IZA

IZA DP No. 9846

When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled

Patricia Cortés
Jessica Pan

March 2016

When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled

Patricia Cortés
Questrom School of Business, Boston University

Jessica Pan
*National University of Singapore
and IZA*

ABSTRACT

When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled[*]

This paper explores the relationship between gender differences in hours worked, the returns to working long hours, and the gender pay gap among highly educated workers. Using a cross-section of occupations, Goldin (2014) documents that occupations characterized by high returns to overwork are also those with the largest gender gap in earnings. To provide a causal link between the demand for long hours and how it relates to gender wage gaps, we exploit supply side shocks – generated by intercity variation in low-skilled immigrant flows – to examine whether reductions in the cost of supplying longer hours of work allow women to close the gap in hours of work and benefit from higher wages. We find that low-skilled immigration leads to a reduction in a city's gender gap in overwork, as well as in the gender pay gap in occupations that disproportionately reward longer hours of work.

Corresponding author:

Jessica Pan
Department of Economics
National University of Singapore
1 Arts Link
Singapore 117570
Singapore
E-mail: jesspan@nus.edu.sg

1   Introduction

Over the past five decades, women have made substantial labor market gains in labor force participation, earnings, and representation in professional occupations. Accompanying these changes has been a striking reversal in the gender education gap. Yet, despite the converging roles of men and women, gender pay differentials have remained remarkably persistent, even among the highly-skilled. While women's relative earnings converged rapidly beginning in the late 1970s through the 1980s, the convergence slowed in the 1990s, and appears to have stalled since the 2000s (Blau 2012, Blau and Kahn, 2006). Moreover, the rates of convergence of the gender pay gap has been quite different across the education and skill distribution. As shown in Figure 1A, in 1980, the gender pay gap for college women was between 5 to 10 percentage points smaller than that of other education groups – by 2010, the opposite appears to be true. The size of the gender pay gap among those with some college education or just a high school degree is about 5 percentage points smaller relative to those with a college degree or more. Similarly, using PSID micro data from 1980 to 2010, Blau and Kahn (2016) document that the gender wage gap is currently largest at the top of the wage distribution, and has decreased more slowly at the top relative to other points in the distribution.[1] These patterns are all the more remarkable as college-educated women today are characterized by high labor force attachment and are increasingly well-represented in many professional spheres.[2]

The existence of a "glass ceiling" even among groups of women who are arguably as skilled and well-trained as men highlights the need to look within occupations to understand how jobs are organized and compensated. In her 2014 AEA presidential address, Claudia Goldin points to the demand for temporal flexibility as a major cause of the gender wage gap. Specifically, she argues that some jobs disproportionately reward individuals who are willing to work long (and particular) hours, and that these same occupations tend to impose the largest penalties for workforce interruptions (Goldin 2014). As women tend to place a higher value on temporal flexibility given their dual roles in the home and in the labor market, such rigid occupational demands are likely to work to the disadvantage of women (Gicheva 2013, Wasserman 2015). Using a cross-section of occupations from the 2010 ACS, Goldin (2014) provides some empirical support of this view – occupations that are characterized by high returns to long hours, and greater temporal inflexibility, are also those with the largest gender earnings gaps.

Aggregate time-series patterns also provide some suggestive evidence that the increasing returns to providing long hours of work, coupled with relatively stable gender differences in hours worked, might account for the persistence of gender pay gaps, particularly among highly educated workers.

---

[1] Arulampalam, Booth and Bryan (2007) document an acceleration in the gender wage gap at the top of the wage distribution in Europe from 1995-2001.

[2] For example, as documented by Goldin (2014), there is near equal representation of men and women in law and medical fields.

Figure 1B depicts trends in the elasticity of annual earnings with respect to weekly hours worked (a measure of the returns to working long hours) from 1980 to 2010 for males of different education levels. As observed in the figure, the premium for working longer hours has increased consistently for all education groups, with college-educated workers experiencing the largest increase over time. Given that the gender gap in hours worked has remained relatively constant over time for all education groups (Figure 1C), these patterns suggest a potential role for changes in the returns to working long hours in explaining the relatively slower convergence of the gender pay gap among the highly educated. In a similar vein, Cha and Weeden (2014) document that rising returns to overwork, coupled with the gender gap in the propensity to work overtime, worked to slow the convergence of the gender wage gap during the 1979-2009 period.

While the cross-occupation and time-series evidence of the relationship between the returns to working long hours and the gender pay gap are highly suggestive, they do not address the issue that occupations that disproportionately reward individuals who work long hours are likely to differ on other important dimensions that may have an independent effect on the gender pay gap. For example, occupations where the incidence of overtime is common, such as financial managers and lawyers, are also characterized as being highly competitive. Recent research suggests that males tend to outperform females in competitive settings and women are more likely to "opt-out" of competition (Gneezy, Niederle and Rustichini, 2003; Flory, Leibbrant and List, 2014). Furthermore, some of the proposed causes of the occupational differences in the returns to overwork, as well as its secular increase of time, such as such as globalization, changes in the organizational structures within firms, and technological advancements, might have also affected the returns to other worker attributes and job characteristics that are correlated with working long hours on the job.[3] Simply stated, the observed relationship between the returns to working long hours and the gender wage gap may be confounded by other job characteristics and worker sorting.

The main contribution of this paper is to provide a causal link between the demand for long work hours and the gender wage gap among highly educated workers. In the absence of an obvious source of exogenous variation in the returns to working long hours, our approach to establishing causality focuses on the supply side. We test if the higher cost to women of engaging in longer hours of market work is an important factor in explaining the persistent pay gap in skilled occupations, particularly in occupations that disproportionately reward individuals who are willing to supply longer hours. To identify this effect, we use a triple difference strategy that compares changes in the gender wage gap over time, in occupations that vary in terms of their demand for long hours of work, across cities where women face different costs of supplying longer hours of work. Following Cortes (2008) and Cortes and Tessada (2011), we use plausibly exogenous variation in low-skilled immigrant flows across US cities, and over time, to proxy for changes in the costs of outsourcing

---

[3]See Kuhn and Lozano (2008) for a discussion of some of the potential factors that may explain the rising returns to overtime.

household production. The exogeneity of the immigration flows is based on using the historical distribution of immigrants to allocate future flows of low-skilled immigrants at the national level. The intuition behind our empirical strategy is the following – cities that receive a large influx of low-skilled immigrants have greater availability of market substitutes for household production, thus enabling highly-skilled women in these cities to increase their market work (Cortes and Tessada, 2011). This reduction in the costs of outsourcing household production is likely to have the largest effects on women who work in occupations that demand more temporal flexibility and reward longer hours of work.

This empirical approach also allows us to address issues relating to whether our measure of the occupation-specific returns to working long hours – based on the observed cross-sectional relationship between earnings and weekly hours of work for workers in each occupation – are indeed capturing the true wage returns of additional hours of work, or unobservable worker characteristics that might be correlated with longer work hours and independently affect earnings. For example, if more able individuals are also the ones who tend to work longer hours, the observed returns to working long hours would capture this element of worker sorting. Moreover, if worker sorting occurs on the basis of worker attributes that differ by gender, this could lead to a spurious correlation between the observed returns to overwork and the gender pay gap. By focusing on supply side shocks, our empirical strategy isolates variation stemming from changes in the costs of providing long hours of work, providing us with an estimate of the causal effect of reducing the gender gap in work hours on women's relative earnings.[4]

We limit our focus to highly educated workers in skilled occupations for a number of reasons. First, as discussed above, high-skill workers have experienced the highest increase in the returns to working long hours and have the most persistent wage gender gaps. It is likely that occupation-specific factors such as the demand for temporal flexibility and the degree of substitutability of workers matter more for highly educated workers. Furthermore, unskilled occupations are likely to be subject to overtime laws, which imply that the sources of the returns to working longer hours for these groups are different from skilled occupations and may be determined by institutional forces. A final reason is that by focusing on highly educated workers in skilled occupations, we are able to focus on a relatively homogenous sample of workers and occupations, which would help limit the potential set of confounding factors when making comparisons across workers and occupations.

Using data from the 1980 to 2000 US Census and 2011 3-year aggregate American Community Survey (ACS), we begin by documenting that the simple cross-occupational correlations between the returns to working long hours and the gender pay gap as documented by Goldin (2014) hold for a sample of skilled occupations in each of the four decades from 1980 to 2010 and are robust to the

---

[4]Put differently, if the observed returns to working long hours largely captured cross-occupation differences in worker sorting, an exogenous increase in women's willingness to work longer hours should have little or no impact on the gender wage gap.

inclusion of occupation-specific controls. Next, we illustrate the challenge in interpreting these cross-occupation correlations as causal. In particular, we use data from O*NET to construct a measure of the competitiveness of an occupation, and show that this measure is strongly correlated with both the returns to working long hours and the gender pay gap in that occupation. Moreover, controlling for differences across occupations in this measure of competitiveness reduces the magnitude of the observed correlation between the returns to working long hours and the gender pay gap. This exercise suggests that unobserved differences across occupation are likely to bias the observed cross-sectional relationship between the returns to working long hours and the gender pay gap.

In order to implement the triple difference strategy, we start by establishing that women's labor supply and the gender gap in overwork do indeed respond to low-skill immigration. Extending the analysis in Cortes and Tessada (2011) to our present sample and outcomes of interest, we find that highly skilled women in cities that experienced larger inflows of low-skilled immigrants increased their hours worked and the likelihood of working overtime. The magnitude of the estimates imply that an increase in the predicted low-skilled immigration flow from 1980 to 2010 led to a decline in the gender gap in the probability of working 50 or more hours a week of about 1.3 percentage points, or 10%.

Turning to the triple difference estimates, our variable of interest is the interaction between an occupation-specific measure of the returns to working long hours (measured in 1980 at the national level) and the gender gap in the likelihood of working long hours, which we instrument using the predicted flow of low skilled immigrants. Our hypothesis is that if the higher cost of working long hours for women is harming their potential earnings, then an exogenous shock that reduces this cost should lead to a reduction in the gender gap in hours worked, and consequently, reduce the gender pay gap, particularly in occupations that place the highest demands on long hours of work. Our IV estimates imply that a one standard deviation decrease in the gender gap in the likelihood of working 50 or more hours a week of a city reduces the gender gap by about a half of a standard deviation more in occupations with the highest returns (top tercile), relative to occupations with the lowest returns (bottom tercile). This decline is approximately 30% of the average gender gap across occupations in 1980 and 50% of the cross-occupation standard deviation of gender pay gaps. The large magnitude of the effects suggests that gender differences in supplying long hours of work is an important cause of the persistence of gender gaps in some highly skilled occupations and that the returns to working long hours are mostly capturing real effects, and not just workers sorting or the returns to other skills. We show that these results are robust to the inclusion of a flexible set of fixed effects that capture unobserved shocks at the city*year, occupation*year, and city*occupation level, alternative measures and functional forms for the returns to working long hours, as well as to dropping occupations with the highest and lowest returns.

Besides contributing to the literature on gender gaps, this paper also adds to the literature on

the effects of migration on the receiving country. Most of the research on immigration flows have focused on the effects on the labor outcomes of natives via changes in the relative supplies of skilled vs. unskilled workers and the substitutability or complementarity of native-born versus foreign-born workers in the production function. Cortes and Tessada (2011) examined other potential channels, and provided evidence that immigration from low skilled countries by lowering the prices of services that are close to household production, have enabled highly skilled women to work more hours in the market. This paper extends their findings, by showing that low-skilled immigrants not only increase the probability that highly skilled women work long hours, but also leads to a reduction in the gender pay gap in the upper tail of the skill distribution, thereby indirectly contributing to raising the glass ceiling.[5] Our reduced form estimates suggest that the increase in the predicted low-skilled immigration flow from 1980 to 2010 led to a decline in the gender pay gap in occupations at the top tercile of the returns to working long hours – relative to the bottom tercile – of between 2 and 3 percentage points.

The rest of the paper is organized as follows. Section 2 describes the data, the construction of key variables, and presents the cross-occupation correlations. Section 3 discusses the empirical strategy, presents the main empirical specifications, and discusses the results. Section 4 concludes.

2   Data and Descriptive Statistics

The data is drawn from the 1980 to 2000 Censuses and 2011 American Community Survey (ACS) 3-year aggregate (2009-2011).[6] The main sample that we use to estimate the gender wage gaps and the returns to working long hours at the occupation level is restricted to native-born individuals age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week.[7]

To ensure that we have a consistent set of occupations over the sample time period, we use Dorn's (2009) occupational classification which modifies the *OCC1990* Census classification to create a consistent set of occupations from 1980 to 2010.[8] This consistent occupational coding scheme creates a panel of occupations from 1980 to 2011 and ensures that our results are not affected by changes in the set of occupations over time. Finally, our main empirical analyses focus on skilled occupations. We defined skilled occupations as those satisfying at least two of the following three

---

[5]Note that, theoretically, the effect could have gone in the opposite direction if males and females are not perfect substitutes in production, and the increase in the labor supply of highly skilled women lowered their wages.

[6]In the text, we refer to the data from the 2011 ACS as corresponding to the 2010 time period.

[7]We focus on full-time workers as selection into working part-time, particularly for men, is likely to be very strong. These unobserved factors that drive certain individuals to choose to work part-time are likely to distort our estimates of gender gaps and the return to working long hours.

[8]As Dorn's (2009) crosswalk only provides a consistent classification scheme for occupations until 2009, we extended the crosswalk to include the 2011 ACS occupation classification.

conditions: (1) classified by the Census as Managerial and Professional Specialty Occupations (codes 3-200), (2) share of college educated workers in 2010 is higher than that in the working population (35%), and (3) median full-time male annual income in 2010 is greater than the median full-time male annual income across occupations ($52,000)[9,10] These occupations account for approximately 72% of female and 74% of male college graduates employed in the US labor market. The full list of the 95 skilled occupations included in our sample is presented in Appendix Table 1. For the IV analysis, we further classify these 95 occupations into 16 broad occupational categories based on the classification scheme available in the Census.

Occupation-specific Estimates of the Returns to Working Long Hours

To estimate the returns to working long hours in each occupation o in decade t, we follow the procedure outlined in Goldin (2014). Specifically, we restrict the sample to male workers and estimate the following regression separately for each decade:

$$\ln(yearly\_earnings)_{io} = \alpha + \sum_o \beta_o * I(occ_o = 1) * \ln(hours\_week)_{io}$$
$$+ \eta * \ln(weeks\_year)_{io} + \pi_o + X'_{io}\delta + \varepsilon_{io} \quad (1)$$

where $yearly\_earnings_{io}$ is the annual wage and salary income of individual i in occupation o, $hours\_week$ refers to the usual hours worked per week, and $weeks\_year$ is the number of weeks an individual worked in the previous year.[11,12] π_o includes occupation fixed effects and X_i is a vector of demographic characteristics that includes a quartic in age, race fixed effects and indicators for whether an individual has a masters or doctoral degree. β_o provides a measure of the returns to working long hours, and indicates the elasticity of annual earnings with respect to weekly hours. Specifically, $\beta_o > 1$ implies that annual earnings increase more than proportionally for a given change in weekly hours worked, suggesting a convex relationship between earnings and working long hours. Conversely, $\beta_o < 1$ implies that a given increase in hours worked is associated with a less than proportional change in annual earnings. Therefore, occupations with higher β_o's are characterized by higher returns to working longer hours. For the main analysis, we use estimates of the returns estimated using only full-time male workers to avoid the complex selection issues that are likely to affect the annual wages and hours worked of female workers and workers who choose to work part-time. Nonetheless, in some of the analyses that follow, we will also discuss

---

[9]We decided on this selection rule because it combines three reasonable criteria for determining the skill level of an occupation (education, earnings, and nature of the work). The main results in the paper are robust to restricting the sample of skilled occupation to those satisfying each condition separately.

[10]We drop occupations with fewer than 30 males and 30 females in the each Census/ACS year.

[11]Weeks worked in the previous year is only available in intervals in the ACS. For each interval, we assign the mode of the interval as measured in the 2000 Census. For example, for the interval 50-52 weeks, most people report working 52 weeks in the 2000 Census.

[12]We drop observations in cases where based on the information proved on annual salary, hours per week and weeks per year individuals have implied hourly wages (in 1990$) below 3.5 or above 150.

the magnitudes of the returns that are estimated including full-time female workers, and evaluate the sensitivity of our results to using different samples. Appendix Table 2 reports the estimated returns to working long hours for each of the sixteen broad occupational groups from 1980 to 2010 for the sample of males, males and females, and using the alternative specification which we will discuss in equation (3). The table also indicates the corresponding terciles.

The interpretation of the OLS estimate of β_o from equation (1) as the true return to working longer hours in an occupation is subject to several important caveats. First, there is an issue of measurement. Our procedure measures the contemporaneous returns among individuals who choose to work different numbers of hours each week. For some occupations, one might reasonably expect that contemporaneous earnings might underestimate the long-run returns of working longer hours. For example, in business and finance, workers may be expected to work long hours at lower wages at the beginning of their career in order to advance to management positions that have significantly higher wages. For such workers in these occupations, β_o is likely to be an underestimate of the true (longer-run) returns.[13]

We investigate this issue empirically by conducting additional analysis using panel data from the Survey of Income and Program Participation (SIPP). As the SIPP follows individuals over time for up to four years, we can examine whether working longer hours is predictive of wages in the longer-run (3 to 4 years later). To the extent that some of the returns to working long hours are only accrued in the future, we would expect to see a positive correlation between hours worked today and future wages, even after controlling for future hours. In particular, we would like to assess whether occupations with high contemporaneous returns to working long hours are also those with high future returns.

Using data from the first wave and the last wave of the 1996, 2004, and 2008 SIPP panels,[14] we estimate (a) the elasticity of annual income in wave 1 with respect to current weekly hours worked in wave 1 and (b) the elasticity of annual income 3 to 4 years later in wave 12 with respect to weekly hours worked in wave 1, conditional on the hours worked in wave 12. The regression specification for (a) is similar to that in equation (1), with additional controls for the calendar year that each respondent was surveyed. For (b), we estimate the following modification of equation (1):

---

[13]Gicheva (2013) shows that among a sample of GMAT takers, working more hours, conditional on having worked at least 47 hours, is associated with a significant increase in annual wage growth as well as the likelihood of promotion. Interestingly, she does not find a similar relationship among employees working fewer than 47 hours.

[14]We use these panels as they had information available for at least 12 waves.

$$\ln(yearly\_earnings)_{io,w12} = \alpha \quad + \quad \sum_o \beta_o * I(occ_{o,w1} = 1) * \ln(hours\_week)_{io,w1} + \eta \ln(weeks\_year)_{io,w1}$$

$$+ \quad \sum_o \gamma_o * I(occ_{o,w1} = 1) * \ln(hours\_week)_{io,w12} + \theta \ln(weeks\_year)_{io,w12}$$

$$+ \quad \pi_o + X'_{io}\delta + \varepsilon_{io} \tag{1'}$$

where the subscripts $w1$ and $w12$ refer to wave 1 and wave 12, respectively. The estimates of β_o from equation (1) and equation (1$'$) using the SIPP are reported in Columns (1) and (2) of Appendix Table 3, respectively.[15] With some exceptions, the estimates in Column (2) are generally positive, indicating that longer hours worked in wave 1 are generally associated with higher income three to four years later in wave 12, even after controlling for weekly hours worked in wave 12. The cross-occupation rank correlation between the contemporaneous returns reported in Column (1) and the estimates of the additional returns to working long hours on future earnings in Columns (2) is 0.69 (significant at the 1% level).[16] These results suggest that measures of the contemporaneous returns are likely to understate the true returns to working long hours, particularly in occupations with high contemporaneous returns.[17]

Furthermore, given our focus on the labor market for skilled workers, top-coding of income in the Census is likely to affect a relatively large share of our observations and introduce measurement error. To address this issue, we follow the literature and multiply the income top-code for the 1980 Census by 1.5. We do not modify the income variables from 1990 to 2010 as the wages in the top-code are assigned the state median and mean of values above the top-code, respectively in the 1990 and 2000 Census. In the 2011 ACS, the top-code is assigned the mean earnings of individuals above the 99.5th percentile of income within each state. Nevertheless, while the Census/ACS top coding procedure from 1990 to 2010 ensures that the average income among individuals earning the top-code is accurate at the state level, at the occupation level, we are likely to underestimate the returns for occupations with a large share of workers with incomes at the very top of the earnings distribution. Finally, measurement error in reported weekly hours worked is also likely to lead to

---

[15] Note that caution has to be exercised when comparing the estimated returns in Column (1) directly to the returns reported in Appendix Table 4 since these were computed using a different dataset and are estimated using much fewer observations.

[16] Since individuals may not be in the same occupations in wave 1 and wave 12, we also estimate a version of equation (1$'$) where we include additional controls for the hours worked in wave 12 interacted with wave 12 occupation dummies – this exercise yielded similar results. The results are also robust to restricting the sample to individuals who were in the same occupation in wave 1 and wave 12.

[17] Given that occupations that tend to have higher contemporaneous are also those with higher returns in the future, for our analysis, in addition to entering the contemporaneous returns linearly in our models, we will also consider specifications that focus on differences across three different groups of occupations (low, medium, and high occupations), classified based on their contemporaneous returns. By focusing on occupations in separate terciles, we are able to capture both the effects of contemporaneous as well as longer-run returns to some extent.

a downward bias in the estimated elasticities. Overall, these limitations inherent in our measure imply that $\hat{\beta}_o$ is likely to underestimate the true returns to working long hours.

The second issue is that of causality. In general, the OLS estimate $(\hat{\beta}_o)$ provides a causal estimate of the returns to working longer hours only if an individual's willingness to work long hours is exogenous to his/her other skills. If an individual's leisure preferences and other skills are correlated, then the measure of the returns to working long hours would not capture a causal relationship. For example, in Gicheva's (2013) promotion model, learning-by doing depends on the worker's ability level – in other words, more able individuals are the ones who benefit the most from working more hours. Therefore, in occupations where working longer hours and worker ability are strategic complements, more able workers would tend to sort into higher levels of hours, and our estimate of the returns $(\hat{\beta}_o)$ would be biased upwards. The willingness to work long hours might also be correlated with other non-cognitive skills – for example, in Landers et al.'s (1996) study of lawyers, they found that billable hours were used as a signal for ambition for success and the willingness to pursue the interests of clients aggressively. In the same vein, Jeff Bezos, CEO of Amazon, reportedly wrote in a letter to shareholders "When I interview people I tell them, 'You can work long, hard, or smart, but at Amazon.com you can't choose two out of three.'" (Kantor and Streitfeld, 2015). Our empirical strategy will allow us to determine if the relationship between working long hours and earnings is at least partially causal.[18]

Occupation-specific Estimates of the Gender Pay Gap

To estimate the gender pay gap in each occupation, we estimate the following specification for each Census/ACS year for our main sample:

$$\ln(yearly\_earnings)_{io} = \alpha \;\; + \;\; \sum_o \lambda_o * I(occ_o = 1) * female_{io} + \beta * \ln(hours\_week)_{io}$$
$$+ \;\; \eta * \ln(weeks\_year)_{io} + \pi_o + X'_{io}\delta + \varepsilon_{io} \tag{2}$$

The controls used in this equation are identical to that in equation (1). The coefficient, λ_o, is our estimate of the residual gender earnings gap in occupation o. Note that in this specification, we are not allowing β to vary by occupation.

Alternative Measures of the Returns to Long Hours and the Gender Pay Gap

For the main analysis of the paper we follow Goldin (2014)'s functional forms for the estimation of our key variables. However, we also construct measures using an alternative functional form and present results using them as robustness tests. In particular, we construct a dummy variable for

---

[18]To our knowledge, this is the first paper to tackle this issue; neither Goldin (2014) nor Cha and Weeden (2014) address the issue of causality.

overwork (working 50 hours a week or more) and estimate the following models:[19]

$$\ln(hourly\_wage)_{io} = \alpha + \sum_o \theta_o * I(occ_o = 1) * Dummy\_overwork_{io}$$
$$+\pi_o + X'_{io}\delta + \varepsilon_{io} \tag{3}$$

and

$$\ln(hourly\_wage)_{io} = \alpha + \sum_o \lambda_o * I(occ_o = 1) * female_{io} +$$
$$+\pi_o + X'_{io}\delta + \varepsilon_{io} \tag{4}$$

In this case, we interpret a positive θ_o as indicating a premium for working long hours, and a negative θ_o as a penalty. Appendix Figure 1 shows a very strong correlation between the two measures of the returns to working long hours.

2.1 Correlations across Occupations

To highlight the potential relationship between the returns to working long hours and gender pay gaps, we begin by presenting cross-occupation correlations separately by decade. As shown in Figure 2, there is significant variation in the premium to working long hours and the size of the gender pay gap across occupations. For example, in 2010, the elasticity of annual earnings with respect to weekly hours worked was lower than 0.3 for occupations such as teachers, dentists, physicians, and veterinarians, but higher than 1.2 for lawyers, financial managers, actuaries, accountants, and other financial specialists. The average return increased substantially from 0.43 in 1980 to 0.7 in 2010. The residual gender gap in earnings also varies considerably across occupations – the earnings gap is less than five percent for many scientific, engineering, and teaching occupations in 2010, but is larger than 25 percent for financial managers, physicians, dentists, and occupations in insurance and financial services.

Consistent with the evidence presented by Goldin (2014), for most years, there is a statistically significant negative correlation between the returns to working long hours and the female-male earnings gap. Occupations that reward long hours of work are also those with higher gender pay gaps. Table 1 presents the regression version of Figure 2. Since the precision with which we measure the gender earnings gap and the returns to working long hours depends on the numbers of observations from which they are calculated, we weight all the regressions with the number of observations in each cell.

---

[19]We follow Kuhn and Lozano (2008) and Cha and Weeden (2014) in choosing 50 hours per week as the threshold for "overwork".

Columns (1), (3), (5), and (7) report the coefficient estimates of the univariate relationship for each decade. The correlations are negative, large in magnitude, and statistically significant from 1990 to 2010. Note, however, that the estimated correlations might be biased if there is positive selection of females that choose or decide to stay in occupations where the returns are high (Olivetti and Petrongolo, 2008), or might be driven by confounders such as wage levels and overwork prevalence.[20] Columns (2), (4), (6), and (8) show that the observed cross-occupation correlations keep their statistical significance but the magnitudes become somewhat smaller when occupation-level controls such as the female share of full-time workers in the occupation, the labor force participation rates of women who report being in that occupation,[21] the log of the average wage of males working exactly 40 hours, and the share of males working 50 or more hours a week are included.[22] Similar results are obtained when the alternative measures of the returns to long hours and the gender pay gap are used (see Appendix Table 5).

There are two potential issues in interpreting these cross-occupation correlations as the causal effect of the returns to working long hours on the gender earnings gap. The first stems from an issue that we discussed earlier – which is that our measure of the returns to working long hours, $\hat{\beta}_o$, may in fact capture the returns to other worker attributes that are correlated with the propensity to work longer hours. To the extent that there are gender differences in these worker attributes (e.g. bargaining skills, ambition, leadership, competitiveness), this might result in a spurious correlation between the returns to working long hours and gender pay gaps. The second, related, issue is that occupations vary on many dimensions, and the observed correlation between the long hours premium and gender pay gaps might be driven by other (unobserved) characteristics of occupations that are correlated with the factors that drive the returns to working long hours.

A specific example is the observation that occupations that tend to have the highest returns to working long hours (e.g. lawyers and business professionals) are also typically characterized as occupations that are highly competitive. The observed correlation may therefore capture both the effect of the returns to working longer hours as well as the returns to competitiveness. To the extent that women tend to shy away from competition and perform relatively poorer relative to men in more competitive settings (e.g. Gneezy, Niederle, and Rustichini, 2003, Ors, Palomino and Peyrache, 2012, Flory, Leibbrant and List, 2014), this suggests that the observed correlation between the returns to long hours and the gender pay gap might partially reflect differences in other characteristics of occupations such as the degree of competition, and not just intrinsic properties

---

[20] Kuhn and Lozano (2008) find an inverse relationship between changes in wage levels and changes in the residual inequality in occupation. They also find that residual inequality is positively correlated with the prevalence of long hours in an occupation.

[21] Note that we could only assign occupations to individuals currently in the labor force or who left the labor force within the last 5 years.

[22] Appendix Table 4 presents the cross-occupation correlations between the returns to long hours and female labor supply outcomes. Although most of the coefficients are negative, only the effect of the returns to working long hours on the female share in an occupation in 2000 is marginally significant.

of the occupation such as the imperfect substitutability of workers (Goldin, 2014). In other words, if the observed relationship between the returns to working long hours and the gender pay gap largely reflects differences across occupations in the value placed on worker attributes such as competitiveness (apart from number of hours worked), then it is unlikely that closing the gender gap in work hours, or reducing the convexity of the pay structure in occupations, will have a significant impact on the gender pay gap.[23]

To examine this possibility more concretely, we use the data from O*NET online to construct a measure of competitiveness in an occupation.[24] To construct a measure of how competitive an occupation is, we use answers to the following question in O*NET: "How competitive is your current job?" Respondents provide answers on a 1 to 5 scale (1: not competitive at all, 5: extremely competitive). We use the average reported competitiveness in each occupation, standardized to have a mean of zero and a standard deviation of one in the full sample of occupations.[25] Panels A and B of Figure 3 graphically depict the cross-occupation relationship in 2010 between the competitiveness measure from O*NET and the gender earnings gap, and the returns to working long hours, respectively.[26] As observed in Figure 3, we find that more competitive occupations tend to have larger gender pay gaps as well as higher returns to working long hours.

Table 2 presents the regression version of these correlations. Similar to Table 1, all the regressions are weighted by cell size. Column (1) confirms the positive relationship between an occupation's degree of competitiveness and its returns to working long hours (Panel A), as well as the size of its gender pay gap (Panel B). Column (2) shows that these correlations are robust to including controls for the same set of occupational characteristics as that in Table 1. In Columns (3) and (4), we re-estimate the regression of the gender pay gap on the competitiveness measure, including the returns to working long hours as an additional explanatory variable. Relative to the univariate correlation reported in Columns (7) and (8) of Table 1, the inclusion of the competitiveness measure reduces the size (between 23 and 60 percent) and statistical significance of the observed relationship between the returns to working long hours and the gender pay gap.[27] Overall, these results highlight the potential role that unobserved occupation-specific characteristics might play in confounding the

---

[23]On the other hand, if greater competitiveness is the source of higher returns to working long hours – for example, greater competition within an occupation may increase the returns to working longer hours as workers compete to realize larger potential earnings gains – then reducing the gender gap in hours worked, or reducing the convexity of the pay structure (possibly by reducing the degree of competition in an occupation) may reduce the gender pay gap. See, for example, Landers, Rebitzer and Taylor (1996)'s rat race model.

[24]O*NET online is a comprehensive database of worker attributes and job characteristics for over 900 occupations.

[25]To merge the O*NET occupations to Census/ACS occupations, we use the crosswalk by Autor and Acemoglu (2011). There are about twice as many O*Net occupations than Census occupations and the crosswalk weights each O*Net characteristic levels with the relative number of individuals in each O*Net occupation to get the characteristic values for each of the Census occupations. In order to use the crosswalk, we use version 14.0 (2009) of the O*NET online database.

[26]Similar results are obtained if we use the other Census years.

[27]See Appendix Table 5 for results using the alternative measures of the outcomes.

relationship between the returns to working long hours and the gender pay gap.

3 Empirical Strategy and Results

To establish a causal interpretation of the relationship between the returns to working long hours and the gender pay gap, we need to isolate exogenous variation in either the demand or supply of long hours of work. One natural option is to look for exogenous variation in the returns (β_o) across occupations and over time, ideally generated from underlying changes in the way occupations reward longer hours of work. Unfortunately, occupation-specific changes in the demand for long hours of work, due in part by technological change, globalization, or the diffusion of performance pay schemes, are likely to be correlated with other unobserved demand shocks faced by occupations, which could have an independent effect on the gender pay gap. Moreover, such an approach will also have to confront the inherent difficulty in establishing whether the changes in the observed returns are due to changes in the true returns to working long hours, or increases in worker sorting.

The empirical approach that we pursue in this paper is to utilize supply-side shocks to the cost of providing long hours to provide a causal link between the long hours premium and the size of gender pay gaps. Specifically, we examine if the higher cost to women of working long hours is an important factor in explaining gender pay gap in skilled occupations, particularly in occupations that disproportionately reward individuals who are willing and able to supply longer hours. To identify this effect, we exploit cross-city variation in the cost of providing long hours for skilled women – in particular, we build on earlier work by Cortes (2008) and Cortes and Tessada (2011) that demonstrate that the influx of low-skilled immigration leads to lower prices of services that are close substitutes for household production, and increases the supply of market work among highly skilled women.[28] Following both papers, we utilize plausibly exogenous variation in low-skilled immigrant flows across cities to proxy for changes in the prices of outsourcing household production, thereby providing us with an arguably exogenous shifter of the cost to women of providing long hours in the labor market.

Figure 4 provides a graphical illustration of our empirical strategy. Suppose that there are two occupations, A and B, that differ in terms of their elasticity of annual earnings with respect to weekly hours worked. Assume that occupation A has higher returns to weekly hours of work – that is, $\beta_A > \beta_B$. Suppose that in both occupations, on average, males work λ_m hours per week, while females work λ_f hours per week, where $\lambda_f < \lambda_m$. As depicted in Figure 4, given the greater

---

[28] Using confidential data from the CPS, Cortes (2008) shows how the inflow of low-skilled immigrants to the US has lowered significantly the prices of services in which they concentrate, in particular, of housekeeping, babysitting, and gardening. Due to price data limitations, her analysis is restricted to the US 25 largest city. To be able to extend our analysis to more cities, we follow Cortes and Tessada (2011) and use a reduced form. Note that the functional form of the key explanatory variable (Log of number of low-skilled immigrants) is derived from Cortes (2008)'s model.

convexity of annual earnings with respect to weekly hours worked, the gender gap in earnings is larger in occupation A $(w_m^A - w_f^A)$ as compared to occupation B $(w_m^B - w_f^B)$. A decrease in the cost of providing longer hours of work, that allows women to put in longer hours of work (e.g. λ_f'), would tend to reduce the gender gap in earnings more in occupations with larger returns to working long hours i.e $\Delta G_A > \Delta G_B$.

Expanding this simple framework to many cities and more than two occupations, the low-skill immigration shocks create exogenous variation in the gender gap in hours worked across cities and over time $\Delta\lambda_{ct} = \lambda_{ct,m} - \lambda_{ct,f}'$. Our basic hypothesis is that, if there is indeed a causal relationship between the returns to working long hours and the gender pay gap across occupations, cities that experience larger changes in the gender gap in hours worked, $\Delta\lambda_{ct}$, as a result of low-skilled immigration shocks, would experience a larger reduction in the gender pay gap of highly skilled women in occupations within that city that have a (pre-existing) high demand for long hours of work (larger $\beta_{j,1980}$). To test this hypothesis, we use a triple-difference strategy that examines how the gender pay gap has evolved over time, across cities that vary in terms of their low-skilled immigration inflows, for occupations with higher versus lower returns to long hours.

3.1 Variation in Gender Gaps across Cities and Occupation Groups

Our empirical strategy seeks to explain variation in the gender pay gap across cities and occupation groups over time. To ensure a reasonable sample size to construct the key variables, we aggregate the occupations into 16 broader groups, and restrict the sample to the 59 largest cities.[29,30] The unit of analysis is an occupation-group–city–year. A city in our analysis corresponds to a Metropolitan Statistical Area (MSA) defined by the US Census Bureau.

To estimate the gender pay gap for each MSA c in year t for occupation group j, we estimate regressions of the form of equation (2) but allow the coefficient on the female dummy to vary at both the broad occupation and city levels, and include controls for city, occupation group, and city*occupation group fixed effects. Appendix Table 7 reports the mean, interquartile range, and the standard deviation of the gender pay gap for the full sample (Panel A), within cities (Panel B), and within occupation groups (Panel C) for each year. As observed, there is significant variation in the gender pay gap, even within cities and occupations. Most occupations follow the aggregate trend presented in Figure 1 – gender gaps narrowed from 1980 to 1990, but remained relatively constant or declined only slightly during the last two decades. The exception is "Executive, Administrative, and Managerial Occupations," which has experienced a steady narrowing of the gap, although it is

---

[29]Broad groups are based on the Census classification of occupations (see Appendix Table 1). In the regressions we drop "Police and Detectives in Public Service" because there are not enough observations to construct city level outcomes for this occupation.

[30]Appendix Table 6 provides a list of the MSAs included in the analysis.

15

still at the top in terms of the earnings disparity between men and women.[31]

3.2 Predicted Low-skilled Immigration as an Instrument for the Gender Gap in Overwork

Before turning to the main regression specification, we begin by establishing that low-skilled immigration allows highly skilled women to work longer hours and reduces the gender gap in hours worked. We show that the findings in the Cortes and Tessada (2011) carry through to our sample which includes the most recent time period, and also to a different outcome variable, the gender gap in work hours.

Following Cortes and Tessada (2011), we use intercity variation in low-skilled immigration inflows to identify their effect on the labor supply of high-skilled women. To account for the potential endogeneity in the location choice of low-skilled immigration, we construct a measure of predicted low-skilled immigration that isolates a plausibly exogenous component in the cross-city distribution of low-skilled immigrants by exploring the tendency of immigrants to settle in a city with an existing enclave of immigrants from the same country (Card 2001, Munshi, 2003). Specifically, we use the 1970 distribution of immigrants from a given country across US cities to allocate future aggregate flows of low-skilled immigrants at the national level to individual cities. For example, if a third of Mexican immigrants in 1970 were living in Los Angeles, the predicted low-skilled immigration measure allocates one third of all Mexicans in the 1990s to Los Angeles.

Formally, the instrument for the number of low-skilled immigrants in city c and decade t can be written as:

$$Predicted\_LS\_Immigrants_{ct} = \sum_p \frac{Immigrants_{pc,1970}}{Immigrants_{p,1970}} * LS\_Immigrants_{pt} \qquad (5)$$

where p are all countries of origin included in the 1970 Census, $\frac{Immigrants_{pc,1970}}{Immigrants_{p,1970}}$ is the share of immigrants in 1970 originating from country p living in city c, and $LS\_Immigrants_{pt}$ stands for the aggregate number of low-skilled immigrants from country p to the United States in year t.

To examine the effect of predicted low-skilled immigration on the labor supply of highly-skilled women and the gender gap in hours worked, we estimate the following regressions:

$$Hours\_Outcome_{ict} = \alpha + \delta * \ln(Predicted\_LS\_Immigrants)_{ct} + \gamma X_{ict} + \theta_c + \mu_t + \varepsilon_{ct} \qquad (6)$$

---

[31]Note that the differential trend for this broad occupation might be explained by compositional changes, as it includes managers at very different levels, from Chief Executives to Funeral Directors.

$$Gender\_Gap\_Hours\_Outcome_{ct} = \alpha + \delta * \ln(Predicted\_LS\_Immigrants)_{ct} + \theta_c + \mu_t + \varepsilon_{ct} \quad (7)$$

where the subscripts i refers to the individual, c the city and t the Census/ACS year. $Hours\_Outcome$ is either usual hours worked per week or an indicator variable for working at least 50 hours a week. θ_c and μ_t denote city and year fixed effects, respectively. The vector X_{ict} represents individual-level demographic controls which include education dummies, race dummies, a quartic in age, and an indicator for being single. All specifications include city and year fixed effects. Standard errors are clustered at the city level to allow for the possibility of serial correlation within cities across years.[32]

Panel A of Table 3 reports the individual-level regression estimates of equation (6) for the female sample for two outcomes – an indicator for working 50 or more hours a week (Column (1)) and the usual weekly hours worked (Column (2)). The coefficients on log predicted low-skilled immigration are positive and statistically significant, confirming that in cities that experience a larger inflow of low-skilled immigrants, highly-skilled women are more likely to increase their time spent on market work. The magnitude of the estimates imply that an increase in the predicted low-skilled immigration flow from 1980 to 2010 led to an increase in the usual hours worked of full-time college educated women by about 15 minutes per week and increased the probability that she works 50 or more hours per week by 1.3 percentage points (relative to a baseline of about 20 percent).[33] Columns (3) to (6) show that the same immigration flow reduced the gender gaps in these labor supply outcomes by approximately the same magnitude at the city level. The latter finding supports the idea that low-skilled immigration flows impacted the labor supply decisions of highly-skilled women, but had little effect on the labor supply of highly-educated men.

3.3 Triple-Difference Estimates

The triple-difference procedure compares changes in the gender wage gap over time in cities with high vs. low immigration inflows, for occupations with high vs. low returns to weekly hours worked. To implement this approach, we require an occupation-specific measure of the returns to overtime that captures its pre-existing demand for long hours of work. We construct the returns to long hours at the occupation-level using data from 1980. As this variable is meant to capture an

---

[32]These specifications are similar to that in Cortes and Tessada (2011) with a few extensions and differences. First, we extend the sample period to include the 2011 ACS 3-year aggregate (2009-2011) and restrict the sample to the 59 MSAs for which we can construct occupation-specific outcomes. We also restrict the sample to individuals who work full-time. Second, we also consider the gender gap in weekly hours worked as the key dependent variable. Compared to using changes in the levels of hours worked, it is less likely that the 1970 distribution of immigrants is correlated with unobserved determinants of changes in gender gaps.

[33]This result is very similar to the one reported by Cortes and Tessada (2011) who found that the low-skilled immigration flows from 1980 to 2000 increased the probability of working 50 hours or more by 1.8 percentage points, for women working in occupations in which men work very long hours.

intrinsic component of the occupation, we do not allow it to vary at the city level.[34] To estimate the occupation-group specific returns to working long hours, we estimate equation (1) using the sample of male full-time college educated workers from the 1980 Census and the broad occupation classifications (see Appendix Table 2)

To implement the triple-difference strategy, we begin by estimating reduced form regressions of the form:

$$
\begin{aligned}
Gender\_gap\_pay_{cjt} = \alpha \quad & + \quad \Phi * Return\_overwork_{j,1980} * \ln(Predicted\_LS\_Immigrants)_{ct} \\
& + \quad \kappa * X_{cjt} + \pi_c + \theta_j + \mu_t + \varphi_{cj} + \tau_{ct} + \gamma_{jt} + \varepsilon_{cjt}.
\end{aligned}
\tag{8}
$$

where c refers to the city, j the occupation group, and t the Census year/ACS year. π_c, θ_j, and μ_t are fixed effects for city, industry, and year, respectively. We also include all the relevant two-way fixed effects – namely, city*industry (φ_{cj}), city*time (τ_{ct}), and occupation*time (γ_{jt}). In some specifications, we include a vector of controls, X_{cjt}, for the share of full-time workers in the occupation who are female, the labor force participation of women who reported being in that occupation, log average wage of males working exactly 40 hours, and the share of males working 50 or more hours a week. The first two variables account for the possibility that predicted low-skilled immigrant inflows may affect women's occupational choice, or their decision to participate in the labor market, which might result in potential selection effects when considering the effects on the gender wage gap.[35] The latter two controls account for the possibility that the instrument may be correlated with city*occupation shocks to wage levels and the overall prevalence of overwork.

We cluster standard errors simultaneously at the city and occupation-group levels (Cameron, Gelbach, and Miller, 2011).[36] Our key variable of interest is the interaction term: $Return\_overwork_{j,1980} * \ln(Predicted\_LS\_Immigrants)_{ct}$. The first component in the interaction is the occupation-specific measure of the returns to overwork for occupation-group j in 1980. The second variable in the interaction, $\ln(Predicted\_LS\_Immigrants)_{ct}$, captures the exogenous shifter of the cost of providing long hours across cities c over time t, as discussed above. If our hypothesis is true, we

---

[34]Goldin (2014) provides some examples of occupation characteristics that might be correlated with a high return to working long hours. Fundamentally, she points out that nonlinearities in the relationship between work hours and earnings is likely to arise whenever an employee does not have perfect substitutability. Imperfect substitutability of workers is more likely arise in occupations that are more client-oriented, are more structured, have greater degree of time pressure and require more flexibility of decision making.

[35]In Appendix Table 8, we show that replacing the outcome in equation (8) with either the share of full-time workers in the occupation who are female or the labor force participation of women who reported being in that occupation suggests that a reduction in the gender gap in overwork does not appear to lead to more women choosing to work in occupations with higher returns to overwork. Most coefficients are not statistically significant, and the two that are marginally significant (at the 10% level) go in the opposite direction.

[36]Note that Cameron, Gelbach, and Miller (2011)'s method works reasonably well when at least one of the dimensions has many clusters—we have 59 cities—even if there are few clusters in the other dimension, as in our case, where we have 15 broad occupations.

expect the coefficient on the interaction term (Φ) to be positive – that is, an increase in predicted low-skilled immigrant flows should decrease the cost to providing long hours, thereby reducing the gender gap in occupations where the returns to overwork are the highest.[37]

We also present 2SLS estimates using $\ln(Predicted\_LS\_Immigrants)_{ct}$ as an instrument for the gender gap in the likelihood of working overtime. The regression specification is:

$$
\begin{aligned}
Gender\_gap\_pay_{cjt} = \alpha \quad + & \quad \phi * Return\_overwork_{j,1980} * Gender\_gap\_long\_week_{ct} \\
+ & \quad \kappa * X_{cjt} + \pi_c + \theta_j + \mu_t + \varphi_{cj} + \tau_{ct} + \gamma_{jt} + \varepsilon_{cjt}
\end{aligned}
\tag{9}
$$

where we instrument for the interaction term using the measure of predicted low-skilled immigrants interacted with $Return\_overwork_{j,1980}$. Similar to the interpretation for the reduced form model shown in equation (8), $\phi > 0$ implies that a decrease in the gender gap in working long hours (i.e. $Gender\_gap\_long\_week_{ct}$ becomes less negative) leads to a reduction in the gender pay gap, particularly for occupations with a high return to working long hours. It is worth highlighting that one important advantage of the triple difference approach is that we are able to include a full set of city*year fixed effects in the regression. This is likely to mitigate concerns about the exogeneity and interpretation of the instrument since the identification assumption would only be violated if unobserved determinants of the location choice of immigrants in 1970 are correlated with shocks to gender gaps in particular occupations in a given city.

The results for the reduced form specification as detailed in equation (8) are reported in Table 4. All the regressions include year, city and occupation-group fixed effects, as well as year*city, year*occupation, and city*occupation fixed effects and are weighted by cell size. The baseline estimate of the effect of $Return\_overwork_{j,1980} * \ln(Predicted\_LS\_Immigrants)_{ct}$ on the gender pay gap is positive and statistically significant at the 10% level (Column (1)), confirming our hypothesis that an increase in low-skilled immigration reduces the gender gap in earnings in occupations with higher returns to working long hours, relative to occupations with lower returns. The estimate becomes larger and more statistically significant with the inclusion of additional time-varying controls at the city and occupation level (Column (2)). The results do not appear to be driven by particular occupations with unusually high or low returns – as shown in Columns (3) and (4) the estimated effects remain largely similar when we exclude occupations with the two highest and two lowest returns.[38]

In Columns (5) to (8), we allow $Return\_overwork_{c,1980}$ to enter non-linearly in the interaction term. We classify occupation groups into three groups based on their 1980 returns, and interact each category with our measure of predicted low-skilled immigration. Across the four specifications, we

---

[37] Note that the gender gap in pay is defined as female - male, and thus it is negative.

[38] We also run models excluding each of the 15 broad occupation groups in turn. All coefficients are statistically significant at the 5 percent level, except for one that is significant at the 10 percent level.

find consistently that low-skill immigration inflows significantly reduce the gender gap in earnings in occupations with the highest returns (top tercile), relative to occupations with returns in the middle or lowest terciles. The magnitude of the estimates indicate that a 10% increase in predicted low-skill immigration reduces the gender earnings gap in occupations with the highest returns (top third) by 0.18 to 0.3 percentage points, relative to occupations with the lowest returns (bottom third). Put differently, these estimates imply that, for the average city, the increase in low-skilled immigration from 1980 to 2010 led to a 2-3 percentage point decline in the gender pay gap for occupations in the top tercile of returns relative to occupations in the lowest tercile. This corresponds to about 20 to 30 percent of the standard deviation of the gender pay gap across occupation groups.

Table 5 presents the OLS and 2SLS estimates of equation (9). Panel A reports the coefficient on the interaction of $\beta_{j,1980}$ and the gender gap in working 50+ hours per week, while Panel B reports the coefficient on the interaction of $\beta_{j,1980}$ and the gender gap in weekly hours worked. In the 2SLS specifications, we instrument for each interaction term using $\beta_{j,1980}*\ln(Predicted\_LS\_ImmigrantFlow)_{ct}$. As in the previous specifications, we include the full set of city*year, occupation*year, and city*occupation fixed effects, and use cell sizes as weights in all the regressions.

As shown in Column (1), the baseline OLS estimates of the relationship between gender pay gaps and the interaction between the occupation-specific returns to working long hours and each of the two measures of the gender gap in labor supply (probability of working 50+ hours per week, and average weekly hours) are positive and almost statistically significant at the 10% level. The point estimate remains positive, but is smaller and not statistically significant when we include the usual set of time-varying controls at the city and occupational level (Column (2)). The key difficulty in attaching a causal interpretation to the OLS estimates is that the interaction of changes in the gender gap in work hours within cities and over time and the occupation-specific returns to working long hours are likely to be correlated with unobserved demand and supply shocks that may independently affect gender pay gaps. The instrument that we use isolates a plausibly exogenous supply-side shock – predicted low-skilled immigration – that affects the gender gap in work hours of highly skilled workers, but is unlikely to have any direct effect on earnings.

Columns (3) to (10) of Table 5 present the 2SLS estimates. The IV estimates from the most basic specification (Columns (3)) using the full sample of occupations are positive, significant at the 10 percent level, and larger than the corresponding OLS estimates. Controlling for the set of additional city*occupation*time varying controls tends to increase the magnitude and statistical significance of the baseline IV estimates (Column (4)). The magnitude of the estimated effects remain similar when we exclude occupations with the highest and lowest returns (Columns (5) and (6)). As observed in Columns (7) to (10), consistent with the results observed in the last four columns of Table 4, reductions in the gender gap in the likelihood of working 50+ hours per week or the gender gap in weekly hours worked reduces the gender pay gap in occupations with the highest

20

(top third) returns to working long hours, relative to occupations with returns in the middle or lowest tercile. The Angrist and Pischke F-statistics for the instruments are close to or larger than 10 for most of the specifications instrumenting for the interaction of the returns to long hours and the gender gap in the probability of working 50+ hours per week (top panel). The F-statistics are somewhat smaller in the bottom panel when instrumenting for the interaction of the returns to long hours and the gender gap in average weekly hours worked.

Overall, these results provide empirical support of a causal link between the returns to working long hours and the gender pay gap, and suggests that reductions in the cost of supplying longer hours of work allow women to close the gap in work hours and to benefit from higher wages. The magnitudes of the estimates imply that, for the average city, a one standard deviation reduction of the gender gap in overwork will decrease the gender pay gap between occupations in the top tercile and the bottom tercile of 1980 returns by between 5 and 7.5 percentage points. This corresponds to between 50 to 75 percent of the cross-occupation standard deviation of gender pay gaps.

Finally, in Table 6, we present a number of robustness tests to explore the sensitivity of our results to alternative measures of the returns to long hours and the gender pay gap. As discussed earlier, although we believe that there is good reason to use the returns to working long hours constructed from the male sample in the earliest time period in our sample (1980) to proxy for occupation-specific demand for overwork, our maintained assumption is that these returns capture a feature that is inherent to the occupation that presumably does not change substantially over time. Moreover, we also assume that these returns are also relevant for females. Columns (1) to (3) report the reduced form and 2SLS estimates when we replace $\beta_{j,1980}$ in the interaction term with similar returns constructed using the 2010 data ($\beta_{j,2010}$). While the point estimates are somewhat smaller when we use the 2010 returns, the overall results are qualitatively similar, and the estimates are statistically significant at the 10% level. Columns (4) to (6) show that the results are also robust to using both the male and female sample to construct the returns to working long hours in 1980.

The last three columns consider specifications where we use alternative measures of the gender gap and the returns to overwork based on hourly wages instead of annual earnings. The hourly wage premium for overwork is defined as the difference in average log hourly wages between males who work 50 or more hours per week and full-time males who work fewer than 50 hours per week (see equation (3)). We construct this variable separately for each occupation using the 1980 data, analogous to the returns measure constructed using annual earnings. We use the gender gap in hourly wages as the outcome variable in these specifications (see equation (4)). As observed in Columns (7) to (9), the coefficient estimates for the corresponding reduced form and 2SLS specifications using these variables are all positive and statistically significant, suggesting that our main findings in Tables 4 and 5 are robust to using an alternative functional form for the specification of the returns to overwork, and to using hourly earnings instead of annual earnings in

the measure of the gender pay gap.

4 Conclusion

Women have made enormous gains in reversing the education gap and increasing their representation in skilled occupations. Despite these gains, the gender gap at the top of the skill distribution has stalled for the past three decades. This paper highlights the importance of a particular occupational characteristic – the returns to working long hours – and examines the causal mechanism through which it contributes to the perpetuation of gender pay gaps among highly skilled workers. Our results suggest that high returns to overwork, coupled with the persistent gap in the propensity to work overtime across genders, is an important factor that limits the convergence in gender pay gaps in many highly-skilled occupations. Furthermore, the fact that the returns to working long hours appears to have risen particularly rapidly for highly skilled workers over time, suggests that these forces may play an increasingly large role in slowing the convergence in gender wage gaps in the future.

How should we address this mechanism? One possibility is to focus on policies that increase the supply of substitutes to household production, thus enabling women to increase their labor supply. While we provide some evidence in this paper that there might be some scope for affecting the gender pay gap through this channel, the fact that women shoulder a disproportionate share of housework and child care, coupled with slow-changing social norms that dictate the traditional division of labor within households (Bertrand, Kamenica, and Pan, 2015, Fortin, 2013), suggests that attempts to address the gender gap in earnings through policy innovations that target the supply side of labor are likely to be limited in their effectiveness.

Focusing on the sources of the returns to working long hours may offer a more promising solution. Goldin (2014) suggests several examples of occupations and sectors that have moved toward greater hours flexibility such as physicians, pharmacists and veterinarians. The causes of these changes are varied, ranging from re-organizing work to take advantage of economies of scale, lower labor costs or because of employee pressure. Some countries such as Korea and Japan have moved toward legislations that restrict the maximum number of hours of work per week with the explicit aim to reduce working hours and to promote work-life balance. The effectiveness and desirability of such a policy will depend on the source of the returns to working long hours. If the returns to overtime are the result of market imperfections such as incomplete information as suggested by Landers, Rebitzer and Taylor (1996), a policy to reduce hours of work through government intervention may be welfare enhancing. On the other hand, if the higher returns to hours worked is an optimal response to technological change and globalization (Cha and Weeden, 2014), or an intrinsic characteristic of how work in an occupation is organized (Goldin, 2014), such policies could have detrimental effects

on firm productivity. If the key driver of the returns to overwork are the latter, then policies to promote the redesign and the reorganization of work to enhance temporal flexibility are likely to be more effective.

References

[1] Acemoglu, Daron and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." Handbook of Labor Economics Vol. 4, Orley Ashenfelter and David E. Card (eds.). Amsterdam: Elsevier, forthcoming.

[2] Arulampalam Wiji, Alison L. Booth and Mark L. Bryan. 2007. "Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution." *Industrial and Labor Relations Review*. Vol. 60(2): 163-186.

[3] Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender Identity and Relative Income within Households." *Quarterly Journal of Economics*. Vol 130(2): 571-614.

[4] Blau, Francine D. 2012. "The Sources of the Gender Pay Gap." Pp. 189-210 in The New Gilded Age: The Critical Inequality Debates of Our Time, edited by D. B. Grusky and T. Kricheli-Katz. Stanford: Stanford University Press.

[5] Blau, Francine D. and Lawrence M. Kahn. 2006. "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence." *Industrial and Labor Relations Review*. Vol. 60 (1): 45-66.

[6] Blau, Francine D. and Lawrence M. Kahn. 2016. "The Gender Wage Gap: Extent, Trends, and Explanations." NBER Working Paper 21913.

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

[8] Cameron, Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering". *Journal of Business & Economic Statistics*. Vol. 29(2): 238-249

[9] Card, David. 2001. "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*. Vol. 19: 22-64.

[10] Cortes, Patricia. 2008. "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data." *Journal of Political Economy*. Vol. 116(3): 381-422.

[11] Cortes, Patricia and Jose Tesseda. 2011. "Low-skilled Immigration and the Labor Supply of Highly Skilled Women." *American Economic Journal: Applied Economics*. Vol 3(3): 88-123.

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

[13] Fortin, Nicole. "Gender Role Attitudes and Women's Labor Market Participation: Opting Out, AIDS, and the Persistent Appeal of Housewifery," University of British Columbia Working Paper, December 2013.

[14] Flory, Jeffrey A., Andreas Leibbrandt and John A. List. 2014. "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job-Entry Decisions." *Review of Economic Studies*, forthcoming.

[15] Gicheva, Dora. 2013. "Working Long Hours and Early Career Outcomes in the High-End Labor Market". *Journal of Labor Economics*. Vol. 31(4): 785-824.

[16] Gneezy, Uri, Muriel Niederle and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*. Vol 118: 1049-1074.

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

[18] Kantor, Jodi and David Streitfeld. "Inside Amazon: Wrestling Big Ideas in a Bruising Workplace." *New York Times*, August 15.

[19] Kuhn, Peter J. and Fernando A. Lozano. 2008. "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S.Men, 1979–2006." *Journal of Labor Economics*. Vol. 26: 311-43.

[20] Landers, Renee M., James B. Rebitzer, and Lowell J. Taylor. 1996. "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms." *American Economic Review*. Vol. 86(3): 329-348.

[21] Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economis*. 118: 549-99.

[22] Olivetti, Claudia and Barbara Petrongolo. 2008. "Unequal Pay or Unequal Employment? A Cross-country Analysis of Gender Gaps." *Journal of Labor Economics.* Vol. 26(4): 621-654.

[23] Ors, Evren, Frederic Palomino and Eloic Peyrache. 2012. "Performance Gender-Gap: Does Competition Matter?" *Journal of Labor Economics*. Vol. 31(3): 443-499.

[24] Wasserman, Melanie. 2015. "Hours Contraints, Occupational Choice and Fertility: Evidence from Medical Residents." MIT Working Paper.

Figure 1. Trends in Labor Market Outcomes for Full-time Workers, by Education Level
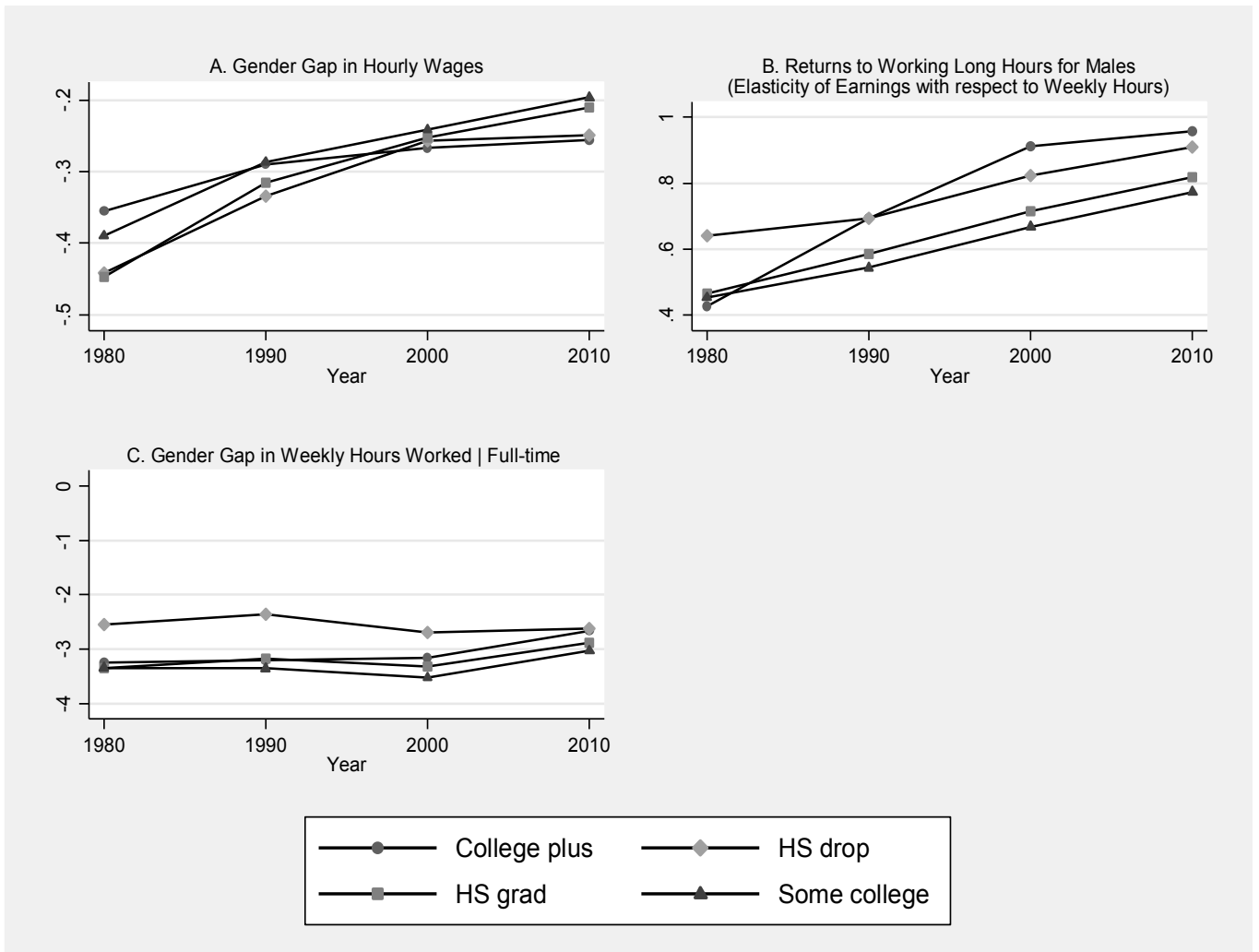
Note. The data is from the 1980, 1990, 2000 Census and 3-year aggregate 2011 ACS (2009-2011). The sample is restricted to native-born age 25-64 who report working full-time (35 hours or more) in a given week. See text for a discussion on the construction of the variables.

Figure 2. Cross-occupation Correlation between Gender Pay Gap and Returns to Working Long Hours
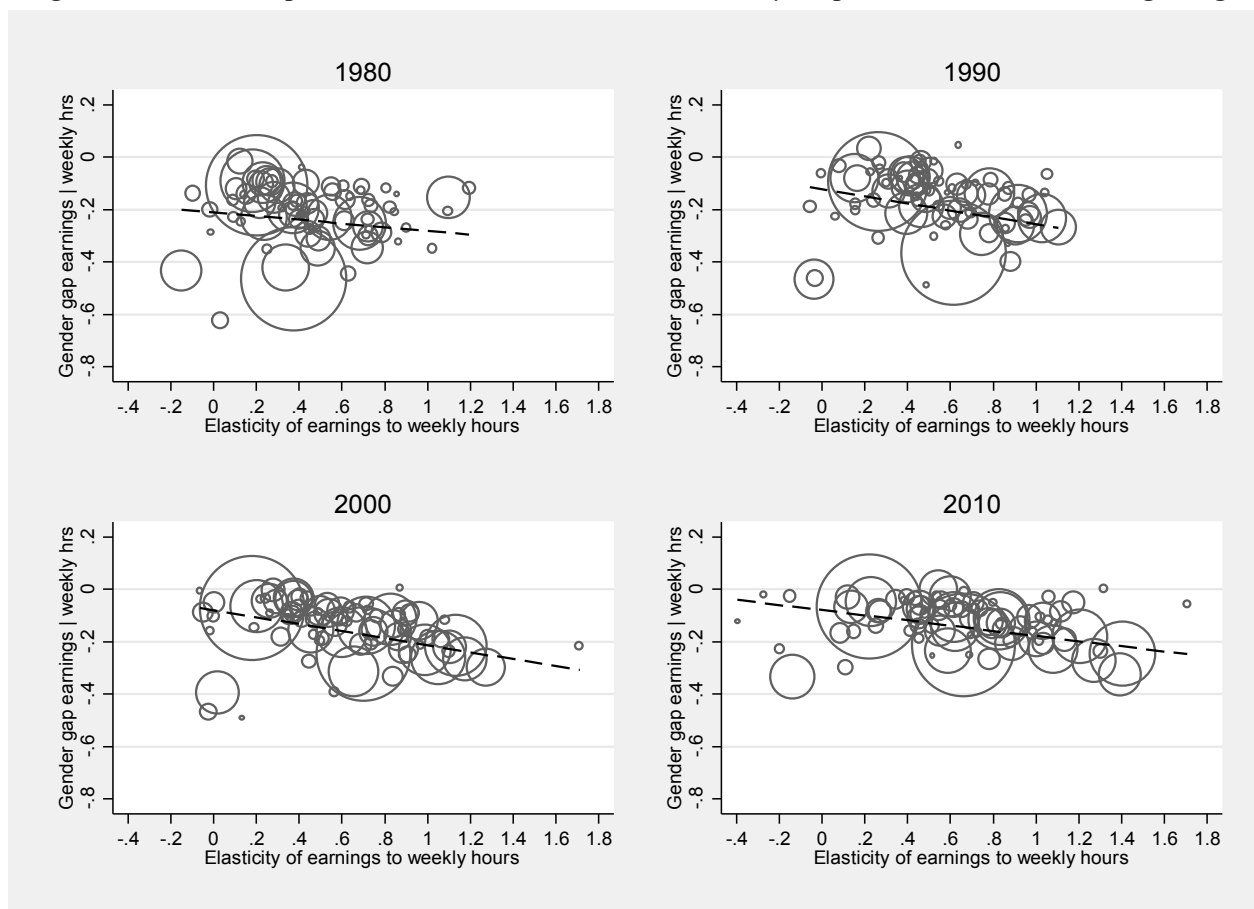
Note. The data is from the 1980, 1990, 2000 Census and 3-year aggregate 2011 ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. Each skilled occupation is weighted by sample size.

Figure 3. Cross-occupation Correlations between Competitiveness, Gender Pay Gap and Returns to Working Long Hours in 2010
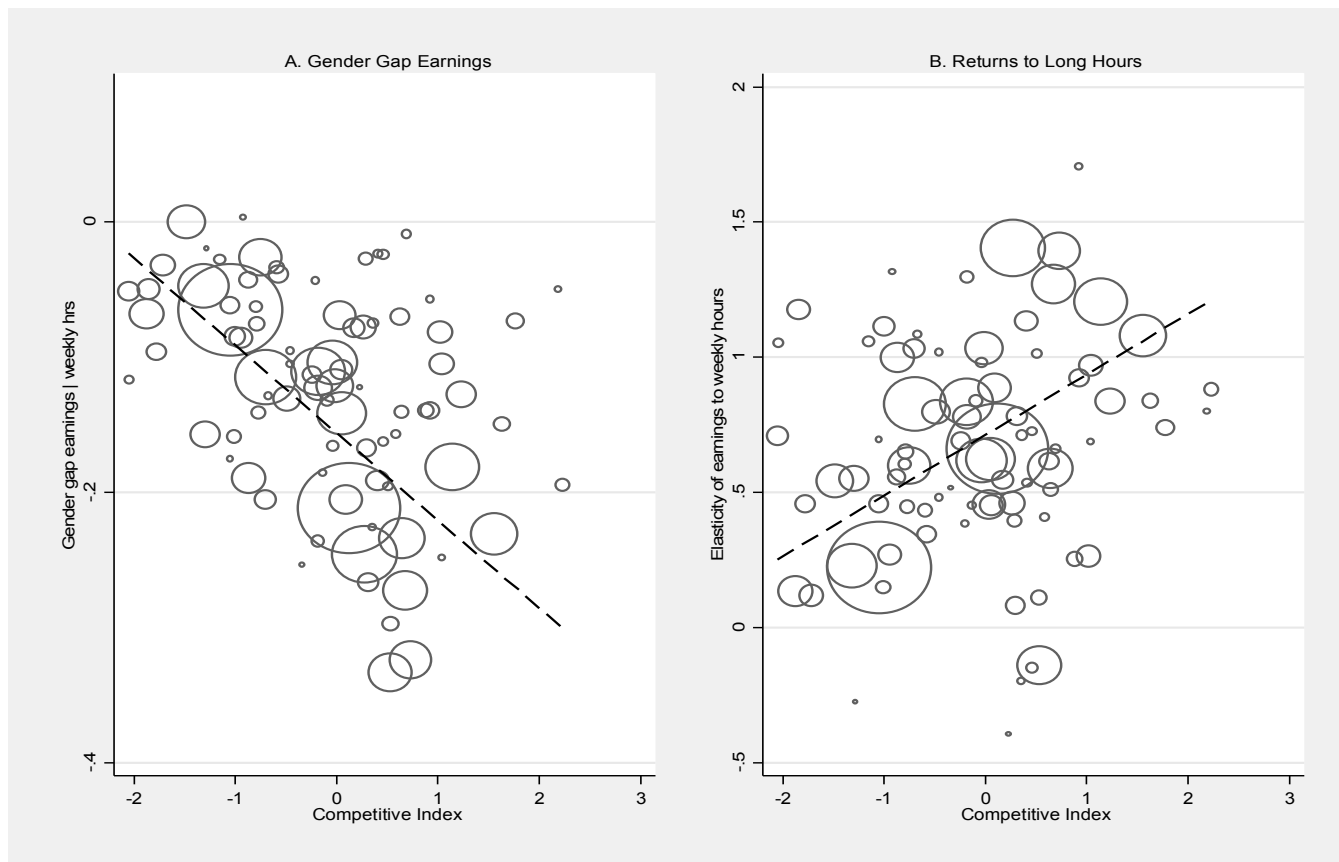
Note. The data is from the 1980, 1990, 2000 Census and 3-year aggregate 2009 ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. Each skilled occupation is weighted by sample size. The competitive index is computed using ONET and standardized to have a mean of 0 and standard deviation of 1.

Figure 4. Relationship between Gender Pay Gap and Gender Gap in Work Hours for Occupations with Different Returns to Working Long Hours

Annual Wages

β_A

β_B

ΔG_A

ΔG_B

35   λ_f λ'_f λ_m Hours worked per week
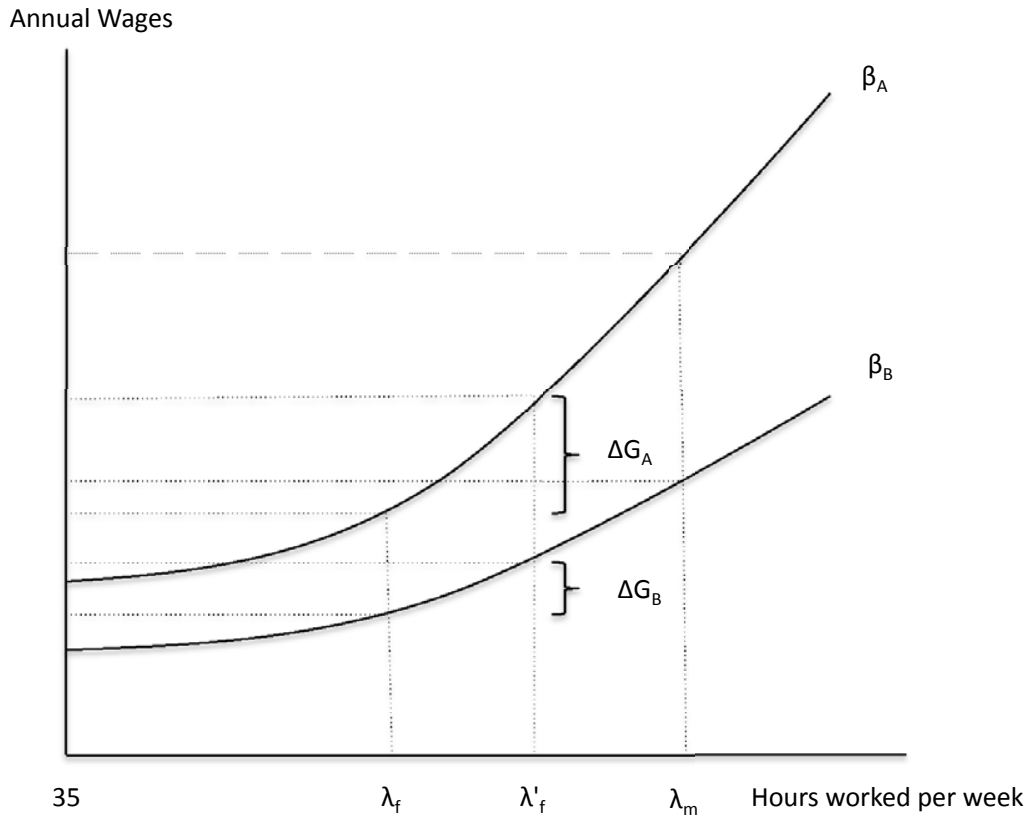
Table 1. Cross-occupation Correlation between Returns to Working Long Hours and the Gender Pay Gap by Year

	Outcome: Gender Pay Gap							
	1980		1990		2000		2010	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns to Long Hours	-0.047	-0.076**	-0.136*	-0.070*	-0.135***	-0.080**	-0.097**	-0.066**
	[0.083]	[0.032]	[0.081]	[0.037]	[0.044]	[0.033]	[0.041]	[0.026]
Controls		X		X		X		X
Observations	85	86	94	93	95	94	95	94

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The unit of observation is an occupation. The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. The gender pay gap is the coefficient on female*occupation dummy in a regression of log annual earnings on the full set of female*occupation dummies controlling for the hours worked per week, weeks worked per year, occupation fixed effects and a vector of demographic characteristics that include a quartic in age, a female dummy, race fixed efects and an indicator for whether an individual has a masters or doctoral degree (see equation (1) for more details). The returns to working long hours is estimated on the sample of male workers and is the coefficient on the interaction between log hours worked per week and the occupation dummy in the regression of log yearly earnings on the full set of interactions between occupations*ln(hours worked per week) conditional on the weeks worked per year, occupation fixed effects and the same set of demographic characteristics listed above (see equation (2) for more details). Both the gender pay gap and returns to working long hours were estimated separately for each Census/ACS year. See the text for the choice of the occupations. The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size in the regression. Robust standard errors are reported in brackets. ***significant at 1%, **5%, *10% level.

Table 2. Correlations of Competitive Index with Gender Gap and Returns to Long Hours in 2010

	(1)	(2)	(3)	(4)
	A. Outcome: Returns to Long Hours			
Competitive Index	0.214***	0.277***		
	[0.056]	[0.062]		
	B. Outcome: Gender Pay Gap (Female - Male)			
Competitive Index	-0.059***	-0.031***	-0.050***	-0.018**
	[0.009]	[0.010]	[0.014]	[0.009]
Returns to Long Hours			-0.039	-0.051*
			[0.043]	[0.027]
Controls		X		X
No. Observations	93	93	93	93

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The unit of observation is an occupation. The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. Please see Table 1 and the text for details on the outcome variables. The competitive index is constructed using data from ONET and is standardized with a mean of 0 and standard deviation of 1 in the full sample of occupation (see text for details). The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size in the regressions. Robust standard errors are reported in brackets. ***significant at 1%, **5%, *10% level.

Table 3. Predicted Low-skilled Immigration flows and High Skilled Female Labor Supply

	A. Micro Data, Female Sample		B. City Level Data			
	Indicator for Working 50+ hrs	Usual Weekly Hours	Gender Gap (Female-Male) for Working 50+ hrs		Gender Gap (Female-Male) in Weekly Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Predicted LS Immigration)	0.011**	0.199***	0.015***	0.013**	0.289***	0.208**
	[0.004]	[0.074]	[0.005]	[0.005]	[0.097]	[0.099]
Weights	Person	Person	Unweighted	Cell size	Unweighted	Cell size
Controls						
Demographic controls	X	X				
City FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Observations	759,273	757,934	236	236	236	236
R-squared	0.040	0.044	0.757	0.804	0.755	0.793

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. See text for details on the construction of the Ln(Predicted LS Immigration) flows and for the gender gaps. The unit of analysis for Panel B is at the city*year level. There are 59 MSAs and 4 time periods. "Demographic controls" include dummies for a masters degree, more than a masters degree, a quartic in age, race dummies and an indicator for being single. Standard errors are clustered at the MSA level and are reported in brackets. ***significant at 1%, **5%, *10% level.

Table 4. Causal Effect of Working Long Hours on the Gender Gap: Reduced Form Estimates

	Outcome: Gender Gap in Earnings | Weekly Hours							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns Long Hours$_{j, 1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$	0.026*	0.042***	0.032	0.049***				
	[0.013]	[0.010]	[0.021]	[0.013]				
Top tercile$_{j, 1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$					0.017***	0.029***	0.029**	0.030***
					[0.003]	[0.003]	[0.010]	[0.005]
Middle tercile$_{j, 1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$					0.003	0.009	0.009	0.009
					[0.011]	[0.008]	[0.008]	[0.008]
Excluded Occupations	None	None	2 Highest Returns	2 Lowest Returns	None	None	2 Highest Returns	2 Lowest Returns
Controls								
Year, City, Occ FE	X	X	X	X	X	X	X	X
Year*City FE	X	X	X	X	X	X	X	X
Year*Occ FE	X	X	X	X	X	X	X	X
City*Occ FE	X	X	X	X	X	X	X	X
Other Controls		X	X	X		X	X	X
Observations	2,348	2,346	2,037	2,027	2,348	2,346	2,037	2,027

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. The unit of observation is an occupation-group*MSA*year. There are 59 MSAs, 15 occupation groups and 4 time periods. "Returns Long Hours" is the elasticity of annual income to weekly hours for the broad occupation group in 1980 obtained from the estimation of equation (1). See text for details on the construction of the Ln(Predicted LS Immigration) flows and for the gender gaps. The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size. Robust standard errors are two-way clustered at the MSA and broad occupation levels and are shown in brackets.***significant at 1%, **5%, *10% level.

Table 5. Causal Effect of Working Long Hours on the Gender Gap: 2SLS Estimates

	OLS		2SLS							
			Outcome: Gender Gap in Earnings | Weekly Hours							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Returns Long Hours$_{j,1980}$ x Gender Gap Work 50+ hrs$_{ct}$	0.604 [0.371]	0.473 [0.328]	2.073* [1.030]	3.394*** [1.120]	2.683 [1.928]	4.117*** [1.292]				
Top tercile $_{j, 1980}$ x Gender Gap Work 50+ hrs$_{ct}$							1.526*** [0.093]	2.542*** [0.755]	3.438 [1.938]	2.662*** [0.793]
Middle tercile $_{j, 1980}$ x Gender Gap Work 50+ hrs$_{ct}$							0.177 [0.835]	0.695 [0.520]	0.689 [0.591]	0.741 [0.519]
F-stat for Instrument			10.76	11.23	9.68	13.23	8.9; 10.4	9.2; 10.1	13.2; 10.4	9.3; 10.7
Returns Long Hours$_{j,1980}$ x Gender Gap Hours Worked$_{ct}$	0.040* [0.021]	0.029 [0.022]	0.135 [0.079]	0.220** [0.098]	0.175 [0.141]	0.258** [0.107]				
Top tercile $_{j, 1980}$ x Gender Gap Hours$_{ct}$							0.103*** [0.035]	0.170** [0.077]	0.270 [0.201]	0.176** [0.080]
Middle tercile $_{j, 1980}$ x Gender Gap Hours$_{ct}$							0.011 [0.054]	0.044 [0.034]	0.043 [0.039]	0.046 [0.032]
F-stat for Instrument			6.57	7.00	6.04	8.78	5.1; 6.5	5.2; 6.5	4.8; 6.7	5.2; 7.1
Excluded Occupations	None	None	None	None	2 Highest Returns	2 Lowest Returns	None	None	2 Highest Returns	2 Lowest Returns
Controls										
Year, City, Occ FE	X	X	X	X	X	X	X	X	X	X
Year*City FE	X	X	X	X	X	X	X	X	X	X
Year*Occ FE	X	X	X	X	X	X	X	X	X	X
City*Occ FE	X	X	X	X	X	X	X	X	X	X
Other Controls		X		X	X	X		X	X	X

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. The unit of observation is an occupation-group*MSA*year. There are 59 MSAs, 15 occupation groups and 4 time periods. "Returns Long Hours" is the elasticity of annual earnings to weekly hours for the broad occupation group in 1980 obtained from the estimation of equation (1). The instrument is Ln(Predicted LS Immigration) -- see the text for a description on how it is constructed. The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size. Robust standard errors are two-way clustered at the MSA and broad occupation levels and are shown in brackets.***significant at 1%, **5%, *10% level.

Table 6. Robustness Tests

Outcome:	Gender Gap in Earnings | Weekly Hours						Gender Gap in Hourly Wage		
Measure of Returns to Long Hours:	Elasticity of Annual Earnings to Weekly Hrs $_{j,\ 2010}$			Elasticity of Annual Earnings to Weekly Hrs $_{j\ f+m,\ 1980}$			Hourly Wage Premium for Overwork$_{j,1980}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	RF	RF	2SLS	RF	RF	2SLS	RF	RF	2SLS
Returns$_j$ x Ln(Predicted LS Immigrant Flow)$_{ct}$	0.019** [0.009]			0.026** [0.010]			0.102** [0.042]		
Top tercile $_j$ x Ln(Predicted LS Immigrant Flow)$_{ct}$		0.018* [0.009]			0.020** [0.009]			0.032*** [0.005]	
Mid tercile$_j$ x Ln(Predicted LS Immigrant Flow)$_{ct}$		0.008 [0.007]			0.011 [0.009]			0.013 [0.008]	
Returns$_j$ x Gender Gap Overwork $_{ct}$			1.544** [0.643]			2.061** [0.879]			8.028** [3.065]
F-stat for Instrument			11.12			11.68			11.45
Controls									
Year, City, Occ FE	X	X	X	X	X	X	X	X	X
Year*City FE	X	X	X	X	X	X	X	X	X
Year*Occ FE	X	X	X	X	X	X	X	X	X
City*Occ FE	X	X	X	X	X	X	X	X	X
Other controls	X	X	X	X	X	X	X	X	X
Observations	2,346	2,346	2,346	2,346	2,346	2,346	2,346	2,346	2,346

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. The unit of observation is an occupation-group*MSA*year. There are 59 MSAs, 15 occupation groups and 4 time periods. The instrument for the 2SLS models is Ln(Predicted LS Immigratio) -- see the text for a description on how it is constructed. The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size. Robust standard errors are two-way clustered at the MSA and broad occupation levels and are shown in brackets.***significant at 1%, **5%, *10% level.

Appendix Table 1. List of Skilled Occupations

Code	Occupation	% with College Degree	Median Male Wage Income	Broad Occupation Category
4	Chief executives, public administrators and legislators	0.66	121968	Executive, Administrative, and Managerial Occps
7	Financial managers	0.59	91476	Executive, Administrative, and Managerial Occps
8	Human resources and labor relations managers	0.56	77000	Executive, Administrative, and Managerial Occps
13	Managers and specialists in marketing, advert, PR	0.69	92493	Executive, Administrative, and Managerial Occps
14	Managers in education and related fields	0.81	76336	Executive, Administrative, and Managerial Occps
15	Managers of medicine and health occupations	0.59	82525	Executive, Administrative, and Managerial Occps
18	Managers of properties and real state	0.40	63957	Executive, Administrative, and Managerial Occps
19	Funeral Directors	0.37	53641	Executive, Administrative, and Managerial Occps
22	Managers and administrators, n.e.c	0.48	78000	Executive, Administrative, and Managerial Occps
23	Accountants and auditors	0.75	76230	Business and Financial Operations Occupations
24	Insurance underwritters	0.51	70000	Business and Financial Operations Occupations
25	Other financial specialists	0.63	82525	Business and Financial Operations Occupations
26	Management analysts	0.75	90000	Business and Financial Operations Occupations
27	Personnel, HR, training	0.53	65000	Business and Financial Operations Occupations
33	Purchasing managers, agents, and buyers, n.e.c	0.46	66066	Business and Financial Operations Occupations
34	Business and promotion agents	0.56	61894	Business and Financial Operations Occupations
36	Inspectors and compliance officers, outside	0.56	65000	Business and Financial Operations Occupations
37	Management support occupations	0.54	66020	Business and Financial Operations Occupations
43	Architects	0.89	75652	Architects and Engineers
44	Aerospace engineers	0.83	98591	Architects and Engineers
45	Metallurgical and materials engineers	0.69	76283	Architects and Engineers
47	Petroleum, mining and geological engineers	0.80	103156	Architects and Engineers
48	Chemical engineers	0.85	100000	Architects and Engineers
53	Civil engineers	0.83	81312	Architects and Engineers
55	Electrical engineers	0.71	87683	Architects and Engineers
56	Industrial engineers	0.66	75000	Architects and Engineers
57	Mechanical engineers	0.69	76230	Architects and Engineers
59	Engineers and other proefessionals n.e.c	0.79	86394	Architects and Engineers
64	Computer systems analysists and computer scientists	0.52	68099	Computer and Mathematical Occupations
65	Operations and systems researchers and anlaysts	0.57	72210	Computer and Mathematical Occupations
66	Actuaries	0.98	121968	Computer and Mathematical Occupations
68	mathematicians and statisticians	0.88	82525	Computer and Mathematical Occupations
69	Physicists and astronomists	0.88	100500	Life and Physical Science Occupations
73	Chemists	0.91	75000	Life and Physical Science Occupations
74	Atmospeheric and space scientists	0.84	82009.5	Life and Physical Science Occupations
75	Geologists	0.94	75000	Life and Physical Science Occupations
76	Physical scientists, n.e.c.	0.98	87683	Life and Physical Science Occupations
77	Agricultural and food scientists	0.78	56959.5	Life and Physical Science Occupations
78	Biological scientists	0.96	58951	Life and Physical Science Occupations
79	Foresters and conservations scientists	0.86	55902	Life and Physical Science Occupations
83	Medical scientists	0.98	89079	Life and Physical Science Occupations
84	Physicians	0.99	220000	Health Diagnosing Occupations
85	Dentists	1.00	152460	Health Diagnosing Occupations

Code	Occupation	% with College Degree	Median Male Wage Income	Broad Occupation Category
86	Veterinarians	0.99	96967	Health Diagnosing Occupations
87	Optometrists	1.00	101640	Health Diagnosing Occupations
88	Podiatrists	1.00	120346.5	Health Diagnosing Occupations
89	Other health and therapy occupations	0.94	71148	Health Diagnosing Occupations
95	Registered nurses	0.53	70000	Health Assessment and Treating Occupations
96	Pharmacists	0.98	117598	Health Assessment and Treating Occupations
97	Dieticians and nutritionists	0.75	54943	Health Assessment and Treating Occupations
98	Respiratory therapists	0.25	60000	Health Assessment and Treating Occupations
99	Occupational Therapists	0.90	71148	Health Assessment and Treating Occupations
103	Physical therapists	0.90	76336	Health Assessment and Treating Occupations
104	Speech therapists	0.99	70000	Health Assessment and Treating Occupations
105	Therapists, n.e.c	0.79	50820	Health Assessment and Treating Occupations
106	Physicians' assistants	0.75	97000	Health Assessment and Treating Occupations
154	Subject instructors, college	0.94	68041.5	College instructors
155	Kindergarten and earlier school teachers	0.47	37000	Teaching (except college) and Library Occupations
156	Primary school teachers	0.95	51578	Teaching (except college) and Library Occupations
157	Secondary school teachers	0.96	52853	Teaching (except college) and Library Occupations
158	Special education teachers	0.87	50820	Teaching (except college) and Library Occupations
159	Teachers, n.e.c	0.62	55705	Teaching (except college) and Library Occupations
163	Vocational and educational counselors	0.79	43705	Teaching (except college) and Library Occupations
164	Librarians	0.85	52000	Teaching (except college) and Library Occupations
165	Archivists and curators	0.82	54000	Teaching (except college) and Library Occupations
166	Economists, market and survey researchers	0.79	88000	Economists and other Social Science Occupations
167	Psychologists	0.99	75000	Economists and other Social Science Occupations
169	Social scientists and sociologists, n.e.c.	0.89	62000	Economists and other Social Science Occupations
173	Urban and regional planners	0.93	72165	Economists and other Social Science Occupations
174	Social workers	0.77	43705	Social and Religious Workers
176	Cleargy and religious workers	0.73	44048	Social and Religious Workers
177	Welfare service workers	0.53	45738	Social and Religious Workers
178	Lawyers and judges	0.98	128946	Lawyers
183	Writers and authors	0.84	62001	Arts, Design, Entertainment, Sports, and Media
184	Techinical writers	0.74	69115	Arts, Design, Entertainment, Sports, and Media
185	Designers	0.52	55902	Arts, Design, Entertainment, Sports, and Media
186	Musicians and composers	0.52	49515	Arts, Design, Entertainment, Sports, and Media
187	Actors, directors, and producers	0.69	67083	Arts, Design, Entertainment, Sports, and Media
188	Painters, sculptors, craft-artists, and print-makers	0.54	58000	Arts, Design, Entertainment, Sports, and Media
189	Photographers	0.48	46420	Arts, Design, Entertainment, Sports, and Media
194	Art/entertainment performers and related occs	0.43	47226	Arts, Design, Entertainment, Sports, and Media
195	Editors and Reporters	0.80	58799	Arts, Design, Entertainment, Sports, and Media
198	Announcers	0.39	45369	Arts, Design, Entertainment, Sports, and Media
199	Athletes, sports instructors, and officials	0.60	50820	Arts, Design, Entertainment, Sports, and Media
226	Airplane pilots and navigators	0.75	92841	Technicians, Paralegals and Pilots
227	Air traffic controllers	0.35	90000	Technicians, Paralegals and Pilots
229	Computer software developers	0.73	87683	Computer and Mathematical Occupations
234	Legal Assistants and paralegals	0.41	55000	Technicians, Paralegals and Pilots
253	Insurance sales occupations	0.46	68099	Skilled Sales Occupations
254	Real estate sales occupations	0.46	60000	Skilled Sales Occupations
255	Financial services sales occupations	0.69	101640	Skilled Sales Occupations
256	Advertising and related sales jobs	0.54	64033	Skilled Sales Occupations
258	Sales engineers	0.66	96558	Skilled Sales Occupations
274	Salespersons, n.e.c	0.46	67052	Skilled Sales Occupations
418	Police and Detectives Public Service	0.37	62000	Police and Detectives Public Service

Note: The occupation level statistics are for 2010.

Appendix Table 2. Broad Occupation Characteristics for Fulltime Workers

	Measure of Returns to Working Long Hours																
	Elasticity of Annual Income to Weekly Hours - Males						Elasticity of Annual Income to Weekly Hours - Females+Males					Hourly Wage Premium for Working 50+ hrs- Males				
	Continuous Variable				Terciles		Continuous Variable				Terciles	Continuous Variable				Terciles
Broad Occupational Category	1980	1990	2000	2010	1980	2010	1980	1990	2000	2010	1980	1980	1990	2000	2010	1980
Executive, Administrative, and Managerial Occs	0.36	0.61	0.83	0.75	Mid	Mid	0.45	0.69	0.89	0.82	Mid	-0.15	-0.07	-0.01	-0.03	Mid
Business and Financial Operations Occs	0.66	0.88	1.11	1.19	High	High	0.70	0.89	1.11	1.18	High	-0.08	0.00	0.06	0.08	High
Architects and Engineers	0.35	0.43	0.47	0.42	Mid	Mid	0.35	0.41	0.47	0.42	Mid	-0.18	-0.14	-0.12	-0.15	Mid
Computer and Mathematical Occupations	0.36	0.47	0.68	0.70	Mid	Mid	0.33	0.42	0.65	0.68	Mid	-0.19	-0.14	-0.07	-0.07	Mid
Life and Physical Science Occupations	0.24	0.23	0.34	0.22	Low	Low	0.25	0.20	0.29	0.25	Low	-0.22	-0.21	-0.18	-0.23	Low
Health Diagnosing Occupations	-0.08	0.14	0.27	0.22	Low	Low	-0.09	0.09	0.24	0.12	Low	-0.35	-0.23	-0.18	-0.18	Low
Health Assessment and Treating Occupations	0.39	0.53	0.58	0.60	Mid	Mid	0.34	0.44	0.42	0.38	Mid	-0.19	-0.15	-0.14	-0.12	Mid
College Instructors	0.22	0.25	0.43	0.42	Low	Mid	0.22	0.24	0.39	0.34	Low	-0.22	-0.22	-0.16	-0.18	Low
Teaching (except college) and Library Occs	0.04	0.11	0.20	0.23	Low	Low	-0.03	0.09	0.18	0.16	Low	-0.28	-0.26	-0.23	-0.23	Low
Economists and other Social Science Occs	0.64	0.87	0.95	0.96	High	High	0.62	0.83	0.80	0.91	High	-0.09	0.00	0.01	0.02	High
Social and Religious Workers	-0.13	0.05	0.07	0.10	Low	Low	-0.24	-0.09	-0.04	0.02	Low	-0.38	-0.30	-0.29	-0.26	Low
Lawyers	0.39	0.85	0.93	1.11	Mid	High	0.44	0.87	0.95	1.09	Mid	-0.15	0.00	0.02	0.08	Mid
Arts, Design, Entertainment, Sports, and Media	0.53	0.64	0.78	0.76	High	Mid	0.56	0.64	0.76	0.77	High	-0.14	-0.09	-0.05	-0.05	High
Technicians, Paralegals and Pilots	0.51	0.39	0.51	0.38	High	Mid	0.94	0.59	0.77	0.64	High	-0.06	-0.19	-0.18	-0.20	High
Skilled Sales Occupations	0.37	0.59	0.87	0.94	Mid	High	0.45	0.64	0.91	0.99	High	-0.16	-0.08	-0.01	0.02	Mid
Police and Detectives	0.43	0.43	0.47	0.42	High	Mid	0.44	0.44	0.48	0.44	Mid	-0.11	-0.14	-0.11	-0.10	High

Source: 1980, 1990, and 2000 Census and 3-year aggregate 2011 ACS (2009-2011). See text for details on how the variables were constructed.

Appendix Table 3. Current and Future Returns to Working Long Hours - Estimates from SIPP

Broad Occupational Category	Elasticity of Current Annual Income with respect to Weekly Hours in Wave 1	Elasticity of Annual Income 3 to 4 Years Later with respect to Weekly Hours in Wave 1, conditional on current weekly hours (interacted with wave 1 occupations)
	(1)	(2)
Executive, Administrative, and Managerial Occs	0.65	0.25
Business and Financial Operations Occs	0.79	0.34
Architects and Engineers	0.54	0.25
Computer and Mathematical Occs	0.75	0.17
Life and Physical Science Occupations	0.03	-0.12
Health Diagnosing Occupations	-0.38	-0.47
Health Assessment and Treating Occupations	0.76	0.30
College Instructors	0.71	0.40
Occupations	0.24	0.08
Economists and other Social Science Occs	1.15	0.48
Social and Religious Workers	0.14	0.07
Lawyers	0.57	-0.20
Arts, Design, Entertainment, Sports, and Media	0.86	0.09
Technicians, Paralegals and Pilots	0.61	0.42
Skilled Sales Occupations	0.56	0.08
Police and Detectives	0.68	0.52
Rank correlation between Col(1) and Col (2)		0.69***

Note. The data is from wave 1 and wave 12 of the 1996, 2004, and 2008 SIPP Panels. The sample is restricted to full-time (35 or more hours worked per week) male workers with non-missing earnings data in wave 1 and wave 12. The number of observations used to estimate each regression is 5,115. The first column reports estimates of the elasticity of annual income in wave 1 with respect to weekly hours in wave 1 for males each of the broad occupations following the usual procedure as detailed in the text. The second column reports estimates of the elasticity of annual income in wave 12 (between 3 to 4 years later) with respect to weekly hours in wave 1 for males who reported being in each of the broad occupations in wave 1, controlling for weekly hours worked in wave 12 interacted with wave 1 occupation dummies (see equation (1')). All regressions include controls for the calendar year that each individual is sampled. ***significant at the 1% level, **5% level, *10% level.

Appendix Table 4. Cross-occupation Correlation between Returns to Working Long Hours and Participation Outcomes by Year

	Outcome:							
	Labor Force Participation				Share Female			
	1980	1990	2000	2010	1980	1990	2000	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns to working long hours	-0.004	-0.005	-0.004	0.017	0.052	-0.158	-0.238*	-0.090
	[0.025]	[0.013]	[0.011]	[0.018]	[0.246]	[0.196]	[0.133]	[0.117]
No. Obsv.	86	94	95	95	86	94	95	95

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). All regressions are weighted by the sample size of the cell (occupation x year). The unit of observation is an occupation. See Table 1 and the text for details on the outcome and explanatory variables. Robust standard errors are reported in brackets. ***significant at 1%, **5%, *10% level.

Appendix Table 5. Correlations between Returns to Long Hours, Competitive Index, and the Gender Pay Gap: Alternative Measure of Returns to Long Hours (2010)

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Outcome: Hourly Wage Premium for Working 50+hrs					
Competitive Index	0.081***	0.085***				
	[0.018]	[0.019]				
	B. Outcome: Gender Gap in Hourly Wage					
Hourly Wage Premium for Working 50+ hrs	-0.332***	-0.189**			-0.153	-0.135
	[0.099]	[0.087]			[0.133]	[0.095]
Competitive Index			-0.055***	-0.030***	-0.043***	-0.018*
			[0.008]	[0.010]	[0.015]	[0.009]
Controls		X		X		X
No. Observations	94	93	94	93	94	93

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The unit of observation is an occupation. The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. Please see the text for details on the construction of the variables. The controls are the share of full time workers in the occupation who are female, the labor force participation of women that reported being in that occupation, log of the average wage of males working exactly 40 hours, and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size in the regressions. Robust standard errors are reported in brackets. ***significant at 1%, **5%, *10% level.

Appendix Table 6. List of MSAs

Albany-Schenectady-Troy, NY	Miami-Hialeah, FL
Albuquerque, NM	Milwaukee, WI
Atlanta, GA	Minneapolis-St. Paul, MN
Austin, TX	Nashville, TN
Baltimore, MD	New Orleans, LA
Birmingham, AL	New York-Northeastern NJ
Boston, MA	Nassau Co, NY
Buffalo-Niagara Falls, NY	Newark, NJ
Charlotte-Gastonia-Rock Hill, SC	Oklahoma City, OK
Chicago-Gary-Lake, IL	Orlando, FL
Cincinnati OH/KY/IN	Philadelphia, PA/NJ
Cleveland, OH	Phoenix, AZ
Columbus, OH	Pittsburgh-Beaver Valley, PA
Dallas-Fort Worth, TX	Portland-Vancouver, OR
Dayton-Springfield, OH	Richmond-Petersburg, VA
Denver-Boulder-Longmont, CO	Rochester, NY
Detroit, MI	Sacramento, CA
Fort Lauderdale-Hollywood-Pompano Beach	St. Louis, MO-IL
Greensboro-Winston Salem-High Point, NC	Salt Lake City-Ogden, UT
Hartford-Bristol-Middleton-New Britain,	San Antonio, TX
Honolulu, HI	San Diego, CA
Houston-Brazoria, TX	San Francisco-Oakland-Vallejo, CA
Indianapolis, IN	San Jose, CA
Kansas City, MO-KS	Seattle-Everett, WA
Knoxville, TN	Syracuse, NY
Los Angeles-Long Beach, CA	Tampa-St. Petersburg-Clearwater, FL
Louisville, KY/IN	Tucson, AZ
Madison, WI	Tulsa, OK
Memphis, TN/AR/MS	Washington, DC/MD/VA
	West Palm Beach-Boca Raton-Delray Beach

Appendix Table 7: Variation in the Gender Pay Gap across Cities, Occupations, and Time

	1980		1990		2000		2009-2011	
	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev
	A. Full Sample							
Gender Pay Gap	-0.22	0.12	-0.18	0.11	-0.17	0.11	-0.14	0.11
	B. By Cities							
25th Percentile City	-0.28	0.16	-0.22	0.09	-0.20	0.11	-0.19	0.10
Median City	-0.23	0.15	-0.18	0.09	-0.17	0.08	-0.15	0.13
75th Percentile City	-0.20	0.09	-0.16	0.10	-0.15	0.08	-0.12	0.15
	C. By Occupation Group							
Executive, Administrative, and Managerial Occs	-0.40	0.06	-0.32	0.06	-0.28	0.06	-0.23	0.05
Business and Financial Operations Occupations	-0.26	0.06	-0.21	0.05	-0.20	0.06	-0.21	0.07
Architects and Engineers	-0.14	0.06	-0.11	0.04	-0.13	0.07	-0.10	0.07
Computer and Mathematical Occupations	-0.12	0.05	-0.09	0.06	-0.13	0.05	-0.13	0.07
Life and Physical Science Occupations	-0.20	0.06	-0.13	0.07	-0.12	0.09	-0.12	0.11
Health Diagnosing Occupations	-0.27	0.07	-0.25	0.12	-0.25	0.13	-0.22	0.13
Health Assessment and Treating Occupations	-0.19	0.08	-0.17	0.07	-0.20	0.08	-0.14	0.11
College Instructors	-0.20	0.06	-0.15	0.09	-0.14	0.07	-0.08	0.10
Teaching (except college) and Library Occupations	-0.12	0.04	-0.10	0.05	-0.08	0.04	-0.07	0.04
Economists and other Social Science Occupations	-0.24	0.08	-0.22	0.10	-0.21	0.12	-0.17	0.10
Social and Religious Workers	-0.07	0.07	-0.05	0.08	-0.06	0.08	-0.05	0.11
Lawyers	-0.18	0.07	-0.20	0.11	-0.18	0.11	-0.14	0.11
Arts, Design, Entertainment, Sports, and Media	-0.23	0.08	-0.16	0.10	-0.11	0.08	-0.11	0.09
Technicians, Paralegals and Pilots [a]	-	-	-0.32	0.22	-0.30	0.15	-0.27	0.18
Skilled Sales Occupations	-0.37	0.08	-0.23	0.08	-0.20	0.10	-0.18	0.11

Note. The data is from the 1980, 1990, and 2000 Census and 3-year aggregate 2011 ACS (2009-2011). Each cell is the mean gender pay gap for the full sample (Panel A), by city (Panel B) or by occupation group (Panel C) for each year. The "25th percentile" city is defined based on the average outcome across the occupation groups for each city in each year. The "median" and "75th percentile" cities are defined similarly. See text for details on how the variables were constructed.

[a] There are no values for Technicians, paralegals, and pilots for 1980 due to low number of observations. as mentioned in the text, we exclude observations at the city x occupation x year if there are fewer than 30 females and 30 males.

Appendix Table 8. Alternative Outcomes

	Outcome							
	Share Female				Female LFP			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns Long Hours$_{j,\,1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$	-0.005 [0.031]	-0.006 [0.030]			-0.014* [0.008]	-0.014* [0.008]		
Top tercile$_{j,\,1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$			-0.005 [0.013]	-0.006 [0.013]			-0.009 [0.005]	-0.009 [0.005]
Middle tercile$_{j,\,1980}$ x Ln(Predicted LS Immigrant Flow)$_{ct}$			0.011 [0.011]	0.010 [0.011]			-0.001 [0.003]	-0.001 [0.003]
Controls								
Year, City, Occ FE	X	X	X	X	X	X	X	X
Year*City FE	X	X	X	X	X	X	X	X
Year*Occ FE	X	X	X	X	X	X	X	X
City*Occ FE	X	X	X	X	X	X	X	X
Other Controls		X		X		X		X
Observations	2,348	2,346	2,348	2,346	2,348	2,346	2,348	2,346

Note. The data is from the 1980, 1990, 2000 Census and 2011 3-year aggregate ACS (2009-2011). The sample is restricted to native-born age 25-64 with at least a bachelor's degree who report working full-time (35 hours or more) in a given week. The unit of observation is an occupation-group*MSA*year. There are 59 MSAs, 15 occupation groups and 4 time periods. "Returns Long Hours" is the elasticity of annual income to weekly hours for the broad occupation group in 1980 obtained from the estimation of equation (1). See text for details on the construction of the Ln(Predicted LS Immigration) flows and for the gender gaps. The controls are the log of the average wage of males working exactly 40 hours and the share of males working 50+ hours a week. Observations are weighted by the cell's sample size. Robust standard errors are two-way clustered at the MSA and broad occupation levels and are shown in brackets.***significant at 1%, **5%, *10% level.