THE RETURN TO HOURS WORKED WITHIN AND ACROSS OCCUPATIONS:
IMPLICATIONS FOR THE GENDER WAGE GAP

Jeffrey T. Denning
Brian Jacob
Lars Lefgren
Christian vom Lehn

The Return to Hours Worked Within and Across Occupations: Implications for the Gender
Wage Gap
Jeffrey T. Denning, Brian Jacob, Lars Lefgren, and Christian vom Lehn
NBER Working Paper No. 25739
April 2019
JEL No. J16,J3,J7

ABSTRACT

We document two empirical phenomena. First, the observational wage returns to hours worked within occupation is small, and even negative in some specifications. Second, the wage return to average hours worked across occupations is large. We develop a conceptual framework that reconciles these facts, where the key insight is that workers choose jobs as a bundle of compensation and expected hours worked. As an example, we apply this framework to the gender wage gap and show how it can explain the view expressed in recent work that hours differences between men and women represent a large and growing component of the gender wage gap.

Jeffrey T. Denning
Brigham Young University
435 Crabtree Building
Provo, UT 84604
jeffdenning@byu.edu

Brian Jacob
Gerald R. Ford School of Public Policy
University of Michigan
735 South State Street
Ann Arbor, MI 48109
and NBER
bajacob@umich.edu

Lars Lefgren
Department of Economics
Brigham Young University
130 Faculty Office Building
Provo, UT 84602
and NBER
lars_lefgren@byu.edu

Christian vom Lehn
Brigham Young University
435 Crabtree Building
Provo, UT 84602
cvomlehn@byu.edu

1. Introduction

Employed women work, on average, three to five hours fewer than employed men, even controlling for demographics and occupation.[1] A recent literature, including Bertrand et al. (2010), Gicheva (2013), Cha and Weeden (2014), Goldin (2014), Goldin and Katz (2016), and Cortes and Pan (Forthcoming), has argued these gender differences in hours worked and preferences for job flexibility are crucial to understanding the gender pay gap.[2] In order for hours worked to have a significant bearing on the gender *wage* gap, it must be the case that the wage return to hours worked is meaningful. While prior research has documented significant heterogeneity in the income elasticity of hours worked across occupations, the average estimated income elasticity is close to unity.[3] This implies that the *wage-hours* elasticity is close to zero and thus incapable of explaining, in an accounting sense, a substantial fraction of the gender wage gap across all workers.

Consequently, we carefully examine the relationship between weekly hours worked and wages. We show that weekly hours worked is not strongly associated with higher wages *within* occupations.[4] Using individual data from the 2016 ACS, we find that a 10% increase in hours is associated with 1% lower hourly wages.[5] Even accounting for measurement error, wages tend to be only weakly related to hours worked. In contrast, we show that *across* occupations, there is a large, positive relationship between occupation-level average hours worked and occupation-level average wages. Again, using the 2016 ACS, we estimate a 10% increase in the average hours

---

[1] Based on 2016 ACS data; see Section 4 for more details.
[2] Blau and Kahn (2017) offer an excellent overview of the evidence on the size and explanations for the gender wage gap in the cross-section and over time.
[3] See, for example, Figure 3 of Goldin (2014) and Table 2 of Cortes and Pan (forthcoming).
[4] Some occupations have non-negative wage returns, see Goldin (2014) for discussion.
[5] Based on a Mincer regression with controls for standard demographics and occupation fixed effects. See column 3 of Table 1.

worked in an occupation is associated with a 20.3% increase in average wages in the occupation.[6]

To reconcile the differences in returns to hours within and across occupations, we propose a simple framework relating hours worked and compensation. We envision a job as an employment contract where compensation is a function of both expected hours worked and actual hours worked. We allow the relationship between compensation and *expected* hours to be different than the relationship between compensation and *actual* hours.

A natural example of such a difference is found in salaried jobs, where the compensation implicitly depends on the amount of work to be done (expected hours), but at the margin, wages will have a negative relationship with actual hours worked. In jobs paid at hourly rates, the relationship between hours worked and wages is naturally small, though may be larger than zero when accounting for the possibility of overtime work. However, at the occupation level, compensation for wage workers can increase with occupation level hours worked, as workers who commit to more expected hours may receive a wage premium.

It may be the case that even gender differences in hours within occupation may be due to differences in expected as opposed to actual hours. This is consistent with Cortes and Pan (Forthcoming), who conjecture that the large pay gains associated with closing gender hours gaps suggests that relaxing hours constraints on women "enabled [them] to enter a different job… within the same occupation" and that "shifts in women's position in the males' earnings distribution are unlikely to occur if women simply worked more hours and stayed in the same job." It is also consistent with the findings of Hirsch (2005) who finds, using a person fixed effects design, that the wage losses associated with part time work largely disappear when one

---

[6] Based on occupation regression with residualized log wage and residualized log hours. See table 4 column 2.

3

controls for industry, occupation, and task measures. If the difference in hours worked by men and women within occupation is largely driven by differences in expected hours worked across jobs, then pricing male-female hours differences may be best accomplished using the occupation-level relationship between wages and hours.[7]

Empirically, we find that if we account for gender differences in hours using the price implied by a regression of wages on *individual* level hours, this actually widens the residual gender wage gap, as increased hours are associated with lower wages. In contrast, if we price hours based on the cross-occupation wage-hours relationship, even controlling for occupation task characteristics, the residual wage gap drops by half. If we perform the same exercise, but do not control for occupation level task characteristics, the residual gender wage gap closes almost entirely. These pricing exercises indicate that, as argued in prior literature, hours worked may play a critical role in accounting for the gender wage gap, and the degree to which this is important depends on the nature of hours differences between men and women.

We also document that the relationship between hours worked and wages has been increasing over time both at the individual and occupational level. Given the gender gap in hours, this finding suggests that women have been swimming upstream in terms of achieving wage-parity with men. When pricing gender differences in hours at the individual level from 1980 to 2016, these changing returns to hours worked exacerbate the gender wage gap by nearly 20%, similar to the conclusions of Cha and Weeden (2014). However, when we use the occupational-level relationship between hours and wages to price gender hour differences, our estimates imply that the gender wage gap would be as much as 46% lower if returns to hours

---

[7] Supporting evidence for this is found in Cortes and Pan (Forthcoming), who find that when hours constraints on women are relaxed, this results in different occupational choices and not simply more hours worked within an occupation.

worked had remained at 1980 levels. Thus, even as the educational attainment of women has surpassed that of men and they have been increasingly employed in higher wage occupations, this increasing returns to hours worked has worked against wage-parity.

We view our paper as complementary to the existing literature on the role of hours in explaining gender gaps in compensation.[8] Much of this literature has documented important heterogeneity in the returns to hours worked across occupations, such as Bertrand et al. (2010), Goldin (2014), and Goldin and Katz (2016), who show much higher returns to working longer hours in law and business occupations than occupations such as pharmacists, and correspondingly larger gender wage gaps. We focus on the average level of the gender wage gap and returns to hours worked for all workers and not the variation in gaps and returns across occupations. Our analysis demonstrates that the returns to expected hours worked may be higher *on average* than previously believed, suggesting an important role for hours worked in mediating the gender wage gap.

By examining the relationship between wages and hours, it is natural to question how our paper relates to the voluminous literature on labor supply. Historically, this literature has typically assumed that workers can choose their hours in response to wages. However, in many settings workers do not have complete discretion over hours worked which complicates interpretation of the relationship between wages and hours worked.[9] In particular, Altonji and Paxson (1988, 1992) focus on changes in hours and wages among job switchers, arguing that job switching is a primary channel for workers to adjust their hours of work and associated

---

[8] We acknowledge that other job characteristics beyond hours are important in thinking about the gender wage gap. For example, a job with relatively low hours but rigid requirements about scheduling may be disproportionately unappealing to women. We abstract from them in the current paper as is common in this literature (Goldin (2014), Cortes and Pan (Forthcoming), Cha and Weeden (2014)).

[9] In some atypical settings it is quite plausible that workers can adjust hours such as stadium vendors (Oettinger 1999) or taxi cab drivers (Farber 2015).

compensation. More recently, Chetty et al. (2011) study labor supply decisions in a framework where firms set hours constraints on workers and thus changes in hours worked induced by wage changes for individuals typically come when switching jobs. Our analysis of the cross-occupation wage-hours relationship is consistent with these studies, suggesting that choices across jobs and occupations are crucial to individuals' labor supply decisions.

We continue our analysis by describing a set of empirical facts regarding the relationship between hours worked and wages. We then sketch a theoretical framework to rationalize these facts. We finish by relating our empirical findings and theoretical insights to the gender wage gap.

2. Empirical Facts

We begin this part of the analysis by establishing three facts. First, the relationship between hours worked and wages at the individual level is consistently small. Second, the relationship between hours worked and wages at the occupation level is substantively larger. Third, since 1980, the occupation-level relationship between hours worked and wages has grown.

We use individual-level data on hours worked and income from the Current Population Survey, the Census and the American Community Survey (ACS). We focus on the civilian non-institutionalized population of workers between ages 25 and 55. We drop observations with missing occupation data, as well as individuals who are self-employed. Our primary measure of weekly hours worked is given by an individual's reported usual weekly hours worked for the prior year. Hourly wages are computed by dividing total reported wage and salary income for the prior year by the product of usual hours worked and weeks worked in the previous year.

6

Where incomes are top coded, we set them to be 1.5 times above the top coded level, and we drop observations with reported wages below half the minimum wage. We also drop observations with imputed usual hours, weeks worked, income or occupation.

The OLS Mincerian relationship between hours worked and wages may not reflect a causal relationship. They do, however, serve as a useful benchmark for illustrating the challenge of accounting for the gender wage gap using conventional estimates of the return to hours worked. Additionally, both Goldin (2014) and Cortes and Pan (Forthcoming) use this relationship to categorize occupations into ones in which there is a high or low return to hours worked. Our cross-occupation analysis is novel and suggests that the returns to hours worked are much higher than can be explained by conventional Mincer estimates—strengthening the empirical foundation for the conjecture that hours worked are central to understanding the gender wage gap. We view our analysis as complementary to that of Cortes and Pan (Forthcoming) who present quasi-experimental evidence that when barriers to labor supply are relaxed, high-skill women work more and the gender pay gap is reduced. Of course, one must also exert caution in interpreting these cross-occupation returns to hours worked, which we discuss in greater length later.

Figure 1 shows the relationship between raw wages and hours at both the individual (left panel) and occupational level (right panel). To illustrate the relationship between actual individual hours worked and wages, we bin individuals into 1000 groups corresponding to the logarithm of number of weekly hours they report usually working. For occupations, we compute the average log hours and average log wages for occupations at the three digit level. The average log wages and average log hours for each individual bin or occupation are plotted as open circles on the figure, with the size of the circles corresponding to the size of the bin or occupation, as

determined by survey weights and the number of workers in each bin/occupation. The dark line is the best fit line from a bivariate regression.

For individual level hours and wages, the relationship is not particularly strong, with a slope of roughly 0.25. Without any adjustment for individual characteristics, this implies that individuals who work more hours per week only earn slightly more per hour than those who work fewer hours. In contrast, the relationship between occupational averages for hours and wages is much stronger, with a slope of 2.35. Thus, individuals working in more time intensive occupations have a substantially higher hourly wage than individuals working in less time intense occupations.

This difference between the wage-hour relationship at the occupation level versus individual level is robust across various samples and specifications. We first illustrate the consistency of the wage-hour relationship at the individual level for several specifications. Table 1 reports estimates from a standard Mincer wage regression using individual level data from the 2016 American Community Survey (ACS). Specifically, we estimate variants of the following model:

$$(1) \qquad ln\, w_{io} = \beta_0 + \beta_1\, ln\, h_{io} + X_i \beta_2 + \delta_o + \varepsilon_{io}$$

where i indexes individuals and o indexes measured 3 digit occupation. The variables w and h represent wages and hours respectively, and X_i is a vector of individual covariates including binary indicators for race (black, Hispanic, Asian and other race) and gender, indicators for educational attainment (high school only, some college and BA or higher), and a quartic in age. δ_o represents a set of occupation fixed effects. All regressions are weighted with the person weights provided by ACS, and robust standard errors are clustered by three-digit occupation.

Column 1 of Table 1 presents estimates from the bivariate regression of log wage on log hours, and corresponds closely to Figure 1. This indicates that a 10 percent increase in usual hours worked is associated with 2.4 percent higher hourly wages. The inclusion of race, gender, age and educational attainment in column 2 further reduces the coefficient on log hours.

To this point, we have not controlled for occupation fixed effects, however, and thus these point estimates represent some weighted average of the returns to hours within and across occupations. Including occupation fixed effects in column 3, we isolate the return to hours within occupations and we find that this causes the estimated return to flip signs. Within occupation, a 10 percent increase in hours is associated with a 1.1 percent *lower* hourly wages.

Columns 4-5 show that this same basic pattern holds true for men and women. In results available upon request, we show that the same general relationship holds for those with and without a BA degree, and in samples limited to full-time, full-year workers, defined as those working at least 40 weeks in the previous year and at least 35 hours per week.

Tables 2 and 3 shows that these relationships persist when using individual level data from the March CPS pooled across years 2012-2017, including corrections for measurement error and separately addressing hourly wage and salaried workers. Columns 1-3 mirror the results from Table 1 that use ACS data. As in the ACS data, the bivariate relationship is positive and significant (elasticity of .195) but falls to zero with the inclusion of basic controls and becomes negative with the inclusion of occupation fixed effects. However, the hours worked variable may be measured with error in CPS and ACS, which will introduce a negative bias in our estimate of the impact of hours worked. To address this measurement error, the specification in column 4 instruments for a worker's reported usual hours this year with the usual hours the individual reported in the prior March. Although the coefficient increases from -.14 to .15, it

9

remains small. Alternatively, column 5 instruments for a worker's reported usual hours this year

using the number of hours the worker reported actually working in the prior week. The

coefficient is -.0004 and is statistically insignificant.[10]

Table 3 uses the CPS's Outgoing Rotation Groups and shows that the same general

pattern is present for wage earners and salary workers.[11] With the inclusion of occupation fixed

effects, the coefficient on log hours is .17 for wage earners. This rises slightly to .23 when we

instrument log hours using lagged hours. In contrast, salary workers have a negative relationship

between log hours and wages of -.22, after including individual level controls and occupation

fixed effects. Instrumenting for log hours with lagged hours increases the coefficient to .02,

which is statistically indistinguishable from zero.

In contrast to the consistently small estimated relationship between hours and wages at

the individual level, the relationship between hours and wages at the occupational level is

consistently large. Table 4 reports estimates from wage-hour regressions at the three-digit

occupation level in the 2016 ACS, corresponding to the aggregate relationship shown in Figure

1.[12] Specifically, we estimate variants of the following model:

$$(2) \qquad \overline{w}_o = \beta_o + \beta_1 \overline{h}_o + \beta_2 tasks_o + \varepsilon_o$$

where $\overline{w}_o (\overline{h}_o)$ indicate average of log wages (hours) in occupation o. The coefficient of 2.35 in

column 1 indicates that a 10 percent increase in hours worked is associated with a roughly 23.5

percent higher hourly wage. This relationship remains large when we residualize hours and

---

[10] Using either of these instruments reduces the sample size to some extent, however, while unreported, re-estimating the specifications in Columns 1-3 of the table with this smaller sample size does not meaningfully change the coefficients.

[11] Sample restrictions are identical to the previous two samples. We link individuals across years as in the ASEC sample. One distinction from the ASEC and ACS/Census samples is that wages for wage earners in the ORG sample are collected directly rather than computed, and wages of salary earners are computed by dividing weekly earnings by usual hours worked.

[12] We calculate respondent-weighted averages for each occupation.

wages. To construct these residuals, we regress log wage (hours) on the same demographics used at the individual level (shown in column 3 of Table 1) as well as a full set of occupation fixed effects. We use the coefficients on the occupation fixed effects as measures of average residualized log wage and log hours. Column 2 shows the results of the bivariate regression of residualized log wage on residualized log hours. The relationship between hours and wages is still very large, with an elasticity greater than 2.

However, average hours might be correlated with other aspects of the occupation that influence wages. To control for specific traits of each occupation, we follow the literature on occupational tasks, and in columns 3-4 control for four measures of the tasks associated with each occupation (Autor, Levy & Murnane 2003, Deming 2017).[13] While the return to average hours decreases, it remains significantly large. In column 4, for example, the estimates suggest that a 10 percent increase in average hours worked is associated with 11.7 percent higher hourly wages. In columns 5-8, we replicate the specification from column 1 with different groups of occupations. Column 5 (6) shows the results for high (low) skill occupations, which we define as those above (below) the median in terms of fraction of workers in the occupation with at least a BA degree. Column 7 (8) shows the results for female (male) dominated occupations, which we define as those above (below) the median in terms of fraction female in the occupation. In all cases, the coefficients are large, positive and significant. In results not shown, we find that these results are very similar using the CPS.

The final empirical fact we establish is how these relationships in the cross-section have changed over time. Figure 2 shows how these relationships have changed over time as measured

---

[13] In particular we include abstract analytical, manual, routine and social tasks constructed from the ONET 4.0. For more detail on how these are constructed, see Appendix A. There are seven occupations for which task measures are unavailable, and thus these are dropped in these regressions. Running columns 1 and 2 with only occupations for which we have task data does not materially change our results.

from the Census and the ACS.[14] We plot the coefficient on log hours in each year corresponding to the regressions with individual controls and occupation fixed effects (Table 1 Column 3) and the residualized occupation regressions (Table 4 Column 2).

Looking at the individual level wage-hours estimates, we see the relationship has increased modestly between 1980 and 2016, but is always negative and small in absolute magnitude. These individual-level results are consistent with those documented by Kuhn and Lozano (2008), Cha and Weeden (2014), and Cortes and Pan (Forthcoming). Examining the relationship between occupation level wages and hours, we see a much more dramatic change. In 1980, a 10 percent increase in occupation average hours was associated with a 9.4 percent increase in average wages. By 2016, the relationship had doubled so that a 10 percent increase in average hours was associated with a 20.2 percent increase in average wages.[15]

As the returns to certain skills (e.g., cognitive ability and social skills) have increased over this period (Autor, Murnane and Levy 2003; Deming 2017), we show in Figure 3 the coefficients on hours at the occupation level controlling for tasks in each year. [16] We use the same measures of tasks as in Table 4. The inclusion of task measures attenuates the coefficients but the large in magnitude and positive coefficient on log hours remains across all time periods. For example, without task controls the coefficient is 2.02 in 2016 and falls to 1.12 with task controls.

---

[14] In running these regressions, we use contemporaneous occupation codes. However, though we do not report it, using time-consistent occupational codes does not impact our findings.

[15] We have calculated analogous measures for the CPS, looking at 5-year moving averages of the coefficients to take into account the smaller sample size. We also examined both OLS and IV estimates of the return to individual hours worked. For the IV specifications, we instrument usual hours worked in the reference year with usual hours worked in the prior year to account for measurement error. All these specifications generate a similar pattern.

[16] Ideally, we would be able to control for the changing task composition within occupations over time. However, O*NET only covers a limited period of time and is not easily mapped to its precursor, the Dictionary of Occupational Titles. As a result, we use time-invariant task characteristics from O*NET 4.0. See Appendix A for more details regarding task measurement.

To summarize, we find that the individual-level relationship between hours and wages is small, negative within occupations and that it has changed only modestly over time. The occupational level relationship between hours and wages is large has been growing over time.

3. Conceptual Framework

Our descriptive analysis demonstrates that the return to average hours worked at the measured occupation are systematically higher than the corresponding return to a person's actual hours worked within the occupation. In this section, we present a framework for thinking about this empirical result that has the potential to shed light on other economic phenomena of interest.

We assume that a job, j, is an employment contract offered by a firm and is associated with a compensation level, $c(h_j^e, h_j)$, which is a function of expected hours worked, h_j^e, and actual hours worked, h_j. These contracts are determined in hedonic equilibrium by firms offering contracts consistent with profit maximization and workers choosing jobs consistent with utility maximization. Consistent with our empirical results, we allow the return to expected hours to differ from actual hours.

It is helpful to consider a few illustrative examples of common compensation schedules that have this form and are consistent with our empirical observation. Consider first the case of salaried workers. In this case, workers agree to a level compensation based on an expected workload but the salary does *not* adjust in the short run to actual hours worked. This compensation is determined in a simple hedonic equilibrium between workers and firms. Actual hours worked may vary from expected hours due to variation in skill across workers, how busy the firm is, idiosyncratic expectations of a demanding or lenient supervisor, or other reasons. In the context of our framework, this implies that the compensation function has the following

form: $c(h_j^e, h_j) = c(h_j^e)$. This leads in a straightforward fashion to a potentially large return to

expected hours across jobs if $c'(h_j^e)$ is large, even as the individual's personal compensation to

actual hours worked is zero. The implication for the actual hourly wage, $\frac{c(h_j^e)}{h_j}$, is that the effect of

a marginal hour worked within a job is to *reduce* the wage, a phenomenon consistent with some

of our descriptive analysis.

While many jobs are salaried, other jobs pay wages per hour worked. In such jobs, it

seems natural to model the compensation function in the following way: $c(h_j) = wh_j$, where w

is the hourly wage. In such jobs, earnings scale linearly with hours worked. Even in these cases,

however, the firm may place restrictions on the range of hours offered to the worker. As a

consequence, it may not be possible to obtain a high wage hourly job without committing to an

expected hours of work, h_j^e. In this case, the menu of compensation across jobs may take the

following form: $c(h_j^e, h_j) = w(h_j^e)h_j$. For an individual working an extra hour within a given

job, the wage is fixed at $w(h_j^e)$. However, moving across jobs, the wage may increase in

expected hours if in equilibrium $w'(h_j^e) > 0$. Hence, in hourly jobs the wage return of an

increase in expected hours across jobs may be quite high, even as the wage is constant within a

job.

A natural exception to this is the case where workers are paid a higher hourly rate when

they are required to work overtime. This would suggest a higher relationship between actual

hours worked and wages, which is consistent with what we observe when splitting our results

into hourly wage workers and salaried workers in Table 3. However, if the possibility of

overtime is only available in jobs with already high expected hours, this still implies that

expected hours play a potentially substantial role in a worker's compensation.

Both of these examples suggest that the wage return to expected hours of work across jobs may be substantially larger than the wage return of actual hours worked. Our descriptive analysis of the relationships between hours and wages is consistent with this conjecture. We interpret the returns to hours across occupations as representing a useful approximation of the return to expected hours, h_j^e. We believe that occupation average hours represents a good proxy for expected hours since it is averaged across many jobs. Our preferred estimate of this is 2.03 from column 2 of Table 4.

The estimate that most closely relates to the returns to actual hours within jobs (or deviations from expected hours) is -.11 from column 3 of Table 1 (where we include occupation fixed effects). This estimate only partially reflects the returns to actual hours because some of the variation in hours worked within occupation may represent differences in expected hours of different jobs within three digit occupation. However, given that this estimate is so small, this is consistent with the returns to actual hours being small, if not substantially negative.

4. Application to the Gender Wage Gap

Recent research emphasizes the role that job intensity (as measured by hours worked) plays in the gender wage gap. Work in economics (Goldin 2014) and sociology (Cha and Weeden 2014) suggests that the increasing prevalence of long work hours as well as the high return to job intensity has slowed the convergence of the gender wage or earnings gap. It is well documented that, on average, female workers work fewer hours than working men. In the 2016 ACS, for example, employed women work an average of 39.0 hours compared with 43.5 for men, a difference of 4.5 hours. After accounting for race, age and education, this difference increases to 4.8 hours. Controlling for three-digit occupation fixed effects as well, we find that

15

women still work 3.2 fewer hours than men.[17] In 1980, the raw gender difference was 6.7 hours and falls to 5.0 hours after accounting for comparable demographics and occupation fixed effects.

When thinking about how hours worked impacts the gender wage gap, it is helpful to write a simple decomposition of the gap that focuses on the role of hours worked. Consider the following regression similar to the earlier specifications but focusing on the residual gender wage gap, which we denote by α. Given our focus on the role of hours worked, we abstract from other covariates for expositional clarity, but they can be added easily.

$$(3) \qquad ln\, w_i = \pi + \alpha female_i + \beta\, ln\, h_i + \varepsilon_i$$

If we take expectations of this equation separately by gender, we obtain the following:

$$(4) \qquad \overline{ln\, w}_{female} = \pi + \alpha + \beta\, \overline{ln\, h}_{female}$$

$$(5) \qquad \overline{ln\, w}_{male} = \pi + \beta\, \overline{ln\, h}_{male}$$

We can thus express the gender wage gap as

$$(6) \qquad gap_w = \alpha + \beta gap_h$$

where $gap_w = \overline{ln\, w}_{female} - \overline{ln\, w}_{male}$ and $gap_h = \overline{ln\, h}_{female} - \overline{ln\, h}_{male}$.

Note that the gender gap in hours is priced at rate β and α represents the residual wage gap. Given the positive hours gap between men and women, as β rises, the residual wage gap will decline. Indeed, if β becomes large enough, it is possible that the residual wage gap could be negative, even with a substantial unconditional gender wage gap. As we have demonstrated above, estimates of β at the individual level regressions show hours worked has a weak positive or even negative correlation with wages (see Table 1). Thus, a simple regression adjustment for hours worked does little to close the residual gender wage gap.

---

[17] If we limit our analysis to full-time full-year workers, women work 1.8 fewer hours than men in 2016.

Our conceptual framework offers different insights on how to handle gender differences in hours worked. If the gender difference in hours worked is due to differences in expected hours worked, h_j^e, correcting for differences in hours worked using individual estimates will systematically *understate* the contribution of hours worked to the gender wage gap. If instead the gender difference in hours worked is related to actual hours worked, h_j rather than expected hours h_j^e, then using individual level estimates of the returns to hours worked would be more appropriate.[18] Because we do not observe whether an individual's actual hours worked is due to expected hours in the job or deviations from expected hours, researchers need to make an assumption regarding which estimate of the returns is appropriate.

In column 1 of Table 5, we present a baseline residual wage gap in which we do not account for hours or occupational choice at all. We do so by regressing log wages on a female indicator variable as well as race, age, and education level. The coefficient on the female indicator indicates a residual wage gap of .25 log points. We now consider alternative specifications in which we make different assumptions regarding how hours should be taken into account.

In column 2, we add to our minimal regression a simple control for the log of actual hours worked. In this specification we do not control for occupation fixed effects as occupational choice may be driven by differences in hours. This model implicitly assumes that there is no difference in the returns between actual and expected hours. The coefficient on log hours worked is quite small at .06, and reflects the return to some combination of actual and expected hours (along with other occupation-specific factors that might be correlated with expected hours

---

[18] Actual hours worked in individual level regressions will inherently be a mix of across-job variation in hours worked and within-person variation in hours worked.

and wage). As a consequence, it does little to shrink the residual wage gap, which falls to .24. This clearly illustrates the difficulty in explaining the gender wage gap by a simple regression adjustment for hours worked.

In column 3, we add occupation fixed effects but eliminate the control for hours worked. This specification implicitly controls for differences between men and women in the expected hours associated with their occupational choices at the level of the 3-digit occupation codes we observe. The inclusion of occupation fixed effects substantially reduces the residual gender wage gap to .16.[19] This substantial reduction is due to the fact that the fixed effects control for not only the expected occupational hours, but also any other occupational differences (e.g. tasks) that drive compensation.

In column 4, we again include occupation fixed effects but also control for actual hours worked. This regression prices the hour-wage relationship at the individual level. In this model, the coefficient on log hours should be interpreted as the return to actual hours worked driven by idiosyncractic factors such as supervisor requests, speed at completing tasks, etc. We estimate a residual wage gap of .17, *larger* than what we find in the prior specification (column 3) in which we do not control for hours. This is due to the fact that actual hours have a negative price in the regression. As a consequence, the fact that women work fewer hours within occupations means that they should have higher wages than men.[20] Hence, when we price hours based on the individual-level wage equations, the residual gender wage gap essentially does not change at all.

---

[19] Additionally, we have experimented with further including interaction terms between hours worked and occupation fixed effects, which would allow for an occupation-specific wage-hours elasticity. However, including these does not significantly change the point estimate on the residual gender wage gap.

[20] In results available upon request, we show that accounting for measurement error in hours worked using the IV strategy in Tables 2 and 3 does not substantively change the estimated residual gender wage gap.

However, at the other extreme, one can instead assume that the hours differences between men and women, both across and within occupations, is driven entirely by gender differences in choices about expected hours worked. Under this assumption, researchers would wish to price the gender difference in hours at the expected hours wage rate. This implicitly assumes that the expected hours wage rate is the same across measured occupations as well as across jobs within a measured occupation. One tractable way to implement this thought experiment is to instrument individuals' actual hours with the occupation level average hours, omitting the individual's own contribution to this average.[21]

Column 5 of Table 5 shows these results. Consistent with our expectation, the coefficient on hours worked is 1.8, quite similar to the aggregate cross-occupation return to average hours worked presented earlier. In this specification, the residual gender wage gap virtually disappears. These results suggest that gender-correlated choices over job intensity could account for the lion's share of the difference in wages between men and women.

As we showed in Section 3, occupational average hours worked is correlated with job-related tasks. It is unclear whether it is appropriate to control for such factors. If we control for occupation task measures, this reduces the estimated coefficient on hours worked, which has the effect of increasing the residual wage gap. It may be that women prefer to work in jobs with task requirements correlated with high hours, such as abstract reasoning and team management, but do not enter these jobs because of the concomitant expectation of high hours. In this case, controlling for occupation tasks understates the importance of hours when estimating the gender wage gap. On the other hand, if women choose not to work in some high-hours jobs not because

---

[21] We do not control for occupational fixed effects, since they are collinear with the occupational average hours. This method is similar to that used in Angrist (1991).

of the hours, but rather because of the associated tasks, one would wish to account for the task mix of the job and other correlated characteristics including scheduling flexibility.

In column 6 of Table 5, we consider this possibility. Again, we instrument hours worked with average occupational hours (omitting the reference individual). However, we control in the second stage for the set of occupational tasks discussed earlier. This has the effect of lowering the return to hours worked to 1.2. Relative to the prior specification, controlling for task mix expands the residual gender wage gap to 0.08. The gap, however, remains substantially narrower than cases in which hours are priced at the actual hours rate implicit in the individual level regressions.[22]

Ultimately, with available data it is not knowable the extent to which the gender hours gap is driven by actual versus expected hours. However, evidence suggests that the difference in realized hours is a reflection of desire for lower expected hours among women and should be priced accordingly. In particular, recent work has shown that men and women have different willingness to pay for flexible work arrangements (e.g. Mas and Pallais 2017, Wiswall and Zafar 2017). Though unreported, we also find a substantial correlation between the gender hours gaps and wage gaps within occupations, even when controlling for demographic characteristics.[23] Thus, there is evidence that gender gaps in observed individual hours worked may be priced at a higher level than implied by standard Mincerian specifications.

---

[22] The number of observations is slightly lower in column 6 compared with columns 1 through 5. This is because a small number of occupations do not have ONET task information. If we replicate the specifications in columns 1-5 on the sample used in column 6, the results are qualitatively the same.

[23] Using the pricing implied by the slope on within occupation regressions of gender wage gaps and hours gaps with empirical Bayes corrections for measurement error implies a similar reduction in the gender wage gap as in Column 4 of Table 5, where we use the pricing of hours from occupation-level regressions controlling for tasks.

Gender Wage Gap over Time

The difference between women's log wages and men's log wages has shrunk considerably since the 1980s, but flattened in recent years, as shown by the blue trend line in Figure 4. Over this same period, the gender gap in log hours worked has closed somewhat (red trend line in Figure 4), although by much less, and in 2016 women still work notably fewer hours than men on average. The green trend line in Figure 4 illustrates the increasing returns to hours worked at the occupation level, which we estimated earlier and also show in Figures 2 and 3. Taken together, these facts suggest that wage convergence between men and women would have been larger if the wage premium for hours worked had remained the same. In this section, we consider what the evolution of the gender gap would have been if the returns to hours worked had remained at 1980 levels.

To begin, we estimate individual level wage regressions separately by year for 1980, 1990, 2000 and 2016. We do so both estimating the relation between wages and hours worked via OLS and also via IV in which we instrument individual hours with occupational average hours. This allows us to consider how the counterfactual wage gap would have evolved had the return to hours not increased. Our primary empirical specification is given by:

$$(7) \qquad ln\, w_{it} = \pi^t + \alpha^t female_i + \beta_1^t\, ln\, h_{it} + \gamma^t demos_{it} + \varepsilon_{it}$$

Because changing demographics and the coefficients on those demographics are not the focus of the analysis, while we control for them in our analysis, we ignore them for the purposes of presentation. Building on the logic we developed in our static analysis, we can write the gender wage gap (conditional upon demographic covariates) as

(8) $gap_w^t = \alpha^t + \beta_1^t gap_h^t$

This expression allows us to consider how the gender wage gap would have increased if the residual wage gap evolved according to the observed time series but the prices on hours remained constant at 1980 levels. Put another way, we allow the unexplained part of the gender wage gap to change period by period but fix the returns to hours worked at 1980 levels.

Again, we are faced with the question of what price should be used to relate the hours gap to the wage gap period by period.   In Table 6 we present these results for four different prices for log hours worked. In each of the cases, row 1 shows the evolution of the gap described in equation (8) above—that is, the residual wage gap plus the wage gap that can be explained by the gender gap in hours. Row 2 considers the counterfactual evolution of the gap assuming that the residual gap evolved according to the observed time series but the prices on hours are held constant at 1980 levels. Row 3 shows the percentage difference between Row 1 and Row 2—how different the gender wage gap would have been if returns to hours remained constant at 1980 levels.

First consider the top panel, where we price the hours gap at the rate implied by the return to actual hours, controlling for the standard covariates. We do not control for occupation fixed effects, instead pricing both hours within and across occupations at the single price indicated by this regression. We find that the total residual and hours related wage gap declined from 0.43 to 0.25 over the time period, with the bulk occurring between 1980 and 1990. Performing our counterfactual exercise, we show that gender wage gap would have been roughly 20 percent smaller in the year 2016 had the returns to hours worked not increased over the time period. Note that even though the return to hours worked is very small in 2016, it was *negative* in 1980. Thus even though hours does little to explain the residual wage gap in the contemporary cross-

section, the change in the return to actual hours still can still account for a portion of the change in the gender wage gap.

In the second panel, we base our analysis on specifications in which we control for occupation fixed effects. Doing so removes from the residual wage gap the contribution of hours associated with differences in expected hours across occupations. As a consequence, the level of the gender wage gap falls relative to the other specifications. However, we find that the reduction in the gender wage gap is comparable to the first case, with the counterfactual wage gap being roughly 24 percent lower than the actual wage gap.

We next consider the case in which the hours gap is priced at the rate implied by the cross-occupation wage-hours schedule.[24] As we did in the prior section, we implement this by instrumenting actual hours worked with the leave-out mean of hours in each occupation. We do not control for occupation fixed effects and instead price out all hours, both within and across occupation, at the rate estimated in our instrumental variables specification. In the third panel of Table 6, we show the results corresponding to this specification. With the large increase in the return to expected hours, the counterfactual wage gap would have been substantially lower had the hours premium remained at 1980 levels. By 2016, the gap would have been 46 percent lower.

For the bottom panel, we perform the same analysis but control for the occupational task mix. The results are qualitatively similar to the prior panel. Had the hours premium remained fixed at 1980 levels, the counterfactual gender wage gap would have been 34 percent lower in 2016. Thus, although the hours gap between men and women has been closing over this time

---

[24] Note that the residual wage gap plus the wage gap explained by the hours gap is nearly identical to what we saw in the first panel. The primary difference is the portion of the gap that is explained by the difference in hours.

period, it has not closed as rapidly as the return to hours has grown and thus this has exacerbated the gender wage gap from 1980 to 2016.

5. Conclusions

In this paper we consider the manner in which wages are related to hours worked. We demonstrate that while the hours worked by an individual has only a weak relationship with wages, the average hours within an occupation is much more strongly related to wages. This relationship holds, controlling for individual characteristics and occupation tasks. Indeed, a 10 percent increase in occupational hours worked is associated with between a 10 and 20 percent increase in wages depending on the specification.

We examine how this relationship has changed over time. We find that while the relationship between actual hours worked and wages has increased modestly, the relationship between occupational average hours and wages has doubled between 1980 and 2016. This substantial increase is observed whether or not we control for the occupational task mix. This is consistent with prior literature showing that the return to hours worked has risen over time.

We present a framework to rationalize this finding in which equilibrium compensation is a function of expected and actual hours. In the context of this framework, we provide clear settings in which expected hours would be priced at a much higher level than the deviations from expected hours. These findings are consistent with what we observe in the data.

Informed by these empirical findings and our conceptual framework, we revisit the gender wage gap. Because hours have such a weak relationship with wages, they explain very little of the gender wage gap. In the context of our framework, we demonstrate that the extent to

which the gender wage gap is mediated by hours worked depends crucially on whether we assume the hours difference is priced according to the empirical schedule for expected hours or rather hours deviations. If we assume that the gender difference in hours worked should be priced according to the relationship we observe across occupations, then hours worked accounts for virtually all of the unexplained gender wage gap. If we price hours according to the cross-occupation schedule and control for the occupational task mix, hours differences account for approximately half of the gender wage gap. These results imply that the findings of Cortes and Pan (Forthcoming) that increases in labor supply have important wage effects generalize to a broader sample of U.S. workers.

Given that return to hours worked has risen over time, we examine how the gender wage gap would have counterfactually evolved had the return to hours worked not risen relative to 1980. We do so under various assumptions regarding the pricing of hours worked. Consistent with Cha and Weeden (2014), the increase in returns to hours worked imply that the counterfactual wage gap would have been substantially narrower had the price of hours not increased over time. Significantly, if we price hours worked according to the occupation-level relationship, the counterfactual gender wage gaps would have been between 31 and 46 percent smaller. This suggests that the increase in the return to occupational intensity made it increasingly difficult for women to achieve wage parity with men.

Further research is needed to obtain more precise estimates of the pricing of hours differences between men and women at the individual level. We also leave for further research a greater understanding of why the returns to hours have increased over time and additional implications of this change.

References

Acemoglu, Daron, and David Autor. "Skills, tasks and technologies: Implications for employment and earnings." *Handbook of Labor Economics*. Vol. 4. Elsevier, 2011. 1043-1171.

Altonji, J.G. and Paxson, C.H., 1988. Labor supply preferences, hours constraints, and hours-wage trade-offs. Journal of labor economics, 6(2), pp.254-276.

Altonji, Joseph G., and Christina H. Paxson. "Labor supply, hours constraints, and job mobility." Journal of Human Resources 27, no. 2 (1992): 256-279.

Angrist, J. D. (1991). Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics*, 47(2-3), 243-266.Autor, David H., Larry F. Katz, and Melissa S. Kearney. "Trends in U.S. Wage Inequality: Revising the Revisionists." *The Review of Economics and Statistics* 90.2 (2008): 300-323.

Autor, David H., Frank Levy, and Richard J. Murnane. "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics* 118.4 (2003): 1279-1333.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics* 2, no. 3 (2010): 228-55.

Blau, Francine D., and Lawrence M. Kahn. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature* 55.3 (2017): 789-865.

Cha, Youngjoo and Kim A. Weeden (2014). "Overwork and the Slow Convergence in the Gender Gap in Wages." *American Sociological Review*. 79(3): 457-484.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *The quarterly journal of economics* 126, no. 2 (2011): 749-804.

Cortes, Patricia and Jessica Pan. Forthcoming. "When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled." Journal of Labor Economics

Cortes, Patricia and Jessica Pan. 2016b. "Prevalence of Long Hours and Skilled Women's Job Choices." IZA DP No. 10225.

Deming, David J. "The growing importance of social skills in the labor market." *The Quarterly Journal of Economics* 132.4 (2017): 1593-1640.

Dorn, David. "Essays on Inequality, Spatial Interaction, and the Demand for Skills." Dissertation University of St. Gallen no. 3613, September 2009.

Farber, H. S. (2015). Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. *The Quarterly Journal of Economics*, *130*(4), 1975-2026.

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, Claudia, and Lawrence F. Katz. "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation." *Journal of Labor Economics* 34.3 (2016): 705-746.

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

Hirsch, Barry T. (2005), "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." *Industrial and Labor Relations Review*, 58 (4), 525–51.

Kuhn, Peter and Fernando Lozano (2008). "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979-2006." *Journal of Labor Economics*. 26(2): 311-343.

Mas, A., & Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, *107*(12), 3722-59.

Oettinger, G. S. (1999). An empirical analysis of the daily labor supply of stadium vendors. *Journal of political Economy*, *107*(2), 360-392.

Wasserman, Melanie. 2015. "Hours Constraints, Occupational Choice and Fertility: Evidence from Medical Residents." MIT Working Paper.

Wiswall, M., & Zafar, B. (2017). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, *133*(1), 457-507.

Appendix A

Creating Task Measures

Task measures are constructed from the raw data in the ONET 4.0. We first standardize the raw level variables to be mean zero and standard deviation one across the 900 ONET-SOC occupation codes. We then collapse to 677 soc2000 codes by taking the simple average across ONET-SOC codes associated with a single soc2000 code.[25] The composites are created as the average of the included variables (see details below) and are standardized.

For analysis that uses data from 1980, 1990, 2000 and 2016 we merge on task data using David Dorn's *occ1990dd* classification system. The *occ1990dd* system consists of 330 codes that provide a balanced panel of occupations covering the 1980, 1990 and 2000 Censuses and the 2005 ACS. For the purpose of our analysis, we extend the coverage to the 2016 ACS by creating a crosswalk from the codes used in the 2016 ACS to the occ1990dd system. We start with the composite task measures at the soc2000 level and merge on soc2000 weights. We create soc2000 weights by pooling data from the 2005, 2006 and 2007 Occupational Employment Statistics (OES) survey.[26] We then collapse task measures to the occ2000 and standardize, yielding composite task measures for 445 occ2000 codes. Lastly, we use the occ2000 to occ1990dd from Dorn (2009) and the sum of soc2000 weights for each occ2000 code to collapse task measures to the occ1990dd level. The final dataset merged onto data for 1980, 1990, 2000 and 2016 consist of task data for 325 occupations standardized to be mean zero and standard deviation one. Thus, there are five *occ1990dd* occupations for which we are unable to obtain task data; these map into seven occupations in the 2016 ACS coding system.[27]

We use four composite task measures in our analysis taken previously from the literature. Each measure is constructed as the average of the included variables. For each composite the variable names are given in italics, the variable type in parenthesis and the variable question text in quotations.

1. Social Skills (Deming 2017):
   - *Coordination*: (skill) "Adjusting actions in relation to others' actions."
   - *Negotiation:* (skill) "Bringing others together and trying to reconcile differences."
   - *Persuasion*: (skill) "Persuading others to change their minds or behavior."
   - *Social Perceptiveness*: (skill) "Being aware of others' reactions and understanding why they react as they do."

2. Abstract Analytical (Acemoglu & Autor 2011):
   - *Interpreting the Meaning of Information for Others*: (work activity) "Translating or explaining what information means and how it can be used."

---

[25] For example, onetsoccode 11-1011.01 and 11-1011.02 are collapsed into soc2000 code 11-1011.

[26] Specifically, we follow the procedure used by Autor & Acemoglu (2011) to create soc2000 weights from the 2005, 2006 and 2007. Weights are calculated as the mean of employment across the three survey waves for each soc code.

[27] The occ1990dd occupations for which we cannot construct task data are occupations 76, 346, 37, 349 and 415.

- *Thinking Creatively*: (work activity) "Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions."
- *Analyzing Data or Information*: (work activity) "Identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts."

3. Manual (Acemoglu & Autor 2011):
   - *Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls*: (context) "How much does this job require using your hands to handle, control, or feel objects, tools or controls?"
   - *Manual Dexterity:* (ability) "The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects."
   - *Operating Vehicles, Mechanized Devices, or Equipment*: (work activity) "Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft."
   - *Spatial Orientation*: (ability) "The ability to know your location in relation to the environment or to know where other objects are in relation to you."
4. Routine (Acemoglu & Autor 2011)[28]:
   - *Controlling Machines and Processes*: (work activity) "Using either control mechanisms or direct physical activity to operate machines or processes (not including computers or vehicles)."
   - *Spend Time Making Repetitive Motions*: (context) "How much does this job require making repetitive motions?"
   - *Pace Determined by Speed of Equipment*: (context) "How important is it to this job that the pace is determined by the speed of equipment or machinery? (This does not refer to keeping busy at all times on this job.)"
   - *Importance of Being Exact or Accurate*: (context) "How important is being very exact or highly accurate in performing this job?"
   - *Importance of Repeating Same Tasks*: (context) "How important is repeating the same physical activities (e.g., key entry) or mental activities (e.g., checking entries in a ledger) over and over, without stopping, to performing this job?"

---

[28] The measure of routine used in Acemoglu & Autor 2011 also included the variable *Structured versus Unstructured Work* ("To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?") but the variable is unavailable in the ONET 4.0 and responses were not gathered until subsequent installations of the ONET.

Table 1 - OLS Regressions of Ln(Wage) on Ln(Hours), 2016 ACS Individual-Level

Log Wage	(1) all	(2) all	(3) all	(4) male	(5) female
Log Hours	0.244***	0.063**	-0.111***	-0.144***	-0.102***
	(0.046)	(0.026)	(0.002)	(0.004)	(0.003)
Demo. Controls	No	Yes	Yes	Yes	Yes
Occ FE	No	No	Yes	Yes	Yes
N	633,927	633,927	633,927	320,729	313,198

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Columns (1)-(2) use robust standard errors. Columns (3)-(5) include occupation fixed effects. Occupations aggregated to occs as classified in the 2016 American Community Survey. Observations are weighted using perwt. Sample includes prime-age workers aged 25-55. Demographic controls include a quartic in age, indicators for race/ethnicity (black, Hispanic, Asian, other), sex, and indicators for educational attainment (HS degree, some college, college degree or more). Hours are derived from usual hours worked per week. Hourly wage is calculated as total reported wage and salary income for the prior year divided by the product of usual hours worked and weeks worked in the previous year. We trim wages that are below half of the federal minimum wage and inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures.

Table 2 - OLS Regressions of Ln(Wage) on Ln(Hours), CPS Individual-Level

Log Wage	All workers					Full-time, full-year workers		
	(1) OLS	**(2)** OLS	**(3)** OLS	**(4)** IV: Lagged	**(5)** IV: Actual	**(6)** OLS	**(7)** OLS	**(8)** OLS
Log Hours	0.195***	0.011	-0.140***	0.153***	-0.000	0.329***	-0.100*	-0.304***
	(0.039)	(0.020)	(0.003)	(0.039)	(0.024)	(0.094)	(0.050)	(0.007)
Demo. Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Occ FE	No	No	Yes	Yes	Yes	No	No	Yes
N	317,544	317,544	317,544	65,815	309,502	267,091	267,091	267,091

All columns use standard errors clustered at the occupation level. Columns (3)-(5) and (8) include occupation fixed effects. Column (4) instruments usual hours worked with usual hours worked, reported in the previous March. Column (5) instruments usual hours worked with actual hours worked the week previous to the survey. Occupations aggregated to occs as classified in the 2012-2017 Current Population Survey, Annual Social and Economic Supplement. Observations are weighted using asecwt. Sample includes prime-age workers aged 25-55. Demographic controls include a quartic in age, indicators for race/ethnicity (black, Hispanic, Asian, other), sex, and indicators for educational attainment (HS degree, some college, college degree or more). Hours are derived from usual hours worked per week. Hourly wage is calculated as total reported wage and salary income for the prior year divided by the product of usual hours worked and weeks worked in the previous year. We trim wages that are below half of the federal minimum wage and inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures.

Table 3 - Regressions of Log(Wage) on Log(hours), CPS Individual-Level Data, Wage vs. Salary Workers

	Wage Earners					Salary Workers				
	(1)	(2)	(3)	(4) IV:	(5)	(6)	(7)	(8)	(9) IV:	(10)
	OLS	OLS	OLS	Lagged	IV: Actual	OLS	OLS	OLS	Lagged	IV: Actual
Log Hours	0.304***	0.272***	0.167***	0.228***	0.176***	0.063	-0.104***	-0.215***	0.019	-0.184***
	(0.035)	(0.024)	(0.002)	(0.022)	(0.014)	(0.040)	(0.025)	(0.004)	(0.046)	(0.025)
Demo. Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Occ FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
N	382,353	382,353	382,353	96,325	371,227	315,154	315,154	315,154	89,347	304,644

All columns use standard errors clustered at the occupation level. Columns (3)-(5) and (8)-(10) include occupation fixed effects. Columns (4) and (8) instruments usual hours worked with usual hours worked, reported in the previous March. Columns (5) and (10) instruments usual hours worked with actual hours worked the week previous to the survey. Occupations aggregated to occs as classified in the 2012-2017 Current Population Survey. Observations are weighted using earnwt. Sample includes prime-age workers aged 25-55. Demographic controls include a quartic in age, indicators for race/ethnicity (black, Hispanic, Asian, other), sex, and indicators for educational attainment (HS degree, some college, college degree or more). Hours are derived from usual hours worked per week. Hourly wage is calculated weekly wage and salary income divided usual hours worked. We inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures.

Table 4 - OLS Regressions of Ln(Wage) on Ln(Hours), 2016 ACS at the Occupational Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample = All workers				Sample = All workers			
	Average Log Wage				Average Log Wage			
					High-Skill Occ	Low-Skill Occ	Female-Dominated Occ	Male-Dominated Occ
Average Log Hours	2.352***		1.376***		2.239***	1.550***	2.595***	2.727***
	(0.243)		(0.194)		(0.350)	(0.145)	(0.269)	(0.430)
Average Residual Log Hours		2.027***		1.174***				
		(0.196)		(0.181)				
Task Controls	No	No	Yes	Yes	No	No	No	No
N	474	474	467	467	237	237	237	237

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Columns (1)-(8) use robust standard errors. Columns (1) and (3) regress occupation average of log wage on occupation average of log hours. Occupations are weighted by the occupation total of perwt. Columns (2) and (4) regress occupation average of residualized log wage on occupation average of residualized log hours. Residuals are constructed by regressing log wage (hours) on demographic controls and a full set of occupation fixed effects. We use the coefficients on the occupation fixed effects as measures of average residualized log wage and log hours. Demographic controls used in the residualization include black, hispanic, asian, other race, hs only, some college, ba plus and a quartic in age. Hours are derived from usual hours worked per week. Hourly wage is calculated as total reported wage and salary income for the prior year divided by the product of usual hours worked and weeks worked in the previous year. We trim wages that are below half of the federal minimum wage and inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures. High (low) skill occupations are defined as those above (below) the median in terms of fraction of workers in the occupation with at least a BA degree. Female (male) dominated occupations are defined as those above (below) the median in terms of fraction female in the occupation. Task data is constructed from the ONET 4.0 aggregated to the occ1990dd level. There are five occ1990dd occupations missing task data in the ONET 4.0 which yield missing task data for seven occs. Task controls include social skills, as defined in Deming(2017), and abstract analytical, manual and routine, as in Acemoglu & Autor (2011). See appendix A for an explanation how of these composites were created. Task measures are standardized to be mean zero and standard deviation one across the occ1990dd distribution.

Table 5 - Estimates of the Gender Wage Gap, 2016 ACS

Log Wage	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.248***	-0.240***	-0.157***	-0.167***	-0.011	-0.081***
	(0.022)	(0.021)	(0.007)	(0.007)	(0.023)	(0.021)
Log Hours		0.063**		-0.111***	1.786***	1.193***
		(0.026)		(0.020)	(0.162)	(0.174)
Demo. Controls	Yes	Yes	Yes	Yes	Yes	Yes
Occ. FE	No	No	Yes	Yes	N/A	N/A
Task Controls	No	No	No	No	No	Yes
N	633,927	633,927	633,927	633,927	633,927	628,344

Columns (3) and (4) include occupation fixed effects. Columns (5) and (6) are IV regression with individual log hours instrumented with occupation average log hours, thus occupation fixed effects cannot be included. Occupation average log hours are calculated using a leave-out mean by excluding the individual's hours. Standard errors are clustered at the occ level. Occupations aggregated to occs as classified in the 2016 American Community Survey. Observations are weighted using perwt. Sample includes prime-age workers aged 25-55. Demographic controls include a quartic in age, indicators for race/ethnicity (black, Hispanic, Asian, other), and indicators for educational attainment (HS degree, some college, college degree or more). Hours are derived from usual hours worked per week. Hourly wage is calculated as total reported wage and salary income for the prior year divided by the product of usual hours worked and weeks worked in the previous year. We trim wages that are below half of the federal minimum wage and inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures. We merge on task data constructed from the ONET 4.0 aggregated to the occ1990dd level. There are five occ1990dd occupations missing task data in the ONET 4.0 which yield missing task data for seven occs. Task controls include social skills, as defined in Deming(2017), and abstract analytical, manual and routine, as in Acemoglu & Autor (2011). See appendix A for an explanation how of these composites were created. Task measures are standardized to be mean zero and standard deviation one across the occ1990dd distribution.
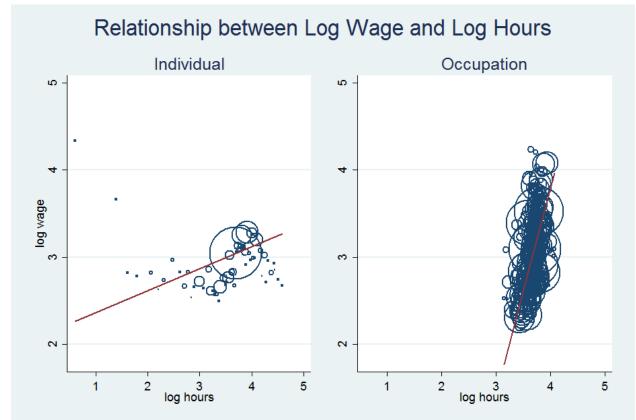
Table 6 – Evolution of the Gender Wage Gap

		1980	1990	2000	2016
Hours priced according to OLS regression of individual log wages on individual log hours
(1)	$\alpha^t + \beta^t gap_h^t$	0.433	0.321	0.260	0.248
(2)	$\alpha^t + \beta^{1980} gap_h^t$	0.433	0.276	0.216	0.200
(3)	% Diff (2) & (3)		14.3%	16.7%	19.2%
Hours priced according to OLS regression of individual log wages on individual log hours with occupation fixed effects
(1)	$\alpha^t + \beta^t gap_h^t$	0.318	0.239	0.174	0.152
(2)	$\alpha^t + \beta^{1980} gap_h^t$	0.318	0.196	0.138	0.116
(3)	% Diff (2) & (3)		17.9%	20.8%	23.8%
Hours priced according to IV regression of individual log wages on individual log hours using occupation hours as instrument
(1)	$\alpha^t + \beta^t gap_h^t$	0.434	0.321	0.257	0.238
(2)	$\alpha^t + \beta^{1980} gap_h^t$	0.434	0.259	0.170	0.130
(3)	% Diff (2) & (3)		19.4%	33.9%	45.5%
Hours priced according to IV regression of individual log wages on individual log hours using occupation hours as instrument, controlling for occupation tasks
(1)	$\alpha^t + \beta^t gap_h^t$	0.435	0.325	0.256	0.232
(2)	$\alpha^t + \beta^{1980} gap_h^t$	0.435	0.268	0.189	0.153
(3)	% Diff (2) & (3)		17.6%	26.0%	34.3%

Occupations aggregated to occs as classified in the 2016 American Community Survey. Observations are weighted using perwt. Sample includes prime-age workers aged 25-55. The gap in hours (gap_h^t) corresponds to the difference between male and female log hours in year t. In the first panel, β^{1980} is the coefficient on log hours fro a regression of log wage on a female dummy and a set of demographic controls. The second panel adds in occupation controls to derive the coefficient on log hours, β^{1980}. In the third panel β^{1980} is the coefficient from log hours in a IV regression of log wage on female, demographic controls and individual log hours instrumented with occupation average log hours. Occupation average log hours are calculated using a leave-out mean by excluding the individual's hours. In the fourth panel add occupation tasks controls to derive the coefficient on log hours, β^{1980}. In each panel α^t is the coefficient on female and β^t is the coefficient on log hours from the corresponding regression in year t. All regressions include a quadratic in age, dummies for black, Hispanic, Asian and other race, high school only, some college and at least a bachelor's degree. Hours are derived from usual hours worked per week. Hourly wage is calculated as total reported wage and salary income for the prior year divided by the product of usual hours worked and weeks worked in the previous year. We trim wages that are below half of the federal minimum wage and inflate wages to 2012 dollars using the Bureau of Economic Analysis (BEA) National Income and Product Accounts Personal Consumption Expenditures. Task data are constructed from the ONET 4.0 aggregated to the occ1990dd level. There are five occ1990dd occupations missing task data in the ONET 4.0 which yield missing task data for seven occs. Social skills are as defined in Deming(2017), abstract analytical, manual and routine as in Acemoglu & Autor (2011). See appendix A for an explanation how of these composites were created. Task measures are standardized to be mean zero and standard deviation one across the occ1990dd distribution.

Figure 1

Relationship between Log Wage and Log Hours

Note: Occupations collapsed to 3 digit occ level. Weights are sum of perwt in each occupation.
Indiividuals split into 1000 bins on lnhours. Weights are sum of perwt in each bin.

Figure 2

Note: Presented are point estimates and 95% confidence intervals. Individual regression includes controls for gender, education, race and age and occupation dummies. Occupation level regresses average residualized log wage on average residualized log hours. Log wages and log hours are residualized on controls for education, race, age and female and occupation dummies. The residuals are the predicted occupation dummies. Occupations aggregated to occ1990dd level.

Figure 3

Occupation Relationship between Log Wage and Log Hours

Note: Presented are point estimates and 95%. Occupations aggregated to occ1990dd level. Occupation average residualized log wage is regressed on occupation average residualized log hours. Log wages and log hours are residualized on controls for education, race, age and female and occupation dummies. The residuals are the predicted occupation dummies. Social, manual routine, and abstract analytical tasks are included. Tasks are time invariant and come from the ONET 4.0.

Figure 4

Note: Plotted is the log difference between female and male average log hours and average log wages.
Occupation estimates are occupation average residualized log wage regressed on occupation average residualized
log hours. Log wages and log hours are residualized on controls for education, race, age and female
and occupation dummies. The residuals are the predicted occupation dummies. Occupations
aggregated to occ1990dd level.