	0.035 (0.043)
Competitive	0.022 (0.027)	0.002 (0.006)	0.032 (0.022)	0.006 (0.031)
Usual hours	0.021*** (0.003)	0.005*** (0.001)	0.023*** (0.002)	0.020*** (0.003)
ln(wage)	0.638*** (0.060)	0.139*** (0.013)	0.573*** (0.045)	0.628*** (0.070)
<i>Females</i>				
Helping	0.158*** (0.033)	0.037*** (0.008)	0.184*** (0.024)	0.137*** (0.038)
Flexibility	0.041 (0.030)	0.017** (0.007)	0.053** (0.024)	0.061* (0.033)
Analytical	-0.088** (0.039)	-0.013 (0.010)	-0.085*** (0.031)	-0.058 (0.042)
Competitive	-0.011 (0.024)	-0.005 (0.006)	-0.025 (0.020)	-0.019 (0.026)
Usual hours	0.008** (0.004)	0.001 (0.001)	0.010*** (0.003)	0.004 (0.005)
ln(wage)	0.554*** (0.069)	0.135*** (0.017)	0.366*** (0.050)	0.549*** (0.081)
Model	FE Logit	FE Linear	RE ordered Logit	FE ordered logit

Notes

Standard errors are in parentheses. Data from the 1982–2014 surveys. The sample is limited to full-time (35+ hours) full-year (48+ weeks) workers in the occupations included in the earlier analysis. The estimation is based on 25,553 observations for 3688 male workers and 24,584 observations for 3805 female workers. The dependent variable in column (1) equals 1 if very satisfied with current job, and 0 otherwise. In all other specifications the dependent variable measures job satisfaction on a 4-point scale. Additional controls include actual experience and experience squared, tenure and tenure squared, age and age squared, and job zone and survey year indicators.

*, **, *** indicate $p < 0.10$, $p < 0.05$, $p < 0.01$, respectively.

The results in column (1) of Table 7 are coefficients from a logit model, in which the dependent variable equals 1 if a worker reports the highest level of job satisfaction, and 0 otherwise. This model includes worker fixed effects. The rest of the models that I estimate measure job satisfaction on a 4-point scale. The results in column (2) are from a linear fixed

effects model; those in column (3) are from a random effects ordered logit model. Finally, column (4) presents the coefficient estimates associated with the ‘blow-up and cluster’ (BUC) estimator introduced in Baetschmann et al. (2015), which is a fixed effects conditional logit estimator.

In all cases, the helping score of an occupation is positively correlated with reported job satisfaction, and the estimated coefficients are larger for women by as much as a factor of 2. This result is consistent with the idea that workers value occupational social value, and women do so more than men. These results should be interpreted with caution, however, because it is also possible that reported job satisfaction is subject to more pronounced social desirability bias in helping occupations. Women's job satisfaction also increases in the occupation's level of flexibility, while the estimated coefficients for men are very close to zero in all specifications. As we would expect, satisfaction is higher in jobs that pay higher hourly wages. Working hours are also strongly positively correlated with job satisfaction for men, but this relationship is much less pronounced for women and is contingent on the choice of specification.

Overall, the results in Table 7 are supportive of the assumptions that are central to the theoretical framework in this paper. Namely, workers tend to be happier in jobs that offer more pro-social benefits, and this relationship is stronger for women.

6 Conclusion

This paper presents a conceptual framework and shows empirical evidence in support of the argument that for salaried employees, the utility from working in a pro-social occupation is a substitute for monetary compensation for long working hours. The results contribute to our understanding of the incentives faced by salaried employees. Few previous studies (e.g. Gregg et al. 2011) have explored the idea that workers in certain occupations may be more likely to donate labour, and there has not been much discussion of the economic significance of such tendencies.

My study is the first to establish empirically that occupations in which workers are involved in helping or providing service to others tend to pay lower wage premiums for long hours, but workers in these occupations report high levels of job satisfaction and similar or higher levels of labour supply compared to workers in less helping but higher-premium occupations. I develop a selection-based conceptual framework to rationalize these trends, in which workers experience lower disutility of long hours in pro-social occupations. I use a composite measure of an occupation's tendency to provide help and service to others, derived from the O*NET database, as a measure of pro-social characteristics. I calculate empirically occupation-specific premiums for working long hours using the large samples of workers available in the 2000 Census and 2000–2016 American Community Surveys, focusing on occupations in which hourly pay is not common.

Further, I show that women earn lower premiums relative to men in pro-social occupations, which is also consistent with the selection-based theory when women derive higher utility from occupational social value. An implication of my results is that gender differences in preferences for occupational social value can explain part of the observed gender pay gap, particularly among long-hour employees.

An important policy-relevant issue related to the findings in the paper is whether extrinsic rewards can crowd out intrinsic motivation in helping occupations. In monopsonistic markets for jobs with pro-social characteristics, such as the one for nurses as modelled in Heyes (2005),

increasing the wage is expected to reduce average worker quality. At the same time, in their recent study of nurse recruitment practices, Ashraf et al. (2018) show that when employers screen on ability when making hiring decisions, adding career incentives to a pro-social job does not lower the average level of intrinsic motivation of new workers. Because the theoretical framework in Section I assumes competitive frictionless markets, it does not allow one to draw conclusions about potential efficiency gains or losses from increasing the wage in helping occupations. Whether increasing the long-hour premium in pro-social occupations would increase or decrease productivity at long hours in occupations in which it is important to attract intrinsically motivated workers remains to be examined in future work.

Acknowledgments

I am thankful to Joseph Altonji, Barry Hirsch, Julie Hotchkiss, Ronald Oaxaca, Jamin Speer and participants at the CSWEP session of the 2016 ASSA Annual Meeting for valuable comments.

Notes

1. Author's calculations based on Census and American Community Survey data.
2. The 2008–16 instalments of the ACS report weeks worked only as an interval. I assign the value of 48 to the 48–49 interval and 52 to the 50–52 interval because among observations from the 2000 Census, 81% of values in the 48–49 interval equal 48, and 92% of values in the 50–52 interval equal 52. Using the midpoints of these intervals instead (48.5 and 51 weeks) has little impact on the empirical results.
3. The social interest definition includes ‘These occupations often involve helping or providing service to others’. The service orientation requirement is defined as ‘Actively looking for ways to help people’. The description of the relationships value states: ‘Occupations that satisfy this work value allow employees to provide service to others and work with co-workers in a friendly non-competitive environment. Corresponding needs are Co-workers, Moral Values and Social Service.’ The other two measures are self-explanatory.
4. The discussion in this paper is focused on salaried employees, but the Census/ACS earnings data do not provide information on time unit of pay. In the NLSY79, 54.5% of observations for full-time full-year workers whose occupations would be classified in job zones 3, 4 or 5 report annual or monthly earnings, with the fraction increasing with job zone. The corresponding fraction is 23.2% for full-time full-year observations that would be categorized into job zones 1 or 2.
5. The conversion procedure starts with the 1980 Census classification. I drop observations from the 1979–81 surveys, when occupations are coded only with the 1970 Census codes.
6. The survey question asks: ‘How do you feel about your job?’ The possible responses are ‘Like it very much’, ‘Like it fairly well’, ‘Dislike it somewhat’ and ‘Dislike it very much’.
7. I also estimated predicted wages using a Poisson model with the level of wages as the dependent variable (see, for example, Nichols 2010). The coefficient of correlation between the two versions of the estimated premiums exceeds 0.9, and the results are robust to using this alternative specification.
8. To simplify the notation, for the remainder of the paper I use p instead of \bar{p}_j .

9. Hirsch and Manzella (2015) find evidence of wage penalties for jobs that score high on the ‘Concern for others’ work style and the ‘Assisting and caring for others’ work activity.
10. Additional occupational characteristics derived from the O*NET data are considered in Table A4 of the Appendix. Because some of the variables are highly collinear in a regression model, I show only the coefficient of correlation between the occupation premium and each O*NET characteristic.
11. I am grateful to a referee for suggesting this distinction.

References

- Aidt, T. and Dallal, B. (2008). Female voting power: the contribution of women’s suffrage to the growth of social spending in Western Europe (1869–1960). *Public Choice*, 134(3), 391–417.
- Ashraf, N., Bandiera, O. and Lee, S. (2018). Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. Working Paper, London School of Economics and Political Science.
- Autor, D. H., Levy, F. and Murnane, R. J. (2003). The skill content of recent technological change: an empirical exploration. *Quarterly Journal of Economics*, 118(4), 1279–333.
- Baetschmann, G., Staub, K. E. and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3), 685–703.
- Bell, L. A. and Freeman, R. B. (2001). The incentive for working hard: explaining hours worked differences in the US and Germany. *Labour Economics*, 8(2), 181–202.
- Bertrand, M. (2011). New perspectives on gender. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 4. Amsterdam: Elsevier, pp. 1543–90.
- Besley, T. and Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, 95(3), 616–36.
- Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. In J. J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol. 5. Amsterdam: Elsevier, pp. 3705–843.
- Brekke, K. A. and Nyborg, K. (2010). Selfish bakers, caring nurses? A model of work motivation. *Journal of Economic Behavior & Organization*, 75(3), 377–94.
- Bureau of Labor Statistics, US Department of Labor (2013). *Occupational Employment Statistics*; available online at www.bls.gov/oes (accessed 27 July 2019).
- Carpenter, J. and Gong, E. (2016). Motivating agents: how much does the mission matter? *Journal of Labor Economics*, 34(1), 211–36.
- and Myers, C. K. (2010). Why volunteer? Evidence on the role of altruism, image, and incentives. *Journal of Public Economics*, 94(11–12), 911–20.
- Chevalier, A. (2007). Education, occupation and career expectations: determinants of the gender pay gap for UK graduates. *Oxford Bulletin of Economics and Statistics*, 69(6), 819–42.
- Collewet, M. and Sauermann, J. (2017). Working hours and productivity. *Labour Economics*, 47, 96–106.
- Cortes, P. and Pan, J. (2019). When time binds: substitutes for household production, returns to working long hours, and the skilled gender wage gap. *Journal of Labor Economics*, 37(2), 351–98.

- Devaro, J. and Brookshire, D. (2007). Promotions and incentives in nonprofit and for-profit organizations. *Industrial & Labor Relations Review*, 60(3), 311–39.
- Eckel, C. C. and Grossman, P. J. (2008). Differences in the economic decisions of men and women: experimental evidence. In C. R. Plott and V. L. Smith (eds), *Handbook of Experimental Economics Results*, Vol. 1. Amsterdam: Elsevier, pp. 509–19.
- England, P., Budig, M. and Folbre, N. (2002). Wages of virtue: the relative pay of care work. *Social Problems*, 49(4), 455–73.
- Erosa, A., Fuster, L., Kambourov, G. and Rogerson, R. (2017). Hours, occupations, and gender differences in labor market outcomes. NBER Working Paper no. 23636.
- Ewing, B. T. (1996). Wages and performance-based pay: evidence from the NLSY. *Economics Letters*, 51(2), 241–6.
- Fehrler, S. and Kosfeld, M. (2014). Pro-social missions and worker motivation: an experimental study. *Journal of Economic Behavior & Organization*, 100, 99–110.
- Filer, R. K. (1986). The role of personality and tastes in determining occupational structure. *Industrial and Labor Relations Review*, 39(3), 412–24.
- Fortin, N. M. (2008). The gender wage gap among young adults in the United States: the importance of money versus people. *Journal of Human Resources*, 43(4), 884–918.
- Gicheva, D. (2013). Working long hours and early career outcomes in the high-end labor market. *Journal of Labor Economics*, 31(4), 785–824.
- Goldin, C. (2014). A grand gender convergence: its last chapter. *American Economic Review*, 104(4), 1091–119.
- Gregg, P., Grout, P. A., Ratcliffe, A., Smith, S. and Windmeijer, F. (2011). How important is pro-social behaviour in the delivery of public services? *Journal of Public Economics*, 95(7–8), 758–66.
- 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), 827–74.
- Hamermesh, D.S. (2001). The changing distribution of job satisfaction. *Journal of Human Resources*, 36(1), 1–30.
- Heyes, A. (2005). The economics of vocation or ‘Why is a badly paid nurse a good nurse?’. *Journal of Health Economics*, 24(3), 561–9.
- Hirsch, B. T. and Manzella, J. (2015). Who cares—and does it matter? Measuring wage penalties for caring work. In S. W. Polachek, K. Tatsiramos and K. F. Zimmermann (eds), *Gender Convergence in the Labor Market*. Bingley: Emerald Group Publishing, pp. 213–75.
- Hotchkiss, J. L. and Rupasingha, A. (2018). Wage determination in social occupations: the role of individual social capital. In S. W. Polachek and K. Tatsiramos (eds), *Transitions Through the Labor Market: Work, Occupation, Earnings and Retirement*. Bingley: Emerald Publishing, pp. 127–81.
- Landers, R. M., Rebitzer, J. B. and Taylor, L. J. (1996). Rat race redux: adverse selection in the determination of work hours in law firms. *American Economic Review*, 86(3), 329–48.
- Lemieux, T., Macleod, W. B. and Parent, D. (2009). Performance pay and wage inequality. *Quarterly Journal of Economics*, 124(1), 1–49.
- Miller, G. (2008). Women’s suffrage, political responsiveness, and child survival in American history. *Quarterly Journal of Economics*, 123(3), 1287–327.
- Nichols, A. (2010). Regression for nonnegative skewed dependent variables. BOS10 Stata Conference, Stata Users Group.

- Penvael, J. (2015). The productivity of working hours. *Economic Journal*, 125(589), 2052–76.
- Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J. and Sobek, M. (2018). IPUMS USA: Version 8.0 [dataset]; available online at <https://doi.org/10.18128/D010.V8.0> (accessed 27 July 2019).
- Scopp, T. S. (2003). The relationship between the 1990 census and census 2000 industry and occupation classification systems. US Census Bureau, Washington, DC.
- Smith, T. W. (2007). Job satisfaction in the United States. NORC/University of Chicago.
- Taylor, L. J. (2007). Optimal wages in the market for nurses: an analysis based on Heyes' model. *Journal of Health Economics*, 26(5), 1027–30.
- Trejo, S. J. (1991). The effects of overtime pay regulation on worker compensation. *American Economic Review*, 81(4), 719–40