NBER WORKING PAPER SERIES

THE IMPACT OF SELECTION INTO THE LABOR FORCE ON THE GENDER WAGE GAP

Francine D. Blau
Lawrence M. Kahn
Nikolai Boboshko
Matthew L. Comey

Working Paper 28855
http://www.nber.org/papers/w28855

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021

We thank Isaac Cohen for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Impact of Selection into the Labor Force on the Gender Wage Gap
Francine D. Blau, Lawrence M. Kahn, Nikolai Boboshko, and Matthew L. Comey
NBER Working Paper No. 28855
May 2021
JEL No. C21,C24,J16,J3,J31,J71

ABSTRACT

We study the impact of selection bias on estimates of the gender pay gap, focusing on whether the gender pay gap has fallen since 1981. Previous research has found divergent results across techniques, identification strategies, data sets, and time periods. Using Michigan Panel Study of Income Dynamics data and a number of different identification strategies, we find robust evidence that, after controlling for selection, there were large declines in the raw and the unexplained gender wage gaps over the 1981-2015 period. Under our preferred method of accounting for selection, we find that the raw median wage gap declined by 0.378 log points, while the median unexplained gap declined by a more modest but still substantial 0.204 log points. These declines are larger than estimates that do not account for selection. Our results suggest that women's relative wage offers have increased over this period, even after controlling for their measured covariates, including education and actual labor market experience. However, we note that substantial gender wage gaps remain. In 2015, at the median, the selectivity-corrected gaps were 0.242 log points (raw gap) and 0.206 log points (unexplained gap).

Francine D. Blau
ILR School
Cornell University
268 Ives Hall
Ithaca, New York 14853-3901
and NBER
fdb4@cornell.edu

Lawrence M. Kahn
ILR School
Cornell University
258 Ives Hall
Ithaca, NY 14853
lmk12@cornell.edu

Nikolai Boboshko
Cornell University
2250 N. Triphammer Rd.
Apt U3A
Ithaca, New 14850
nb562@cornell.edu

Matthew L. Comey
Department of Economics
Cornell University
443 Uris Hall
Ithaca, NY 14853
mlc349@cornell.edu

I. **Introduction**

In one of the more dramatic labor market developments in the last several decades, the observed gender pay gap in the United States has decreased substantially. Yet, there is considerable debate about whether the convergence in the observed gap also represents convergence in wage offers between men and women. The key problem is that we only observe wages among the employed, a group that self-selects into the labor force (Heckman 1979). With a self-selected sample, there is the possibility that the convergence in the observed gender wage gap does not accurately measure convergence in wage offers due to non-random selection into employment. The direction of this bias is unclear. For example, the recent pattern of labor force exit among low-skill and likely low-wage men could artificially inflate the gender wage gap, attenuating our estimates of convergence. Alternatively, the entry of high-skill and likely high-wage women into the labor market could decrease the observed gender wage gap and exaggerate the extent of convergence. Shifts in selection patterns over time are an especially salient concern, since the labor force participation rates of men and women have converged at the same time as the gender wage gap has declined. In this paper, we attempt to account for selection bias using multiple methods to understand what has happened to the gender gap in wage offers.

Previous papers in the literature have found divergent results across techniques, identification strategies, data sets, and time periods. Consequently, as shown by the summary of the main findings from this literature in Appendix Table A1, there is no consensus about selection-adjusted trends in the gender wage gap. Among the methods employed to study selection-corrected convergence of the gender pay gap are structural modeling of selection into the labor force (Blau and Beller 1988; Mulligan and Rubinstein 2008; Maasoumi and Wang 2019); limiting to a sample men and women who have a very high employment probability and for whom selection is unlikely to drive wage gaps—the identification at infinity approach (Mulligan and Rubinstein 2008; Machado 2017); bounding techniques (Blundell, et al. 2007); and imputing unobserved wage offers by assigning wages relative to the median based on observables and employing median regression (Blau and Kahn 2006). On the extreme end of recent findings, Mulligan and Rubinstein (2008) find no evidence of a closing of the gender wage gap between the 1970s and 1990s once selection and covariates are taken into account. In contrast, for example, Blau and Kahn (2006) found that the selection- and covariate-adjusted median wage gap fell substantially between 1980 and 1998

While all of the above methods provide a possible solution to the selection problem, each method has important drawbacks. Structural methods make strong identifying assumptions about variables that affect employment but not wages. Techniques that impute individual wages in relation to the median make a selection on observables assumption that might not hold. The identification at infinity approach, while internally valid, is based on a highly selected sample that may be unrepresentative of the general population, raising external validity concerns. Bounding techniques require fewer assumptions but can easily generate uninformative estimates.

In this paper, we advocate for a data-driven approach to correct for selection, incorporating actual wage observations for individuals who are either not currently in the labor force or not members of wage samples traditionally used in analyses of the gender wage gap. We provide evidence on the raw and unexplained (i.e. covariate-adjusted) gender pay gaps in wage offers across time by using a variety of techniques to adjust for selection, while always limiting the scope for selection bias using our bolstered wage samples. To capture as many wage observations as possible, we use the Michigan Panel Study of Income Dynamics (PSID) to obtain wages for full-time workers with at least 26 weeks of employment, but also part-time workers and those with less than 26 weeks of employment. In addition, following Neal (2004), we employ adjacent survey years to obtain wages for individuals who do not have observed wages in the focal year. Together, these additional wage data create a substantial coverage improvement compared to more traditional wage samples used in wage gap analyses, limiting the ability for selection bias to affect the estimates.

For individuals who still lack wage data after exploiting these features of the PSID — never more than 10 percent of prime-aged men or 20 percent of prime-aged women — our preferred approach is to impute wages. To reduce the potential for bias stemming from these imputations, we focus on the estimation of the wage gap at the median, but we also study mean wage gaps and gaps at other quantiles of the wage distribution. We impute wages using what we call the "probability weighting method." Extending insights from Olivetti and Petrongolo (2008), instead of imputing specific wages we assign individuals without observed wages a set of predicted probabilities of their wages falling into each wage decile based on their observed characteristics. We use the unconditional quantile regression approach of Firpo, Fortin and Lemieux (2009) to compute wage differentials at the median and Oaxaca-Blinder-style decomposition (Oaxaca 1973; Blinder 1973) to identify the portions of the differentials explained and unexplained by covariates.

The probability weighting method has a number of advantages compared to other estimation strategies that have been employed to study our question and which are described in greater detail below. It avoids the strong identifying assumptions required to implement structural methods (Heckman 1979) and produces results that are more stable and less sensitive to small alterations in specification. The method also yields more precise estimates than bounding techniques (Blundell, et al. 2007), although it does require more assumptions in order to do so. Unlike the identification at infinity approach (Chamberlain 1986 and Heckman 1990), it produces results for a sample that is representative of the full population. Moreover, as noted by Olivetti and Petrongolo (2008) for their median approach, our probability weighting method does not assume positive selection, unlike some earlier research (Bayer and Charles 2018; Neal and Johnson 1996). And it does not require an assumption about the exact level of an individual's wage but rather their probabilities of being in particular portions of the distribution (Olivetti and Petrongolo 2008). While our approach uses observables to predict these probabilities, as discussed below, we implement some robustness checks that address the issue of unobservables. Although we believe that this approach has considerable merit, we also implement the other major methods used in the literature to correct for selection, including structural models, bounding techniques, and identification at infinity.

We make several contributions. First, we study the gender pay gap by implementing all the major proposed selection-correction methods using a consistent data set and period. Previous differences in conclusions about convergence in the gender wage gap could be due to different techniques but also due to a myriad of choices that researchers make. Second, for each of the methods, we leverage the longitudinal nature of the PSID to substantially improve wage coverage as described above. Thus, while nearly every method that we use makes assumptions, using additional actual wage data reduces the impact that those assumptions make. Indeed, we find that as wage coverage increases, all the methods converge to the same result. Third, in all our models we account for actual experience, an important variable in explaining the gender wage gap given the greater likelihood of workforce interruptions among women than men (Mincer and Polachek 1974; O'Neill and Polachek 1993; Blau and Kahn 1997) and a likely source of omitted variable bias in analyses using the Current Population Survey (CPS) or the

American Community Survey (ACS), which do not include information on actual experience. Moreover, the gender experience gap has declined sharply over the period, making this variable important in analyses of trends (Blau and Kahn 2017). Fourth, in every method we employ, we correct for sample selection of males as well as females, as is now appropriate given labor force participation trends for both groups. As a final contribution, our preferred probability weighting method implements a richer version of previous median approaches (Neal 2004, Blau and Kahn 2006, and Olivetti and Petrongolo 2008) by estimating predicted wage deciles rather than simple above and below median probabilities.

We find that, after controlling for selection, there are large declines in the raw and the unexplained gender wage gaps over the 1981 to 2015 period. This conclusion largely holds up across all of the methods for correcting for selectivity bias and is similar to what one would have concluded without such corrections. In our preferred specification using the probability weighting approach, the raw median gap declined by 0.378 log points, while the unexplained median gap declined by a more modest but still substantial 0.204 log points. These declines are larger than estimates that do not account for selection. These results are highly robust to a range of alternative specifications, including varying the variables used to predict individual wages, only using individuals with weak labor force attachment to predict wages (as they may be more similar to those without wages on unobserved factors), and finally to adjusting our results based on possible imputation errors. Thus, overall, our results suggest that women's relative wage offers have increased over this period, even controlling for their measured covariates, including education and actual labor market experience. However, we note that substantial gender wage gaps remain. In 2015, at the median, the selectivity-corrected gaps were 0.242 log points (raw gap) and 0.206 log points (unexplained gap).

II. Related Literature and Our Contribution

Blau and Kahn (2017) present the trends in the gender wage gap for the 1980 to 2010 period based on observed wages and Oaxaca-Blinder-style decompositions of OLS regressions (Oaxaca 1973, Blinder 1973). Using PSID data, and consistent with other evidence, they report convergence in the raw gender wage gap over the 1980 to 2010 period, with substantial narrowing in the 1980s and slower convergence thereafter. They find that improvement in women's relative characteristics played a significant role in the decrease, but that much of the substantial convergence during the 1980s was attributable to a fall in the unexplained gender pay gap. However, there was no evidence of subsequent decreases in this measure after the 1980s (see also Blau and Kahn 2006). Blau and Kahn (2017) also explore trends in the gender wage gap at the 10^{th}, 50^{th}, and 90^{th} quantiles of the male and female wage distributions, finding slower convergence at the top, both in raw and adjusted gender gaps.

As discussed above, adjustment for changing patterns in the selection of women and men into the labor force could alter the conclusions based on simple OLS analyses of observed wages. (Again, see Appendix Table A1 for a summary of findings.) The traditional method to account for selection in wage equations uses structural methods initiated by Heckman (1979) where the correlation between sample selection and wages is explicitly modeled and estimated, typically with the use of an excluded instrument that affects employment probability but not the wage itself. This technique has been employed to study trends in the gender wage gap by Blau and Beller (1988); and Mulligan and Rubinstein (2008). Excluded instruments used include non-labor income and age distribution of household members (Blau and Beller 1988), and marital status interacted with number of young children (Mulligan and Rubinstein 2008). Notably, Mulligan and Rubinstein find a complete lack of convergence in the selection corrected unexplained wage gap between the late 1970s and late 1990s, while Blau and Beller (1988) find selection-corrected convergence in the 1970s.

Arellano and Bonhomme (2017) propose a selection correction method in a similar vein to the Heckman (1979) approach that is applicable to quantile regression models. Maasoumi and Wang (2019) employ this method to find broad convergence in the selection-corrected gender wage gap at the mean and various quantiles, using the number of young children as an excluded instrument.[1] However, their year-by-year selection corrected results are highly volatile.

Structural methods for selection bias correction provide estimates of structural parameters that are of interest in themselves. However, potential drawbacks of the structural method for selection bias correction are that the excluded instruments may not be valid and that the results may be sensitive to small changes in specification. Instead of making the arguably strong identifying assumptions needed in order to estimate a structural model, some authors have made a priori assumptions about where workers without wages would place in the wage distribution. For example, some assume that those without wages would place below the median, a perhaps reasonable assumption for men, but less so for women. For examples of this, see Neal and Johnson (1996) and Bayer and Charles (2018), who study black-white gaps among males. It also may be valid to assume that individuals with high levels of education and experience would place above the median, as do Neal (2004) in his study of black-white gaps among females, Blau and Kahn (2006) in their study of male-female differences over the 1979 to 1998 period, and Olivetti and Petrongolo (2008) in some specifications of their study of international differences

[1] Maasoumi and Wang (2019) note that when they only correct for female selectivity, they find little evidence of convergence; however, when they correct for male and female selectivity, some evidence of convergence emerges. Also of note, they find that selection-corrected results show a substantial *increase* in the gender wage gap at the 10th and 25th percentiles during the 1990s and 2000s.

in the gender wage gap. Blau and Kahn (2006) find that this approach generally leads to similar conclusions about convergence patterns in the U.S. gender wage gap at the median as one would conclude from observed wages, but in each wave the inclusion of individuals who do not have observed wages substantially increases the unexplained portion of the median gap (and hence the total gap).

A modification of this method is to estimate the probability that a person without wages would place above the median based on where observed workers with similar characteristics place and impute wages on a probabilistic basis (Olivetti and Petrongolo 2008). Using the PSID and the European Community Household Panel, Olivetti and Petrongolo (2008) also generally find in their cross-sectional data that estimates of the median wage gap increase as more wage observations are added. In this approach, one need not make a priori assumptions about where individuals with unobserved wages would place in the wage distribution. While this method does not require the perhaps strong identifying assumptions needed for more structural approaches, it does not fully address the issue of sample selection bias based on unobservable factors. However, it is less demanding in that it involves placing individuals without observed wages above or below the median with a certain probability, rather than assigning them an exact wage. As discussed above, we expand this approach by imputing the wages for the remaining individuals based on deciles of the wage distribution.

A method relying on even fewer assumptions is to estimate bounds on the distributions of male and female wage offers in the population, using methods devised by Blundell, et al. (2007), who were motivated by the theoretical work of Manski (1994) and Manski and Pepper (2000).[2] In the most extreme case, Blundell, et al. (2007) bound the “true” population wage distribution

[2] Olivetti and Petrongolo (2008) conduct a bounding exercise to compare the gender pay gap across countries at a point in time.

— often by subgroup — at various quantiles by first assuming all unobserved wages fall below a given quantile of interest in the observed wage distribution and then assuming all unobserved wages fall above the quantile of interest. They then show how to tighten these bounds by layering on assumptions about the wage distribution, including stochastic dominance of the observed wage distribution relative to the unobserved distribution and the existence of valid instrumental variables that affect employment but not wages (or at least affect wages in a monotonic fashion). Studying wages in the United Kingdom in 1978 and 1998, they find evidence of a closing of the gender wage gap among the less educated and younger segment of the population. While such analyses do not provide point estimates, if the bounds are relatively tight, we may still be able to make conclusions about the size and convergence of the gender pay gap.

A further approach in the literature at the intersection of the gender wage gap and selection is the identification at infinity method devised by Chamberlain (1986) and Heckman (1990) and employed by Mulligan and Rubinstein (2008). This method involves finding a segment of the population for which the probability of employment is particularly high — ideally approaching one — and estimating the wage gap among this group. For such a sample, there will not be a correlation between the unmeasured factors affecting employment and those affecting wages. Using this approach on CPS samples from the late 1970s and late 1990s, Mulligan and Rubinstein (2008) find almost no convergence in the gender wage gap among samples with high probability of employment (that is, at least 80 percent). While such methods may be internally valid, they may not yield results that are representative of the full population. For example, in Mulligan and Rubinstein's sample of women with high predicted employment probabilities (i.e. at least 80 percent), 76 percent were never married and fully 98 percent had an advanced degree.

As noted, research based on observed wages has found that the gender pay gap fell much more slowly at the 90th percentile than at the median (e.g., Blau and Kahn 2017), and women with an advanced degree are likely to be in the upper portion of the female wage distribution. Thus, the degree of gender convergence among this group may not be informative about the population in general.

Finally, Machado (2017) proposes a kind of hybrid of the structural approaches and the identification at infinity approach. She also analyzes a group with high labor-force attachment, looking at the "always employed" — a group analogous to "always takers" in the local average treatment effect literature — to estimate wage gaps on a group for which selection should not matter. The group is defined with respect to an instrument for employment; she uses the presence children under 6 as an IV. Hence her estimation of the gender wage gap pertains to employed women with children. Unlike Mulligan and Rubinstein (2008), she does find evidence of convergence.[3]

While we focus on the gender wage gap at the median, this sort of comparison recently faced scrutiny from Maasoumi and Wang (2019). They write: "Implicitly assumed and required in the quantile comparisons is an assumption of rank invariance (or similarity); that is, one's relative rank is preserved when endowed with each other's skill sets or market returns" (p. 2439). They claim there is no reason to expect rank invariance to hold with respect to the male and female wage distributions and reject the assumption using a test proposed by Frandsen and Lefgren (2018). While this concern is certainly important to consider, we believe the median wage gap to be an interesting and important descriptive statistic in itself: it is a measure of

[3] Machado (2017) also replicates the structural approach finding of Mulligan and Rubinstein (2008) to contrast her main results, and also finds evidence of convergence using a median bounding approach under a stochastic dominance assumption.

central tendency free of influence from extreme values. However, we also briefly examine other quantiles of the distribution.

As noted, our preferred approach is to use multiple waves of the PSID to obtain wage observations for those without wages in a given wave. Moreover, using our probability weighting approach, extending Olivetti and Petrongolo's (2008) median approach, we then assign decile probabilities for those still without wages in order to create a distribution of wages for the population. Further, unlike Olivetti and Petrongolo (2008), we address the issue of selection on unobservables in some of our robustness analyses. As noted in the Introduction, our approach avoids the strong identifying assumptions required to implement structural methods and does not impose positive selection. It also yields more precise estimates than bounding methods, albeit with more assumptions, for which we will provide some evidence. And, unlike the identification at infinity approach, it produces results for a sample that is representative of the full population.

III. Data and Summary Statistics

a. Data Sources

Our primary data source is the Panel Study of Income Dynamics (PSID). We compute gender wage gap estimates using the 1981, 1990, 1999, 2011, and 2015 survey waves, and we use additional waves to provide supplementary wage information as discussed below. These years were selected to illustrate trends in the gender pay gap during the previous four decades.[4] As noted, the PSID can give insight into what wages might be in a given year if an individual is

[4] The PSID was collected annually through 1997 and biennially thereafter. Note, while the years through 2011 correspond to those used by Blau and Kahn (2017), there are some differences between the samples employed here and in that study — for example, the age range used — so the OLS results are not expected to be identical.

currently out of the labor force, and can be used to construct actual labor market experience (as opposed to a measure of potential experience based on age and years of education). We restrict all analyses to individuals aged 25 through 54, allowing us to largely abstract from issues related to working while in school (on the low end) and working while partially retired (on the high end). We drop the PSID Latino and Immigrant samples as well as individuals for whom we do not have information on key wage and employment determinants, such as work experience and education.[5] All wages are converted to 2010 dollars using the GDP Personal Consumption Expenditures deflator and then log transformed. See the Data Appendix for additional sample construction details.

b. Wage Samples

As part of our main approach to correct for selection bias, we successively add individuals and corresponding wage information to a basic sample of full-time workers employed 26+ weeks so as to limit the incidence of unobserved wages. Specifically, we first add part-time workers and those with less than 26 weeks of employment, then remaining individuals with wage data in nearby PSID waves, and finally all other remaining individuals, for whom we impute wage decile probabilities. This yields one base sample and three progressively more inclusive wage samples for which we compute wage gaps. While each successive sample includes all the wage information of the prior sample, we will reference from time to time sample “entrants” to characterize features of individuals added to form the indicated sample.

The basic sample consists of workers who are currently employed full time and who worked at least 26 weeks during the previous year. This group is relatively homogenous with

[5] We drop the Latino and Immigrant samples because they do not exist for our entire sample period. The Latino sample exists only for 1990 to 1995. The immigrant sample was only added in the 1997 and 1999 PSID waves.

respect to commitment to the labor force, and we denote them as Sample 1. Next, we add those employed part time and those working less than 26 weeks, as long as they have worked at least 100 hours in the previous year.[6] We term the union of Sample 1 and this additional group Sample 2. While Sample 2 is likely to be less homogeneous than Sample 1 with respect to labor force commitment, it has the advantage of covering a larger portion of the population and of course still contains information on earnings. We believe that it is legitimate to include part-time workers, since previous research has found little evidence of a part-time wage penalty once selection is taken into account (Blank 1990; Hirsch 2005). The one exception is a penalty for men who are of retirement age, a group that we exclude from our sample (Aaronson and French, 2004).

We then augment Sample 2 by using the longitudinal nature of the PSID to find wages for individuals for whom we do not have a valid current wage, i.e., who do not satisfy the restrictions to be in Sample 2. Specifically, for such individuals, we assign wages from an adjacent wave of the PSID. Prior to 1997, the PSID was fielded annually and thereafter biennially. In the earlier years, we first assign wages from the surveys that are one year before or one year after the PSID sample wave, if available, and if wages are available in both years, we take the mean. If still no wage has been assigned, we repeat the process using wages two years before or after.[7] In years after 1997, we perform a similar procedure focusing solely on the adjacent surveys (which are two years apart.) We call the union of Sample 2 and these new individuals Sample 3.

[6] We use a 100 hour cutoff in order to increase our confidence in the computation of hourly earnings. Wage gap estimates are robust to completely eliminating the 100 hours restriction and to increasing the restriction to 300 hours.

[7] These additional individuals must still meet the 100 hours restriction in the year from which wage data are taken.

Finding additional wage information from nearby PSID waves may be informative, since we have actual wage observations on individuals rather than imputations. However, a potential drawback of this approach is that past or future wage offers may not be indicative of the current wage offer. For example, on the one hand, a person may have been dismissed from a high wage job, implying that the wage from one or two years ago may be an overestimate of a current wage offer. On the other hand, a wage offer from one or two years ago may be an underestimate of his/her current wage offer if those years coincided with additional human capital formation via education or training. Moreover, a prior or future wage may have resulted from different market conditions, again potentially compromising the accuracy of this method of retrieving wage offers. Hence, we tested several versions of this method, including assigning only past wages or only future wages to individuals without a current wage. Our main results are robust to such alternatives, suggesting that this method of finding wages for those without current jobs has merit. We note that a version of this technique was previously used by Neal (2004), Blau and Kahn (2006), and Olivetti and Petrongolo (2008).

Finally, for any individual still without wage data, we impute wages based on covariates using our probability weighting method. We call the union of Sample 3 and this final group Sample 4. We impute Sample 4 entrant wages in three steps: 1) We use Sample 3 to estimate an ordered probit model relating individual covariates to the probability of inclusion in each decile of the PSID wave-by-sex wage distribution; 2) We use the resulting model to assign each Sample 4 entrant a set of ten probability-weighted wage observations constructed from each entrant's set of predicted decile inclusion probabilities and a corresponding set of within-decile median wages; and 3) We iterate on steps 1 and 2 by updating the wage decile thresholds used for the ordered probit until the final probability-weighted decile assignment of Sample 4 entrants results in an approximate convergence in the decile thresholds. Specifically, this iteration procedure continues until the wage decile end points from the current iteration are each within 25 cents of the corresponding deciles from the previous iteration.[8]

We estimate decile probabilities (rather than, for example, the probability of being above median) in order to more accurately assess each Sample 4 entrant's likely placement in the resulting wage distribution. Since our primary estimation strategy involves estimating median wage gaps, if one were only interested in describing the gender wage gap at the median rather than estimating the effects of explanatory variables, it would not matter in which decile we placed a Sample 4 entrant conditional on being above or below the median. However, we are also interested in *ceteris paribus* gender pay gaps at the median. For such analyses it may matter how far above or below median an individual is, and for this reason, we assign individuals decile probabilities rather than merely probabilities of being above or below median.

In the wage decile prediction model, we include the following covariates: years of schooling, indicator variables for attainment of undergraduate and advanced degrees, quadratics in years of actual full-time and part-time experience, Census region indicators (omitting the Midwest), and race/ethnicity indicators (non-Hispanic white, non-Hispanic black, Hispanic, and omitting Non-Hispanic other races). We refer to this set of variables as our "human capital specification" and use it throughout the paper. From the estimated probit, Sample 4 entrants are assigned a probability of having a wage in each decile. As noted earlier, Sample 4 entrants are then included in the analysis sample ten times assigning a wage equal to the median of each decile. These ten observations are then reweighted using the predicted decile probabilities from

[8] Olivetti and Petrongolo (2008) use a similar iteration procedure to assign wages above or below the median.

the probit so as to keep each entrant's aggregate weight identical to his or her original sample weight. As noted, we iterate this procedure until the decile cutoffs approximately converge.

We note that this method of assigning Sample 4 entrants to wage deciles is based on observed characteristics. It is also possible that results could be sensitive to the exact specification of the ordered probit used to predict wages. In Section V, below, we test the robustness of our main results using a variety of alternative imputation specifications aimed at addressing these concerns. These checks include: varying the wage sample used to estimate the ordered probit (i.e. using Sample 3 entrants, who may be more similar to Sample 4 entrants on unobserved factors); varying the predictor variables used in the ordered probit; varying the exact within-decile wages used as imputation values; varying the number of wage quantiles used as outcomes in the ordered probit; and using Sample 3 entrants to estimate error rates for our wage assignment procedure, and then using those error rates to correct our imputations. We discuss these tests in more detail below, but our results are robust to each.

c. Summary Statistics

Table 1 presents basic sample size and wage summary information by PSID wave and wage sample (Samples 1-4). First, note the degree to which each successive wage sample increases the fraction of individuals with observed wages in each wave-by-sex PSID sample. This fraction is denoted in the table as coverage. In Sample 1, the most traditional of our wage samples, only between 74.4 and 86.1 percent of men and 42.6 and 60.1 percent of women are covered, depending on the PSID wave. Men's coverage was steady through 1999 and then fell sharply through 2011, likely reflecting the Great Recession and its aftermath. While coverage subsequently rebounded somewhat, it remained noticeably below its 1990s levels in 2015.

Women's coverage rose sharply through 1999, fell moderately through 2011, and mostly recovered by 2015. These data indicate that while inclusion in the most committed wage sample (Sample 1) is a larger issue for women than men, it has always affected both men and women to some extent. Moreover, due to opposing male-female employment trends, the gender gap in coverage between men and women has decreased substantially, declining from 42.9 percentage points in 1981 to 19.8 percentage points in 2015.

Adding part-time and workers with less than 26 weeks of employment for Sample 2 increases coverage substantially — by at least 4 percentage points per wave for men and at least 15 percentage points per wave for women. Adding individuals with wages in nearby waves to create Sample 3 provides another substantial increase in coverage — at least a 6 percentage point increase for men in each wave and at least a 13 percentage point increase for women in each wave. In Sample 3, male coverage exceeds 92 percent in each wave and female coverage exceeds 81 percent, dramatically reducing the scope for selection bias relative to traditional wage samples (e.g. Sample 1).

In each wave of the PSID, adding the additional wage observations increases the observed raw gender log wage gap at the median, defined as the male median log wage minus the female median log wage. Further, adding these wages increases the observed convergence in the raw gap with each successive sample. That is, the raw median gap declines by 0.297 log points from 1981 to 2015 in Sample 1, but by 0.326 log points in Sample 2, 0.363 log points in Sample 3, and 0.380 log points in Sample 4.[9] By comparing entrant median wages to incumbent median wages, we can see that there is positive selection into sample incumbency for both men

[9]Male and female mean and median log wages were very similar when we experimented with different methods of locating wages for Sample 3, including looking only two years backward or looking only two years forward, as opposed to our basic approach of finding the nearest wage regardless of whether it was in the past or the future.

and women, on average. For example, for Sample 2 entrant men, the median 2015 log wage was 2.636, while the median incumbent (i.e., Sample 1) log wage was 3.143. In fact, in every wave and for both sexes, the median entrant log wage is less than the incumbent median, a pattern that will be useful in assessing the degree of selection bias using a bounding approach (see below). In some instances, the entrant-incumbent gap is larger for men than women, while in other cases, the gap is larger for women. However, the overall gender gap in wages increases each time we expand the sample, reflecting the higher incidence of entrants among women than men.

Table 2 provides basic summary statistics on key human capital measures: education and full-time work experience. Similar to earlier research (Blau and Kahn 2017), Table 2 shows over our sample period female education catching up to and overtaking male education and the male-female experience gap sharply narrowing. The table indicates that this is the case in each of the four (increasingly inclusive) samples.

By studying education and experience among entrants relative to incumbents in the successive samples, we can assess the human capital characteristics of those not included in most standard wage analyses (such as Sample 2, 3, and 4 entrants) compared to incumbents. For women, the successive wage sample entrant groups are always less educated (with respect to years of schooling) than incumbents on average and nearly always so for men. A similar pattern characterizes full-time experience: for women, entrants always have less full-time experience than incumbents, while for men, this outcome occurs in almost every case. Also of note, women's levels of schooling and full-time experience relative to men's are lower the more inclusive the sample. Thus, just as the gender pay gap is higher for each successive sample (Table 1), so is the human capital gap. Our regression analyses will help us determine whether these human capital differences fully account for the larger gender pay gaps in the more inclusive samples.

IV. Primary Analysis of the Gender Wage Gap at the Median

a. Methods

Our primary analysis consists of estimating the median gender wage gap via unconditional median regression (Firpo, Fortin and Lemieux 2009 and 2018), and then implementing a Oaxaca-Blinder-style decomposition (Oaxaca 1973; Blinder 1973) to identify the sources of the differential.[10] We begin our analysis in 1981 because this is the first year in which data are available on actual experience for both household heads and partners. Fortunately, this is a reasonable starting point because 1980 marks the beginning of convergence in the raw gender wage gap in observed wages as measured in published government statistics (Blau and Winkler 2018).[11] We focus on the median because it is more robust than the mean to our imputation procedure, as described above. When we focus on the adjusted gender gap, the regression coefficients themselves may be sensitive to exactly how far above or below the median one is placed (i.e., the estimated decile probabilities), and we will examine the robustness of our procedures to different methods of imputation.[12] As with the standard Oaxaca-Blinder

[10] An alternative approach would be to estimate conditional quantile regressions and determine the *ceteris paribus* gender gap. However, our imputations involve placing people in the unconditional distribution of wages. We are more confident, for example, in saying that one is below the sample median than that one is below the median among people with his or her measured characteristics. Other authors using unconditional quantile regression techniques to study the gender wage gap include Kassenboehmer and Sinning (2014), Töpfer (2017); and Meara, Pastore, and Webster (2020).
[11] See Figure 7-2, p.173.
[12] For example, Firpo, Fortin and Lemieux (2009) show that the unconditional quantile regression coefficient at, say, the median, is the ratio of the derivative of the probability of being above median and the density of wages at the median. In implementing this technique, the denominator is approximated by a kernel density function, and this can clearly be affected by how far above or below the median an individual is placed. In robustness checks, we replace deciles with octiles, quartiles and halves of the distribution, and find very similar results.

decomposition of the mean gender wage gap, we can report the component of the median wage gap explained by observed characteristics and the component of the gap that is unexplained (i.e., the gap adjusted for observed characteristics). The latter is sometimes taken as an indicator of discrimination but, as is widely acknowledged, may also reflect the impact of unmeasured characteristics. Beyond the issue of measuring discrimination, it is of interest to know the extent to which observed characteristics account for the gender pay gap at a point in time, as well as any changes in the gender wage gap over time. For comparison purposes, we also present results using a similar approach applied to the mean gender wage gap; results are quite similar.

The Firpo, Fortin, and Lemieux (2009) unconditional quantile regression method is computed using a recentered influence function (RIF). At the median, this function is defined as:

$$RIF(Y \mid q_{0.5}, F_Y) = q_{0.5} + (0.5 - \mathbb{I}\{Y \leq q_{0.5}\}) \,/\, f_Y(q_{0.5}) \tag{1}$$

where $q_{0.5}$ is the median of random variable Y, $\mathbb{I}\{Y \leq q_{0.5}\}$ is an indicator function that is equal to one if Y is below the median, and $f_Y(q_{0.5})$ is the density of Y evaluated at the median. Firpo, Fortin and Lemieux (2009) show that unconditional quantile regression can be implemented by regressing the RIF on explanatory variables using ordinary least squares. Coefficients can then be interpreted as the effect of a change in the distribution of covariates on the median, though Firpo, Fortin and Lemieux (2018) caution this interpretation is an approximation that holds well only for small changes. In equation form, this regression is:

$$RIF(Y_i \mid q_{0.5}, F_Y) = \alpha + \beta X_i + \varepsilon_i \tag{2}$$

We calculate the RIFs separately for each PSID wave-by-sex grouping. The decomposition is then estimated within each PSID wave using the same human capital specification variables listed in the previous section, denoted here by X_i.[13] In practice, estimating the determinants of the numerator of the fraction in equation (1) is equivalent to estimating a probability (of being above median) equation, which we estimate separately for men and women. Combining these results with the denominator (estimated by kernel density methods) and with the male and female distributions of personal characteristics gives us the raw material for a Oaxaca-Blinder decomposition.

b. Results

Table 3A shows our estimates for the median wage gap and Oaxaca-Blinder-style decomposition using the RIF approach described above, successively for each wage sample. For comparison, we also present results for an ordinary least squares (OLS) version of the analysis along with the traditional mean Oaxaca-Blinder decomposition, and, in Table 3B, we repeat both analyses using potential experience instead of actual experience to show the importance of actual experience in the decomposition of the gap.[14] We focus on results for the median analysis but results for the mean are very similar. Unless otherwise noted, the discussion below refers to models using actual experience. Standard errors are computed using 599 iterations of a bootstrap procedure that randomly draws PSID observations by wave and recomputes all RIF computations and decomposition results in each iteration. Generally, standard errors are quite small compared to gap estimates and long-term convergence estimates (i.e., from 1981 to 2015), so we largely refrain from discussing inference in this section.

[13] Note that we do not include employment-related variables such as industry, occupation, or union status because these do not exist for those whose wages we impute and because they are endogenous with respect to human capital. We also do not include marital status or children because these are also likely endogenous.

[14] In the potential experience specification, wage imputations for Sample 4 use potential rather than actual experience.

As discussed, Sample 1 focuses on wages for full-time workers with substantial employment (at least 26 weeks) in the previous year. This is the sample most similar to that used in the bulk of research on the gender wage gap. In this sample, we replicate the finding of a declining raw (unadjusted) gender wage gap at both the mean and the median. The gap has fallen by about 0.28 log points at the mean and by about 0.30 log points at the median between 1981 and 2015, with the largest decrease occurring in the 1980s. Nonetheless, in 2015, there was still a substantial wage gap of 0.165 log points at the mean and 0.178 log points at the median. The explained gender gap (due to differences in measured characteristics) fell gradually over the 1981 to 2015 period, by 0.15 log points at the mean and 0.13 log points at the median, and has been statistically insignificant since 2011. The unexplained gender gap (adjusted for gender differences in measured characteristics) fell sharply in the 1980s — 0.12 log points at the mean and 0.15 log points at the median. At the mean, the unexplained gap subsequently decreased very little; at the median, it also fell substantially in the 1990s, but has since risen back to about the 1990 level.

As noted in our discussion of Table 1, the raw gender wage gap increases as we add more wage observations through Sample 2 (adding wages for part-time and <26 weeks), Sample 3 (adding wages from nearby PSID waves), and Sample 4 (adding imputed wages). We saw in Table 2 that the gender gap in human capital also rose across samples. The results in Table 3 indicate that these human capital differences across samples are important in accounting for differences across samples in the gender wage gap. We may see this by comparing the raw gender wage gap to the unexplained gap. The results indicate that the differences across samples are considerably less for the unexplained gap than for the raw gap. Indeed, by 2011 and 2015, the unexplained gap was very similar across the four samples. Thus, differences across samples in human capital are an important factor but, prior to 2011, do not fully account for the wage gap differences across samples.

Turning to our findings for Sample 4, which fully takes selection into account, we see the median wage gap falling by 0.38 log points between 1981 and 2015, larger than the convergence of 0.30 log points for Sample 1, settling at 0.24 in 2015 (compared to 0.18 log points for Sample 1). The unexplained gap also converged more in Sample 4 than in Sample 1 (0.20 log points compared to 0.17 log points), although the differences between Sample 4 and Sample 1 unexplained gap convergence are not statistically significant and smaller than for the raw gap. Regardless, our finding of substantial convergence in the unexplained wage gap after correcting for selection stands in stark contrast to Mulligan and Rubinstein (2008). Another finding is that, in Sample 4, the pace of convergence in the raw median is now virtually the same in the 1980s and the 1990s, although the unexplained gap fell more sharply in the 1980s.[15] As with the Sample 1 analysis, the explained portion of the gap fell gradually between 1981 and 2015 and is now close to 0 although still statistically significant.

Overall, our results suggest that selection does not account for the convergence of the measured gender wage gap among full-time workers with at least 26 weeks of employment over the 1981-2015 period and if anything understates it. This conclusion holds for both the raw gap and the unexplained gap, as we find stronger convergence for both after adjusting for selection. Moreover, our results suggest that selection does not account for the gender wage gap in any given year, but rather that selection results in the magnitude of the gender wage gap being

---

[15] Blau and Kahn (2006) also report faster convergence in both the raw and unexplained gender wage gaps in the 1990s after correction for selection. They suggest this may reflect the large entry of relatively low-skilled female single-family heads during this period.

understated, again for both the raw gap in all years and the unexplained gap in years prior to 2015.

This analysis of selection bias combines the effects of male selection and female selection, and it is instructive to separate out these two components. We focus on the unexplained gap because of its relevance to the role of unobservables. To do so, we first compute the unexplained gender gap in log median wages using men in Sample 1 (the selected sample of male full time workers with at least 26 weeks) and women in Sample 4 (the full population of women). We then compare this gap with the unexplained gap using both men and women from Sample 1. This comparison gives us the effect of *female* selection. The effect of *male* selection is then just the difference between the total selection effect and the female selection effect.[16]

The results of correcting for selection separately for women and men are shown in Figure 1. For comparison, the Figure also shows results when we do not correct for selection and when we correct for both male and female selection. Figure 1 indicates that both men and women were positively selected into Sample 1 in every year, with the exception of a near zero effect for men in 1999 (-0.005). This positive selection is indicated by the unexplained wage gap falling relative to the unadjusted gap when we only correct for male selection and rising when we only correct for female selection. The total selection effect is the net impact of positive female selection reducing the unexplained gap and positive male selection increasing it. Taking 1981 as an example, the uncorrected gap is 0.371 log points. When we control for only female selection, it rises to 0.441 log points, implying that positive female selection into Sample 1 lowers the

[16] This definition of the male selection effect is equal to the difference between the unexplained gap for the Sample 1 men vs Sample 4 women comparison and the unexplained gap for the Sample 4 men vs Sample 4 women comparison. We also used the alternative counterfactual gap (men from Sample 4 and women from Sample 1) to perform the decomposition, with similar results. Findings for mean regression were also similar.

unexplained gap by 0.070 log points. Men are also positively selected, with a selection effect of 0.031 (i.e., 0.441-0.410), working to increase the unexplained gap. The net effect of male and female selection, then, is to reduce the unexplained gap by 0.039 log points (i.e., 0.070-0.031). Performing a similar calculation for 2015 yields smaller selection effects for both men and women: a reduction in the unexplained gender pay gap of 0.022 log points due to female selection and an increase of 0.022 log points due to male selection, resulting in a zero overall selection effect. Women and men are both less positively selected in 2015 than in 1981, with a larger change for women.

Combining these male and female selection estimates for 1981 and 2015, we can calculate the role of female and male selection in the convergence of the unexplained gender pay gap. When we take account only of female selection, we find that the unexplained pay gap fell by 0.213 log points between 1981 and 2015. Recall that in Sample 1 the unexplained gender pay gap fell by 0.165 log points, while the unexplained gap in Sample 4 fell by 0.204 log points. Thus, changes in female selection are more than sufficient to account for the faster convergence obtained when we correct for both male and female selection vs. correcting for neither. The impact of the correction for male selection is actually to slightly reduce convergence.

Figure 1 is also helpful in studying trends in selection within each subperiod. Beginning with the 1980s, we see that both men and women became more positively selected, possibly due to rising returns to skills, with a slightly larger effect for women than men. The latter may be due the loss of some high wage jobs for men during a period of deindustrialization and deunionization. The net effect resulted in slightly slower convergence in the selection-corrected gap vs the uncorrected gap. In contrast, in the 1990s, the degree of positive selection decreased for both men and women, possibly reflecting the booming economy and, for women, the impact of welfare reform. Again, the changes for women were slightly larger than those for men, resulting in somewhat faster convergence in the selection-corrected gap. Between 1999 and 2011, the degree of positive selection again increased for men and women, possibly reflecting the impact of the Great Recession, this time with a larger effect for men. Thus, while the uncorrected gap rose during this period, when we correct for selection bias, the gap stayed roughly constant. Correcting for selection bias thus virtually eliminates a somewhat implausible increase in the uncorrected gap during this period. Finally, the recovery from the Great Recession (2011-2015) reduced the degree of positive selection among both men and women, with a slightly larger effect for women. The result was a roughly constant unexplained gap over this time period when we correct for selection, in contrast to a slight increase when do not.

We note that, like the bounding approach discussed below, our method of correcting for selection combines the possible effects of the relative size of Sample 4 vs Sample 1 and differences in wages between the two groups. Convergence could represent the effects of changes in both factors. However, our results were virtually identical when we applied constant frequency weights to our regression coefficients so as to fix the share of Sample 1 vs Sample 4 observations in each year used in computing the unexplained gaps (see Appendix Figure 1).[17]

Table 3B presents results for potential experience. This is of interest because, as noted, data sources used in analyses of the gender wage gap often do not have data on actual experience. Moreover, if actual experience is endogenous to wage offers then the potential experience specification may be considered a reduced form. The results indicate that we continue to find greater convergence in the raw and unexplained gap after correcting for

[17] Specifically, after estimating each year's regressions, we reweighted each year's data, so that the weighted frequency of Sample 1 observations in Sample 4 was set to the 1981 level. For comparison, we also reweighted based on the 2015 frequency of Sample 1 observations. Changes in the unexplained gaps for these reweighted samples thus only reflect wage differences and not the relative share of Sample 1 observations.

selection bias (i.e., in Sample 4 compared to Sample 1) using potential experience rather than actual experience. However, a comparison of the results for potential and actual experience highlights the fact that actual experience is a crucial factor in accounting for levels and changes in the gender wage gap. Measured factors explain relatively little of the levels or changes in the gender gaps when potential experience is employed.

Finally, in Appendix Table A2 we present our unconditional quantile results for a selection of additional quantiles: the 10th, 25th, 75th, and 90th. As a word of caution, the rationale for our Sample 4 wage imputation strategy is stronger for population gaps at the median than for additional quantiles for reasons discussed above. However, the results are still instructive.

In Sample 1 we find that convergence in the raw gender wage gap is larger for the 10th and 25th percentiles than for the 75th and 90th, with the 50th percentile (our primary results from Table 3A) comprising an intermediate case. Convergence in the unexplained gap tends to fall as we move up the male and female distributions, and the 90th percentile, in particular, exhibits little convergence in the unexplained gap. Adjustment for selection bias (Sample 4) raises our estimates of the levels of the raw and unexplained wage gap in the earlier years in each percentile. As was the case for the median in Table 3A, convergence is larger for the selectivity corrected sample for each percentile, with the exception of the unexplained gap in the 25th percentile. As with the median, this again implies that selection cannot account for the convergence in either the raw or the unexplained gender wage gap, since convergence is if anything larger in the selection corrected sample. The pattern of convergence is broadly similar to that obtained for Sample 1, with convergence in the raw and unexplained gaps tending to be larger at the 10th and 25th percentiles than at the 75th and 90th, with the 50th comprising an intermediate case. Again, an exception is that the convergence in the unexplained gap at the 25th

is relatively small. It is interesting to note that there is now more substantial evidence of convergence in the Sample 4 unexplained gap at the 90th than was the case for Sample 1.

V. Robustness of Our Imputation Procedure

A potential concern with our research design is that our imputed Sample 4 entrant wages may not accurately reflect actual wage offers. We thus test the robustness of our results to several alternatives that probe sensitivity to functional form and explore the possibility of selection into Sample 4 entrant status based on unobservables. We find that our conclusions are highly robust. These results are shown in Appendix Tables A3-A6.

First, we evaluate alternate forms of the wage assignment probit. In our baseline approach (Table A3 Panel A), we use variables from our primary human capital wage equation specification, but the results could in principle be sensitive to changing right-hand side variables. We test the sensitivity to this specification in Table A3 Panel B by augmenting the wage assignment probit with additional predictors of labor force participation: marital status and dummy variables for having exactly one child under 6 and two or more children under 6. Results are nearly identical to our primary specification.

Next, we test robustness to changes in the estimation and prediction samples. A concern with our main approach may be that the use of Sample 3 in its entirety could result in incorrect estimates due to a preponderance of individuals who are highly attached to the labor force and who may differ in unobserved characteristics from the Sample 4 entrants for whom we are imputing decile probabilities. Hence, as a further check on the selection issue, we use the Sample 3 *entrants* (rather than the full Sample 3) to estimate the ordered probits. This group has relatively weak labor market attachment, like the Sample 4 entrants, and may thus be a good comparison group on both observed and unobserved dimensions.[18] These results are shown in Table A3 Panel C. The obvious disadvantage of this approach is the small number of Sample 3 entrants, possibly generating unreliable results. Nonetheless, results are remarkably similar to our baseline model.

In Table A3 Panel D, we test how results change when we limit Sample 4 entrants to observations that we can confidently place above or below the median. These are observations whose combined probability of being above or below the median is greater than or equal to 80 percent. While increasing confidence in our imputation, this approach reduces the coverage of Sample 4.[19] Here, as with our other checks, results are nearly identical to our primary specification.

Next, we test robustness to adjusting the wage assigned conditional on being in a particular decile, with results shown in Table A3 Panel E. In our baseline specification, we assign to each of the ten duplicate observations the median wage within each decile along with the assigned probability weight. As a robustness check, we instead assign each duplicated observation a random wage drawn from a uniform distribution with end points corresponding to the minimum and maximum of each wage decile. These results are nearly identical to our primary approach. Next, in Table A4, we check robustness to alternatives to our wage decile assignment approach by using octiles, quartiles, or halves of the wage distribution. Comparing the new results (Panels B through D) to our baseline (Panel A) shows only minor differences, with our overall conclusions unchanged.

[18]Juhn (1992) similarly used a group of workers with weak labor market attachment to simulate wages of those without jobs in her analysis of male labor supply.
[19] About 29% to 57% of original Sample 4 entrants (depending on wave-by-sex group) are included under this approach.

Our wage assignment approach uses observed variables of Sample 4 entrants to impute decile probabilities based on Sample 3 incumbents. However, it is possible that Sample 4 entrants differ from Sample 3 incumbents in unmeasured characteristics. If so, then our predicted decile probabilities for Sample 4 entrants may have systematic errors. To examine this issue, we explore possible errors in our wage assignment approach by calculating how it performs for individuals for whom we do observe wages. Specifically, we calculate an error rate by estimating our wage assignment probit model on the lower-coverage samples (e.g., Sample 1) and then use the model to predict wages for other observations with observed wages (e.g., Sample 2 entrants). As we observe both the actual wages and the predicted wage, we can calculate an error rate.

Since our focus is on the gender pay gap at the median, our measure of the error rate is the difference between the actual fraction of observations with an above-median wage to the fraction predicted by the model.[20] While the actual errors in assigning deciles to Sample 4 entrants are unknown, the resulting estimated errors from Samples 2 and 3 will be instructive about the performance of the probit model, especially when applied to other low labor-force attachment individuals such as Sample 3 entrants. The error rates are shown in Table A5. In almost all cases, the predicted share of entrants above median is greater than the actual share. These patterns suggest that there is positive selection on unobservables into each sample. Individuals who are more attached to the labor force in most wave-sex cells have better unobserved characteristics than those who are less attached.

[20] These predicted probabilities are based on the probabilities of being in each of the ten deciles based on the ordered probit we described earlier. Specifically, the predicted probability of being above median is the sum of the probabilities of being in the sixth, seventh, eighth, ninth, and tenth deciles.

To explore the consequences of this exercise, we correct our Sample 4 results under the assumption that the error rates obtained from using Sample 2 to estimate Sample 3 entrants are identical to the unobserved error rates in our construction of Sample 4.[21] Intuitively, we use those observed error rates to adjust our probability weights so that the average error becomes zero. To do this, we generate a reweighting factor, which is the actual fraction of individuals belonging to a wage decile divided by the predicted fraction. That is, if the average Sample 3 predicted probability of being in decile k from the Sample 2 ordered probit is p_k, $(k = 1, \ldots, 10)$, and the actual Sample 3 incidence of being in decile k is a_k, we create a Sample 4 entrant adjustment factor of a_k/p_k. If, for instance, the Sample 2 workers are positively selected on unobservables relative to the Sample 3 entrants, then a_1/p_1 (corresponding to the first decile) for Sample 3 will likely be greater than one, while a_{10}/p_{10} (corresponding to the tenth decile) will likely be less than one.[22] We then multiply the probability weights in our baseline model by the reweighting factor and rescale the new weights for Sample 4 entrants to sum to their original PSID weight. This adjustment addresses possible selection on unobservables into Sample 4 entrant status by assuming that the extent of this selection relative to Sample 3 is the same as that of Sample 3 entrants relative to Sample 2.

The results for this correction are shown in Table A6, with our baseline results in Panel A and the corrected results in Panel B. Overall, correcting for possible errors in assigning Sample 4 entrants to deciles in this way has little effect on either the raw gender pay gap, the explained

[21] We use the error rates from Panel B in Table A5. In that panel Sample 2 is used to predict the wages of Sample 3 entrants.

[22] More formally, the adjustment procedure is as follows: 1) For each Sample 4 entrant i, compute the predicted decile probabilities $p_{1i}, \ldots, p_{10i}$ from the Sample 3 ordered probit; 2) Across all Sample 3 entrants, compute the average predicted probabilities and actual incidence for each decile, p_k and a_k; 3) Scale the Sample 4 entrant probabilities by the adjustment ratio, giving a new weight of $p_{ki} \times (a_k/p_k)$; 4) Adjust the new weights such that they sum to one for each Sample 4 entrant.

gap, or the unexplained gap, and does not affect our overall conclusions about convergence. In fact, the convergence in the gender wage gap is actually slightly higher after these potential errors are taken into account.[23]

Finally, although we believe there are considerable gains to fully utilizing all available wage data, we also considered the robustness to more restricted samples. Specifically, we applied our imputation technique to all those not included in Sample 1 and then to all those not included in Sample 2 (i.e. to those without strong labor market commitment or without current year wages). Our results (available on request) were very similar to our baseline estimates. Thus, our conclusion about convergence is not an artifact of our procedure that utilizes data from less committed workers and retrieves wages from nearby waves of the PSID.

VI. Alternative Treatments of Selection

To complement the unconditional quantile regression analysis from above, we next show how the wage information gleaned from the successive wage samples can be combined with other selection correction methods to produce alternative estimates of the gender wage gap and its trends.

a. Structural Selection Bias Correction Models

The first alternative approach we consider is to explicitly account for the correlation between unmeasured factors affecting employment and those affecting wages. We use two alternative methods to implement such an approach. First, we estimate a traditional Heckman (1979)-style selection bias correction model. This involves estimating a wage sample inclusion

[23] When we used Sample 2 entrants to compute the error rates (instead of the full Sample 2), the results were very similar. We use the full Sample 2 to estimate the prediction errors because the sample size is much larger than that of the Sample 2 entrants.

probability model and then a wage model that includes a selection adjustment term derived from the first stage model. Identification of the first equation is aided by the use of instrumental variables, and following the literature, we use variables related to the presence of children under 6 in the family — specifically, a dummy variable for having exactly one child under 6 and a dummy variable for having two or more children under 6.[24] We note that in addition to functional form, the validity of this application rests on this exclusion restriction. Both equations include a vector of controls from our human capital specification. We estimate this model separately for each wave and gender using the Full Information Maximum Likelihood estimator.[25]

Second, we estimate a selectivity bias correction model where the wage equation is a median regression model, using methods developed by Arellano and Bonhomme (2017).[26] Specifically, this method estimates a selection-corrected conditional quantile regression model, which we use to simulate the male and female wage distributions to back out unconditional median gaps. Use of median regression provides a better comparison to our preferred estimates of the gender pay gap using the probability weighting method. In addition, median regression may have more favorable properties than OLS since it is more robust to alternative distributional assumptions about wages (Koenker and Bassett 1978).

---

[24] We experimented with various forms of the children variable, including using only one dummy variable for having any children under 6. In addition, we implemented models with a continuous variable for the number of children under 6 as the instrument. The results, available on request, were similar to those in Tables 4 and 5.

[25] In two instances in Sample 3, we do not obtain convergence in this estimator and instead estimate the gap and decomposition using a traditional two-step approach. In our bootstrap procedure, again based on 599 resamples, we prioritize the FIML estimator, but defer to the two-step approach when that fails to converge, and finally do not correct wages at all in cases when the probit in the first stage of the two-step does not converge. In practice, failure of the first-stage probit to converge is only an issue when wage sample coverage approaches 1, and hence the selection correction is not very relevant. Specifically, this only affected men (largely in earlier waves) and occurred in just 36 instances out of the 8985 male selection correction models that were estimated during the bootstrap procedure.

[26] As mentioned earlier, Maasoumi and Wang (2019) also use these methods in their study of the gender pay gap.

Tables 4 and 5 show results for the selection bias corrected gender wage gaps at the mean and median, respectively.[27] Bootstrapped standard errors are in parentheses.[28] While there are some similarities in the results for both methods, the findings differ along important dimensions as well, illustrating the sensitivity of selectivity bias corrections to functional form. The first column of each table provides the non-selection adjusted raw gender wage gap for reference. As previously discussed, there are large declines over the period in all three samples for both the mean and the median.[29] The pattern of large declines in the raw gender wage gap generally holds when selection is accounted for (column 2 of each table). At the median, the gap decreases at roughly the same rate as the uncorrected raw gap and by comparable amounts across the three samples — 0.346 to 0.374 log points. At the mean, although the sample gap does decline in each sample, it does so by less than the uncorrected raw wage gap in Samples 1 and 2.[30] We observe more convergence in Sample 3, where the decrease is roughly the same as for the uncorrected raw wage gap.

The difference between the selection adjusted and the unadjusted raw gap is the selection bias (column 3 of each table). The results appear to be sensitive to functional form and to vary across samples. For Samples 1 and 2, we find that estimated selectivity bias increases over time

[27] We also estimated our selection corrected models using potential experience. The results for both the mean and the median, available on request, are broadly similar to those in the paper, which use actual experience.

[28] As with our primary results, we draw bootstrap samples by PSID wave and recompute all estimated gaps and decompositions in each iteration. We use 599 iterations in the mean selection model, but only 299 iterations in the median selection model due to computational burden. The 299 iterations in the median selection model matches Maasoumi and Wang (2019). Further, in the quantile selection model, again for computational ease, we follow Maasoumi and Wang (2019) in fixing the wave-by-sex copula parameter rho to the value computed in the main sample. While we do not fix any of the selection parameters in the mean selection correction bootstrap, we find that standard errors are similar when we do so.

[29] Note that the uncorrected gaps in Table 5 differ slightly from those in Table 3. They are not identical because the raw median gaps in Table 5 are calculated by simulating unconditional wage distributions via the conditional quantile regression model (Chernozhukov, Fernández-Val, and Melly 2013) and thus does not exactly equal the true median gap. This is done to facilitate comparison to the quantile selection model, which uses this same method of backing out the unconditional distribution. It is encouraging that the estimates are very close.

[30] Standard errors are somewhat large in Sample 1 and the convergence is not significant at traditional levels.

at the mean but that it decreases over time at the median. However, the only case that is significant at traditional levels is the change at the median in Sample 2. As we move to the more inclusive Sample 3, selection bias becomes smaller both at the mean and at the median and estimated selectivity bias has small effects on the levels and trends in the gender wage gap in both cases.

Turning to column 5 of each table, we see that, as was the case for selectivity bias, the unexplained gap shows markedly different changes over time at the mean vs. the median. At the median (Table 5), we find substantial and highly significant convergence in the unexplained gap of roughly similar magnitude for all three samples. In contrast, at the mean, there is virtually no convergence in the unexplained gap for Samples 1 and 2, though standard errors are very large. These results are very similar to the lack of convergence in Mulligan and Rubinstein's (2008) results. However, there is robust evidence of convergence for Sample 3 at roughly the same magnitude as for the median.

The differences in the selectivity bias corrected results at the mean and those at the median illustrate the sensitivity of these structural methods to functional form, with stronger evidence of convergence at the median. Since analysis of the median is more robust to outliers than the mean, and since our results for Sample 3 indicate convergence at the mean as well, it seems reasonable to conclude that the preponderance of the evidence using structural selection bias correction methods shows convergence in the unexplained pay gap. The magnitudes of the declines at the median and in Sample 3 at the mean are strikingly similar to those we obtained for the unexplained gaps in Table 3 for the corresponding samples using our probability weighting method.

b. Bounding

Blundell, et al. (2007) establish theoretical "worst case" bounds on the wage distribution in the presence of selection and show how the bounds can be shrunk by layering on assumptions. An implication of the Blundell, et al. (2007) bounding formulae is that bounds become tighter the larger the proportion of observations for which wages are available, allowing us to get more informative bounds through our additional wage samples (S2 and S3). This insight was also exploited by Olivetti and Petrongolo (2008), though we extend their exercise by examining *trends* in the median wage gap in the US, and by using results in Blundell, et al. (2007) to further tighten the bounds.

To derive worst case bounds on the gender wage gap, first consider bounds on each gender-specific wage distribution. Without complete sample coverage for wage data, the true gender-specific wage distribution is unknown. However, the true distribution can be decomposed as follows:

$$F(w \mid g) = F(w \mid g, E = 1)P(g) + F(w \mid g, E = 0)[1 - P(g)] \qquad (3)$$

where $F(w \mid g)$ is the wage CDF given gender $g \in m, f$, $P(g)$ is the probability of observing a wage for gender g, and $E \in 0{,}1$ represents no observed wage and observed wage, respectively.[31] Each element of the above decomposition can be observed with the exception of $F(w \mid g, E = 0)$. However, since a CDF must, by definition, fall between 0 and 1, the population wage distribution can be bounded above by

$$F(w \mid g, E = 1)P(g) + [1 - P(g)] \qquad (4)$$

and bounded below by

[31] To fix ideas, *E* can be thought to represent "employed," though we extend this model to bound wage distributions when wages are obtained from our Sample 3 construction, which does not exclusively consist of the currently employed.

$$F(w \mid g, E = 1)P(g) \tag{5}$$

Blundell et al. (2007) note that bounds on quantile q of the gender-specific wage distribution can be obtained by finding the wages that solve

$$q = F(w \mid g, E = 1)P(g) + [1 - P(g)] \tag{6}$$

and

$$q = F(w \mid g, E = 1)P(g) \tag{7}$$

where equation (6) corresponds to the lower bound on the quantile and equation (7) corresponds to the upper bound on the quantile. Note that the lower bound is only identified if $q \geq 1 - P(g)$ and the upper bound is only identified if $q \leq P(g)$ unless we are willing to impose an assumption on the bounds of the wage distribution. In everything that follows, we focus on q = 0.5, or the median. We obtain an upper bound on the median gender wage gap by subtracting the lower bound of the female median wage from the upper bound of the male median wage. Similarly, we obtain a lower bound on the median gender wage gap by subtracting the upper bound of the female median wage from the lower bound of the male median wage. These are "worst case" bounds because we impose no assumptions on the nature of selection. However, because the width of the bounds is decreasing in $P(g)$, we can shrink the bounds by simply adding additional wage information through our wage samples.

Table 6 presents bounds on the gender wage gap for the whole population. The table includes estimates for convergence bounds for the 1981-2015 period calculated as follows. Recall that convergence is indicated by a *reduction* in the gender wage gap over the period and thus by a negative sign on the difference between the 1981 and 2015 gender gaps, with a more negative number indicating more convergence. Bearing this in mind, the algebraically lower bound for convergence (the *most* convergence) is the 2015 upper bound gender gap minus the

1981 lower bound gender gap. The algebraically upper bound for convergence (the *least* convergence) is the lower bound gender gap for 2015 minus the upper bound gender gap for 1981. We compute 95 percent confidence regions for the bounds of the gender wage gaps and their convergence using a bootstrap procedure with 599 replications.[32]

Table 6 begins by showing the results for the worst case bounds. For Sample 1, the bounds are wholly uninformative: they are consistent with very large gender wage gaps in either direction and consistent with both increases and decreases in the gap over time.[33] However, by Sample 3, the worst case bounds become informative. The point estimates suggest at least some degree of a gender wage gap in favor of men in every PSID wave. The point estimates further imply a decrease in the gap between 1981 and 2015, since the lower bound in 1981 is greater than the upper bound in 2015. While the 95 percent confidence region on convergence slightly overlaps zero (no overlap would rule out no convergence at the 5 percent significance level), the preponderance of the 1981 interval lies well above the maximum of the 2015 interval: in 78 percent of our bootstrap iterations, the lower bound of the gender gap in 1981 is greater than the upper bound in 2015.

Next, we present results using two of the assumptions suggested by Blundell, et al. (2007) for shrinking the bounds on wage distributions: first-order stochastic dominance (FOSD)

---

[32] For a given year and sample, we create 599 replications. We then compute a 95 percent confidence region for the gender pay gap in that year by computing the 5th percentile of the lower bound estimates among the 599 replications and the 95th percentile of the upper bound estimates. To create a confidence region for convergence, we compute the lower bound and the upper bound for convergence for each replication. The 95 percent confidence region for convergence is then the interval between the 5th percentile for the lower bound of convergence and 95th percentile of the upper bound for convergence. We use the 5th and 95th percentiles (instead of, for example, the 2.5 and 97.5 percentiles) in light of Imbens and Manski's (2004) reasoning that when a parameter can only be identified in a region with positive width, the "true parameter can be close to at most one of the region's boundaries…Then, asymptotically the probability that the estimate for the lower bound exceeds the true value can be ignored when making inferences on the true parameter" (p. 1845).

[33] For 1981 we are unable to provide assumption free bounds on the median wage in Sample 1 because fewer than half of the women have observed wages. We report bounds in this wave using the assumption that unobserved female wages are bounded by the highest and lowest observed female wage in 1981.

and exclusion restrictions. Under the assumption of first-order stochastic dominance of the observed wage distribution relative to the unobserved wage distribution, the median wage of the unobserved distribution can be no higher than the median wage of the observed distribution. This assumption has no effect on the lower bound for each gender-specific median wage but reduces the upper bound on the unobserved median wage to be exactly that of the observed median wage. Table 1 shows that this assumption is plausible: in every case, the median wage of sample entrants is no greater than the median of the incumbents and in most cases is substantially lower.[34] Note that this does not assume that every entrant's wage is below the incumbent median, but merely that the *median* wage of entrants is no greater than the median wage of incumbents. Table 6 shows that a FOSD assumption substantially shrinks the bounds on the wage gap, especially the lower bound. Now, in Sample 2, there is some evidence of convergence, since the upper bound of the 2015 pay gap at 0.466 log points is only slightly above the lower bound of the 1981 pay gap at 0.464 log points. We note, however, that the 95 percent confidence region on convergence crosses zero.

When we turn to the more inclusive Sample 3, we find that both the point estimates and confidence region strongly suggest convergence. Specifically, the upper bound in 2015 is 0.336 log points, and the lower bound in 1981 is 0.583 log points. Moreover, the 95 percent confidence region on convergence is far from crossing zero, and in 100 percent of the bootstrap iterations the lower bound in 1981 is greater than the upper bound in 2015.

Blundell, et. al (2007) show that another method for shrinking the bounds beyond the worst case scenario can be devised if one is willing to assume that there is a variable, *z*, that

[34] Maasoumi and Wang (2019) implement a test of FOSD and reject it in many years of CPS data, but their test relies on a valid instrumental variable. They use children under 5 as an instrument, but if this variable is itself endogenous, then the rejection of FOSD may not be valid.

affects the probability of employment without affecting the wage — i.e., an exclusion restriction. Under this assumption, the bounds on the population wage distribution become

$$\max_z\{F(w \mid g,z,E=1)P(g,z)\} \leq F(w \mid g) \leq \min_z\{F(w \mid g,z,E=1)P(g,z)+[1-P(g,z)]\} \tag{8}$$

To implement the exclusion restriction, we use a three-valued instrument indicating the number of children under 6, with values for zero, one, or two or more.[35] We apply this to both men and women. Table 6 shows the bounding results for imposing this exclusion restriction, as well as results for imposing it in combination with FOSD. The exclusion restriction shrinks the bounds by similar proportions to the FOSD assumption. Using either the exclusion restriction alone or the exclusion restriction with FOSD, we continue to find strong evidence of convergence in Sample 3, and strong evidence of convergence in Sample 2 when we use both FOSD and the exclusion restriction. However, when combining the exclusion restriction with FOSD, some of the bounds are invalid: in some cases, the lower bound exceeds the upper bound. This implies that one or both of the assumptions fail in the years in question. Given our confidence in the FOSD assumption discussed above, we suspect this result implies a violation of the exclusion restriction. Nonetheless, the bounding exercise shows considerable evidence of convergence when using Sample 3, with some evidence for Sample 2 as well.

One downside of this exercise is that we cannot obtain bounds on a measure directly comparable to the unexplained wage gap. However, we can provide some indirect evidence on this matter by showing bounds for segments of the population with similar covariates. Specifically, we categorize individuals into four mutually exclusive groups divided by education and experience: first by whether they have more than eight years of full-time experience, and

[35] We create one instrument with the three values, rather than two dummy variables, for tractability.

then by whether they have any college education. We present these results in Table 7 and do find considerable evidence of convergence within groups.[36] We only present results for bounds under the assumption of first order stochastic dominance but show full bounding results for each category in Appendix Tables A7A and A7B.[37]

As with the bounding results from the entire sample, the bounds become much more informative as we move through wage samples, and we concentrate on the results for Sample 3. In all four groups, the Sample 3 point estimate for the lower bound on the gap in 1981 is greater than the Sample 3 upper bound on the gap in 2015, meaning the point estimates imply convergence. Only in the low experience-low education group are these two bounds close in magnitude. For the other three groups (low education-high experience; high education-low experience, and high education-high experience), the upper bound in 2015 is 0.10 to 0.12 log points less than the lower bound for 1981. The 95 percent confidence regions on convergence suggest that we can rule out no convergence on both high experience groups, with almost all of the confidence region indicating convergence within the remaining two groups. Moreover, for these latter two groups, the lower bound for the gender pay gap in 1981 is greater than the upper bound in 2015 in 92 percent of the bootstrapped iterations for the high education-low experience group and 64 percent of the iterations for the low education-low experience group. Overall, then, our bounding exercise shows strong evidence of convergence in the gender pay gap within skill groups.

[36] We use 299 bootstrap iterations and the same procedure outlined earlier to compute confidence regions in Table 7.

[37] Using the exclusion restriction or the exclusion restriction combined with FOSD also leads us to conclude convergence within skill groups. However, in several instances, the bounds were invalid (i.e., the lower bound was in some cases larger than the upper bound within a given sample). This suggests that the exclusion restriction may be invalid, an assessment that also calls into question the structural selection bias correction methodology discussed above.

c. Identification at Infinity

The selection bias problem in wage equations arises when the unmeasured factors affecting employment also affect wages. Drawing on work by Chamberlain (1986) and Heckman (1990), Mulligan and Rubinstein (2008) point out that if we can isolate a subsample of workers who have characteristics such that virtually all of them will be employed, then there will be no unmeasured factors affecting employment. If the expected and actual probability of employment are exactly one, there can be no correlation between omitted factors and the error term in a wage equation. This is the identification at infinity method of correcting for selection bias. A drawback of the method is that the sample of workers who have an employment probability near one may not be representative of the broader population. For example, among Mulligan and Rubinstein's (2008) 1970-99 female sample of workers who had at least an 80 percent probability of employment, fully 76 percent were never married, 98 percent had graduate degrees, and the raw gender pay gap was only 0.045 log points (Mulligan and Rubinstein 2008, p. 1104). Note that these women were being compared to a more heterogeneous group of men with, for example, much lower education levels than the women. The authors find little convergence in the gender wage gap among this group and thus conclude that the gender wage gap overall has not fallen. We should point out, however, that this is an elite group of women in the labor force, and earlier research has found suggestive evidence that the gender pay gap has fallen much more slowly at the top of the distribution than elsewhere (e.g., Blau and Kahn 2017). Of course, unlike Mulligan and Rubinstein (2008), such analyses compare high earning women to high earning men.

With these caveats in mind, we implement the technique by first estimating a probit separately for each wave and gender where the dependent variable is employment full-time and

26+ weeks and the independent variables are from our baseline human capital specification. We estimate separate employment probit regressions for each wave because the determinants of employment may have changed over time. For comparison, we also present results using fixed 1981 employment probits. Ideally, we would only include groups whose employment probability is close to 100 percent, as this would lead to the least bias. Practically, this results in small and unworkable sample sizes. Following Mulligan and Rubinstein (2008), we set the threshold for being included in the identification at infinity sample at 80 percent.[38]

The estimates using the identification-at-infinity method are shown in Table 8. In Panel A, we present results where the sample is obtained using a probit model estimated separately for each wave. The raw mean gender wage gap was 0.305 log points in 1981, while the median gap was 0.296 log points. However, by 2015, women in the infinity sample actually outearned men by 0.062 log points at the mean and 0.055 log points at the median. The fact that women earned more than men in 2015 in this sample is related to the construction of the identification at infinity samples. By 2015, virtually all (98.4 percent) of women with high predicted employment probabilities had a BA or advanced degree, in comparison to 60.9 percent of men; women also had more years of full-time work experience than men did. While these descriptive statistics for the infinity sample suggest that this group is not representative of the labor force, we nonetheless note that there is strong, statistically significant convergence in the raw gender wage gap.[39] The unexplained gap also decreases, from 0.375 log points in 1981 to 0.231 in 2015 at the mean (a statistically significant reduction), and from 0.357 to 0.254 at the median (although this reduction is not significant). Thus, even among a group with very high predicted employment rates, there

[38] We also restrict the wage sample to those who satisfy Sample 1 inclusion restrictions.

[39] Not surprisingly, the explained gap favors women in each year, since women in the infinity sample have more education and full-time experience than men do; women's advantage in the explained gap grows over time as their relative education levels rise.

is a substantial unexplained gender pay gap, and it appears to have fallen over time, as our basic approach also showed in Table 3.

In Panel B, we test the robustness of our infinity results to obtaining the infinity sample using an employment probability probit estimated only on the 1981 wave. The results are qualitatively similar to those in Panel A, except that by 2015, there is still a raw gender pay gap favoring men, although it is small and not significant: 0.080 log points at the mean and 0.089 at the median.[40] The unexplained gap is very similar to that in Panel A in each year. Most importantly, using a constant 1981 probit to populate our infinity samples, we continue to see convergence in raw and unexplained gender pay gaps over time.[41]

Thus, all specifications of identification at infinity show convergence in the raw and unexplained gender gaps at both the mean and the median. In terms of magnitude, the resulting estimates of convergence are somewhat lower than in our preferred probability weighting method (Table 3). This may suggest slower overall convergence using this method. However, it could, and more likely does, reflect the slower progress of the sample of elite women who, as noted, other evidence suggests had slower progress than women overall during this period.

Our finding of convergence in the unexplained gender pay gap among the infinity sample contrasts with that of Mulligan and Rubinstein (2008). They used the CPS and of course did not have data on actual experience. Ideally, we would explore the role of specification by implementing our specification on the CPS and Mulligan and Rubinstein's (2008) specification on the PSID. Unfortunately, this is not possible. We obviously cannot calculate actual experience in the CPS since that information is not collected in that data set. Moreover, as the PSID is a

[40] As in Panel A, the explained gap favors women in each year and grows more favorable to women over time.

[41] Out of concern that the sex-specific regression may not perform well given the small number of women in the identification at infinity sample, in Appendix Table A8 we instead estimate a single wage regression including a female dummy variable. Our decomposition results are highly robust to this procedure.

smaller dataset, we cannot obtain reasonable sample sizes of women in the identification at infinity sample using potential rather than actual experience (and other covariates).[42] By including actual experience as an explanatory variable, we are better able to predict strong labor force attachment (at least an 80 percent probability of employment) and generate sufficient samples of women who meet this cutoff in the PSID. While it is unfortunate that we cannot replicate Mulligan and Rubinstein's (2008) specification in the PSID, we would again point out that, whether or not the identification at infinity method reveals convergence, this group is not likely to be representative of the labor force. While conclusions about the infinity sample may be internally valid, they are likely to differ from those for the whole labor force.

VIII. Conclusions

In this paper, we have studied the impact of selectivity bias on the gender pay gap, with special attention to the question of whether the gender pay gap has fallen since 1981, as raw data and ordinary least squares regression analyses without selection bias correction have indicated. Selection bias arises when unmeasured factors affecting employment also affect wages given employment. If this is the case, then the measured gender pay gap among the employed may not be representative of the gender gap in wage offers for the population. Moreover, if selection bias changes over time in particular ways, the apparent reduction in the gender pay gap since 1981 may be an illusion caused by selectivity bias.

Using data from the PSID, we employ a variety of techniques to address the selection issue. The PSID has two important advantages for our purposes that standard data sources such as the CPS or ACS do not. First, it is longitudinal, allowing us to retrieve wage offers for those

[42] This is hardly surprising in that even with the large CPS samples at their disposal, Mulligan and Rubinstein obtain only about 300 female observations per five-year CPS cross section.

who are not currently employed.[43] Second, it has information on actual labor market experience. These two features of the PSID greatly reduce the scope for selectivity bias to influence wage analyses because with the additional wage data, fewer individuals have unobserved wages, and actual experience is an important omitted variable in other data sets likely affecting both wages and employment. For those still without observed wages, in our preferred probability weighing approach, we probabilistically assign them to wage deciles and then compute and analyze the resulting gender gaps. This method does not require restrictive identification assumptions. We find that the gender gap in the population is somewhat larger than that among full-time workers with substantial labor force commitment. However, we find strong convergence since 1981 in both the raw gender pay gap (0.378 log points) and the unexplained gender pay gap (0.204 log points). Indeed, our results suggest, if anything, that correction for sample selection bias increases our estimates of convergence on both measures. We subject our findings to a variety of robustness checks and find our conclusions unchanged.

We then use our data to explore the implications of alternative methods of correcting for selectivity bias. These include structural models of employment and wages, bounding exercises, and the identification at infinity method. The results for these methods almost always also show convergence in the gender pay gap since 1981. While each method has advantages and disadvantages, the overall similarity of the findings across methods gives us considerable confidence that women today receive higher relative wage offers than in 1981.

Nonetheless, women continue to earn substantially less than men. In 2015, based on our main selection corrected estimates, the raw (unadjusted) gender wage gap was 0.242 log points at the median and the unexplained gender wage gap (adjusted for gender differences in covariates

[43] While a short panel can be constructed from the CPS, it can only cover two years.

including education and experience) was 0.206 log points. In considering policies to further narrow the gender wage gap, it is important to know whether there has been convergence in gender gaps. Our results strongly suggest that there has.

References

Arellano, Manuel and Stéphane Bonhomme (2017). Quantile Selection Models with an Application to Understanding Changes in Wage Inequality. *Econometrica*, 85(1): 1-28.

Bayer, Patrick and Kerwin Kofi Charles (2018). Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940. *Quarterly Journal of Economics*, (133(3): 1459-1501.

Blank, Rebecca M. (1990). Are Part-Time Jobs Bad Jobs? *In A Future of Lousy Jobs? The Changing Structure of U.S. Wages*, edited by Gary Burtless. Washington, D.C.: Brookings Institution Press, pp. 123-155.

Blau, Francine D. and Andrea H. Beller (1988). Trends in Earnings Differentials By Gender, 1971–1981. *Industrial & Labor Relations Review*, 41 (4): 513–29.

Blau, Francine D. and Lawrence M. Kahn (1997). Swimming Upstream: Trends in the Gender Wage Differential in the 1980s. *Journal of Labor Economics*, 15(1, Part 1): 1-42.

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

Blau, Francine D. and Lawrence M. Kahn (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3): 789-865.

Blau, Francine D. and Anne E. Winkler (2018). *The Economics of Men, Women, and Work*, Eighth Edition. Oxford, UK and New York, NY: Oxford University Press.

Blinder, Alan S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, 8(4): 436-55.

Blundell, Richard, Amanda Gosling, Hidehiko Ichimura, and Costas Meghir (2007). Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds. *Econometrica*, 75(2): 323-63.

Chamberlain, Gary (1986). Asymptotic Efficiency in Semi-Parametric Models with Censoring. *Journal of Econometrics*, 32(2): 189-218.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly (2013). Inference on Counterfactual Distributions. *Econometrica*, 81(6): 2205-2268.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3): 953-973.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2).

Frandsen, Brigham R. and Lars J. Lefgren (2018). Testing Rank Similarity. *Review of Economics and Statistics*, 100(1): 86-91.

Heckman, James J. (1979). Sample Selection Bias as Specification Error. *Econometrica*, 47(1): 153-161.

Heckman, James J. (1990). Varieties of Selection Bias. *American Economic Review Papers and Proceedings* 80(2): 313-318.

Hirsch, Barry T. (2005). Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills. *Industrial & Labor Relations Review*, 58(4): 525-551.

Imbens, Guido W. and Charles F. Manski. (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica* 72(6): 1845-1857.

Jacobsen, Joyce, Melanie Khamis, and Mutlu Yuksel (2015). Convergences in Men's and Women's Life Patterns: Lifetime Work, Lifetime Earnings, and Human Capital Investment. *Research in Labor Economics*, 41: 1-33.

Juhn, Chinhui (1992). Decline of Male Labor Market Participation: The Role of Declining Market Opportunities. *Quarterly Journal of Economics*, 107(1): 79-121.

Kassenboehmer, Sonja C. and Mathia G. Sinning (2014). "Distributional Changes in the Gender Wage Gap. *ILR Review*, 67(2): 335-361.

Koenker, Roger and Gilbert Bassett, Jr. (1978). Regression Quantiles. *Econometrica* 46(1): 33-50.

Maasoumi, Esfandiar and Le Wang (2019). The Gender Gap between Earnings Distributions. *Journal of Political Economy*, 127(5): 2438-2504.

Machado, Cecilia (2017). Unobserved Selection Heterogeneity and the Gender Wage Gap. *Journal of Applied Econometrics* 32(7): 1348-1366.

Manski, Charles F. (1994). The Selection Problem. In Christopher Sims, ed., *Advances in Econometrics, Sixth World Congress*, Vol. 1. Cambridge, U.K.: Cambridge University Press, 143-170.

Manski, Charles F. and John V. Pepper (2000). Monotone Instrumental Variables: With Application to the Returns to Schooling. *Econometrica*, 68(4): 997-1010.

Meara, Katie, Francesco Pastore, and Allan Webster (2020). The Gender Pay Gap in the USA: A Matching Study. *Journal of Population Economics*, 33(1): 271-305.

Mincer, Jacob, and Solomon Polachek (1974). Family Investments in Human Capital: Earnings of Women. *Journal of Political Economy*, 82 (2 Part 2): S76–108.

Mulligan, Casey B. and Yona Rubinstein (2008). Selection, Investment, and Women's Relative Wages over Time. *Quarterly Journal of Economics*, 123(3): 1061-1110.

Neal, Derek (2004). The Measured Black-White Wage Gap among Women Is Too Small. *Journal of Political Economy* 112 (1, Part 2): S1-S28.

Neal, Derek A. and William R. Johnson (1996). The Role of Pre-Market Factors in Black-White Wage Differences. *Journal of Political Economy*, 104(5): 869-895.

Oaxaca, Ronald L. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3): 693-709.

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

O'Neill, June and Solomon Polachek (1993). Why the Gender Gap in Wages Narrowed in the 1980s. *Journal of Labor Economics*, 11(1, Part 1): 205-228.

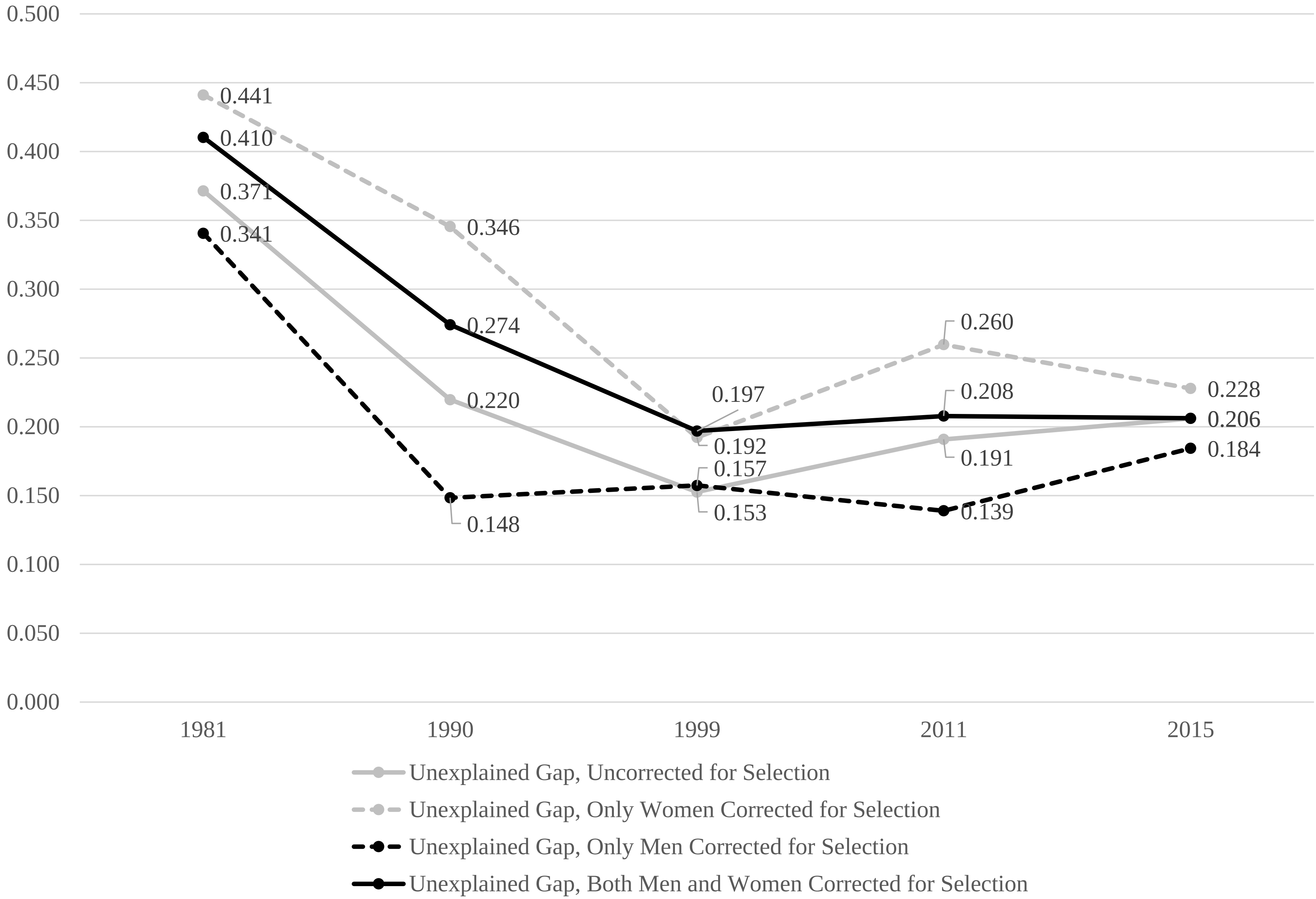
Töpfer, Marina (2017). Detailed RIF decomposition with selection: The gender pay gap in Italy. Hohenheim Discussion Papers in Business, Economics and Social Sciences 26-2017, University of Hohenheim, Faculty of Business, Economics and Social Sciences.

Appendix

A1. Data Construction

Our main PSID sample consists of individuals aged 25 to 54, who are the household head or the spouse of the household head. We drop those who are in the military, the self-employed, and those with non-credible wages. Non-credible wages are either less than or equal to $2 an hour or more than $250 an hour. We also drop the Immigrant and Latino subsamples of the PSID. Using this main sample, we then construct the Samples 1 to 4 referenced in the text. The computation of certain PSID variables, particularly actual experience and educational attainment, is non-trivial. For these, we follow Blau and Kahn (2017), who detail this process in an appendix section.

Figure 1: Unexplained Median Pay Gaps Correcting and Not Correcting for Selection



Notes for Figure 1

Entries are based on the unconditional median regression models of Table 3A. The Unexplained Gap, Uncorrected for Selection is the unexplained gap comparing men in Sample 1 and women in Sample 1. The Unexplained Gap, Only Women Corrected for Selection is the unexplained gap comparing men in Sample 1 with women in Sample 4. The Unexplained Gap, Only Men Corrected for Selection is the unexplained gap comparing men in Sample 4 with women in Sample 1. The Unexplained Gap, Both Men and Women Corrected for Selection is the unexplained gap comparing men in Sample 4 and women in Sample 4.

Table 1: Wage Characteristics by Wage Sample

	Sample Count		Entrant Count		Coverage		Whole Sample Median Wage		Whole Sample Median Wage Gap		Entrant Median Wage		Entrant Median Wage Gap	
	Men	*Women*	*Men*	*Women*	*Men*	*Women*	*All Men*	*All Women*	*Log Gap*	*Ratio*	*Male Entrants*	*Female Entrants*	*Log Gap*	*Ratio*
Sample 1 (Full-time, 26+ weeks worked)														
1981	2207	1413	-	-	0.845	0.426	3.091	2.615	0.476	62.141	-	-	-	-
1990	2619	1999	-	-	0.859	0.533	3.081	2.777	0.304	73.771	-	-	-	-
1999	2384	2087	-	-	0.861	0.601	3.056	2.843	0.213	80.797	-	-	-	-
2011	2038	2070	-	-	0.744	0.551	3.163	2.976	0.187	82.953	-	-	-	-
2015	2192	2154	-	-	0.787	0.589	3.143	2.964	0.179	83.626	-	-	-	-
Sample 2 (Adds part-time and part-year)														
1981	2381	1912	174	499	0.906	0.599	3.083	2.550	0.534	58.645	2.913	2.396	0.517	59.610
1990	2760	2637	141	638	0.906	0.716	3.063	2.694	0.369	69.132	2.725	2.353	0.372	68.905
1999	2507	2617	123	530	0.909	0.768	3.056	2.796	0.260	77.123	3.054	2.584	0.470	62.500
2011	2252	2626	214	556	0.816	0.705	3.145	2.925	0.220	80.251	2.733	2.708	0.025	97.500
2015	2382	2665	190	511	0.841	0.742	3.126	2.918	0.208	81.220	2.636	2.662	-0.026	102.623
Sample 3 (Adds nearby wave wages)														
1981	2642	2698	261	786	0.978	0.818	3.050	2.452	0.597	55.023	2.530	2.192	0.338	71.314
1990	3044	3368	284	731	0.981	0.891	3.024	2.589	0.435	64.752	2.478	2.103	0.375	68.724
1999	2707	3142	200	525	0.974	0.905	3.045	2.746	0.299	74.178	2.549	2.317	0.232	79.323
2011	2636	3334	384	708	0.937	0.864	3.094	2.851	0.243	78.446	2.636	2.503	0.133	87.569
2015	2631	3228	249	563	0.928	0.874	3.086	2.852	0.234	79.167	2.589	2.442	0.147	86.331
Sample 4 (Adds imputed wages)														
1981	2701	3307	59	609	1.000	1.000	3.047	2.426	0.621	53.745	2.853	2.224	0.629	53.329
1990	3112	3838	68	470	1.000	1.000	3.018	2.555	0.462	62.991	2.587	2.287	0.300	74.092
1999	2780	3476	73	334	1.000	1.000	3.036	2.724	0.312	73.221	2.681	2.499	0.183	83.311
2011	2815	3869	179	535	1.000	1.000	3.066	2.815	0.251	77.769	2.833	2.572	0.261	77.052
2015	2821	3722	190	494	1.000	1.000	3.062	2.820	0.241	78.580	2.604	2.596	0.008	99.159

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using education, experience, and marital status. Coverage is defined as the fraction of the wage sample represented in the entire wave-by-sex PSID sample.

Table 2: Sample Characteristics by Wage Sample

	Years of Schooling				Fraction with BA or Advanced Degree				Years Full Time Experience			
	Men		Women		Men		Women		Men		Women	
	Total	*Entrants*	*Total*	*Entrants*	*Total*	*Entrants*	*Total*	*Entrants*	*Total*	*Entrants*	*Total*	*Entrants*
Sample 1 (Full-time, 26+ weeks worked)												
1981	13.528	-	13.355	-	0.289	-	0.236	-	17.216	-	11.895	-
1990	13.917	-	13.872	-	0.314	-	0.278	-	17.028	-	13.721	-
1999	14.191	-	14.352	-	0.349	-	0.334	-	17.731	-	14.736	-
2011	14.251	-	14.566	-	0.380	-	0.419	-	14.268	-	13.413	-
2015	14.340	-	14.905	-	0.378	-	0.465	-	13.293	-	12.850	-
Sample 2 (Adds part-time and part-year)												
1981	13.500	13.105	13.275	13.081	0.286	0.240	0.229	0.210	16.829	11.489	10.358	6.596
1990	13.921	13.988	13.838	13.740	0.313	0.288	0.266	0.232	16.756	11.774	12.365	8.415
1999	14.201	14.383	14.314	14.176	0.349	0.342	0.328	0.304	17.520	13.735	13.724	10.077
2011	14.190	13.563	14.515	14.332	0.368	0.235	0.408	0.368	13.974	10.938	12.571	9.556
2015	14.291	13.592	14.819	14.485	0.366	0.193	0.444	0.362	13.087	10.107	12.013	8.782
Sample 3 (Adds nearby wave wages)												
1981	13.395	12.071	13.160	12.843	0.275	0.139	0.216	0.180	16.586	13.506	9.195	6.008
1990	13.817	12.570	13.704	13.154	0.295	0.083	0.247	0.171	16.545	14.018	11.458	7.767
1999	14.153	13.491	14.244	13.855	0.340	0.212	0.317	0.258	17.328	14.658	12.972	8.752
2011	14.052	13.126	14.400	13.893	0.344	0.183	0.387	0.296	13.864	13.126	12.042	9.694
2015	14.222	13.550	14.652	13.714	0.352	0.209	0.413	0.243	12.977	11.917	11.566	9.055
Sample 4 (Adds imputed wages)												
1981	13.335	10.702	12.934	11.919	0.269	0.000	0.194	0.094	16.593	16.930	8.407	4.866
1990	13.767	11.155	13.553	12.319	0.291	0.047	0.235	0.128	16.498	14.017	10.840	5.776
1999	14.097	11.970	14.126	12.992	0.333	0.068	0.305	0.192	17.177	11.501	12.457	7.526
2011	13.951	12.455	14.218	13.058	0.329	0.109	0.361	0.193	13.673	10.838	11.565	8.532
2015	14.088	12.358	14.468	13.192	0.331	0.061	0.384	0.184	12.854	11.264	11.223	8.838

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using education, experience, and marital status.

Table 3A: Estimates of the Gender Wage Gap and Blinder-Oaxaca-Style Decomposition, Actual Experience Controls

	OLS (Mean)			Unconditional Median		
	Gap	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Sample 1 (Full-time, 26+ weeks worked)						
1981	0.446	0.128	0.318	0.476	0.105	0.371
	(0.021)	(0.015)	(0.021)	(0.027)	(0.016)	(0.028)
1990	0.279	0.083	0.196	0.304	0.085	0.220
	(0.020)	(0.013)	(0.018)	(0.026)	(0.014)	(0.027)
1999	0.245	0.061	0.184	0.213	0.060	0.153
	(0.022)	(0.014)	(0.021)	(0.029)	(0.017)	(0.028)
2011	0.208	0.004	0.204	0.186	-0.005	0.191
	(0.022)	(0.017)	(0.021)	(0.023)	(0.016)	(0.023)
2015	0.165	-0.019	0.184	0.178	-0.029	0.206
	(0.024)	(0.019)	(0.024)	(0.026)	(0.019)	(0.026)
Convergence (2015 – 1981)	-0.282	-0.147	-0.134	-0.299	-0.133	-0.165
	(0.032)	(0.024)	(0.032)	(0.037)	(0.025)	(0.039)
Sample 2 (Adds part-time and part-year)						
1981	0.486	0.173	0.313	0.534	0.133	0.400
	(0.020)	(0.018)	(0.024)	(0.028)	(0.019)	(0.030)
1990	0.353	0.122	0.232	0.369	0.120	0.249
	(0.020)	(0.015)	(0.020)	(0.028)	(0.016)	(0.031)
1999	0.287	0.081	0.206	0.260	0.082	0.178
	(0.022)	(0.017)	(0.025)	(0.027)	(0.018)	(0.029)
2011	0.228	0.021	0.206	0.220	0.009	0.211
	(0.022)	(0.022)	(0.029)	(0.031)	(0.016)	(0.031)
2015	0.197	0.005	0.193	0.209	0.000	0.209
	(0.023)	(0.019)	(0.024)	(0.023)	(0.020)	(0.025)
Convergence (2015 – 1981)	-0.288	-0.168	-0.120	-0.324	-0.133	-0.191
	(0.030)	(0.026)	(0.034)	(0.036)	(0.028)	(0.039)
Sample 3 (Adds nearby wave wages)						
1981	0.555	0.201	0.354	0.597	0.170	0.427
	(0.019)	(0.017)	(0.023)	(0.023)	(0.019)	(0.027)
1990	0.406	0.161	0.245	0.435	0.161	0.273
	(0.020)	(0.015)	(0.019)	(0.027)	(0.018)	(0.026)
1999	0.316	0.111	0.205	0.298	0.100	0.198
	(0.022)	(0.017)	(0.025)	(0.024)	(0.018)	(0.027)
2011	0.239	0.036	0.202	0.242	0.022	0.220
	(0.022)	(0.018)	(0.024)	(0.025)	(0.014)	(0.026)
2015	0.211	0.031	0.180	0.234	0.027	0.207
	(0.023)	(0.017)	(0.023)	(0.027)	(0.019)	(0.026)
Convergence (2015 – 1981)	-0.344	-0.169	-0.174	-0.363	-0.144	-0.220
	(0.029)	(0.024)	(0.032)	(0.035)	(0.027)	(0.038)
Sample 4 (Adds imputed wages)						
1981	0.584	0.228	0.356	0.619	0.209	0.410
	(0.017)	(0.017)	(0.021)	(0.025)	(0.020)	(0.029)
1990	0.433	0.187	0.246	0.462	0.188	0.274
	(0.019)	(0.014)	(0.018)	(0.028)	(0.018)	(0.027)
1999	0.328	0.123	0.205	0.312	0.115	0.197
	(0.021)	(0.016)	(0.023)	(0.022)	(0.017)	(0.023)
2011	0.256	0.053	0.203	0.251	0.043	0.208
	(0.020)	(0.014)	(0.020)	(0.027)	(0.013)	(0.027)
2015	0.217	0.040	0.177	0.242	0.036	0.206
	(0.021)	(0.015)	(0.020)	(0.028)	(0.016)	(0.026)
Convergence (2015 – 1981)	-0.367	-0.188	-0.179	-0.378	-0.174	-0.204
	(0.027)	(0.022)	(0.029)	(0.037)	(0.026)	(0.039)

Notes: See Table 3B.

Table 3B: Estimates of the Gender Wage Gap and Blinder-Oaxaca-Style Decomposition, Potential Experience Controls

	OLS (Mean)			Unconditional (Median)		
	Gap	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Sample 1 (Full-time, 26+ weeks worked)						
1981	0.446	0.024	0.422	0.476	0.022	0.454
	(0.020)	(0.010)	(0.018)	(0.026)	(0.010)	(0.025)
1990	0.279	0.022	0.257	0.304	0.020	0.284
	(0.021)	(0.011)	(0.018)	(0.025)	(0.012)	(0.023)
1999	0.245	0.000	0.245	0.213	-0.005	0.218
	(0.022)	(0.012)	(0.019)	(0.028)	(0.013)	(0.026)
2011	0.208	-0.018	0.226	0.186	-0.016	0.202
	(0.024)	(0.014)	(0.020)	(0.025)	(0.014)	(0.023)
2015	0.165	-0.051	0.216	0.178	-0.061	0.239
	(0.024)	(0.015)	(0.022)	(0.025)	(0.016)	(0.025)
Convergence (2015 – 1981)	-0.282	-0.075	-0.206	-0.299	-0.084	-0.215
	(0.031)	(0.018)	(0.028)	(0.036)	(0.019)	(0.035)
Sample 2 (Adds part-time and part-year)						
1981	0.486	0.013	0.473	0.534	0.012	0.521
	(0.019)	(0.009)	(0.018)	(0.027)	(0.009)	(0.026)
1990	0.353	0.012	0.341	0.369	0.009	0.359
	(0.020)	(0.010)	(0.018)	(0.027)	(0.011)	(0.025)
1999	0.287	-0.002	0.289	0.260	-0.007	0.268
	(0.022)	(0.011)	(0.019)	(0.027)	(0.012)	(0.025)
2011	0.228	-0.030	0.257	0.220	-0.028	0.247
	(0.024)	(0.013)	(0.021)	(0.032)	(0.013)	(0.030)
2015	0.197	-0.055	0.252	0.209	-0.064	0.273
	(0.023)	(0.014)	(0.021)	(0.023)	(0.015)	(0.022)
Convergence (2015 – 1981)	-0.288	-0.067	-0.221	-0.324	-0.076	-0.248
	(0.030)	(0.017)	(0.027)	(0.036)	(0.018)	(0.034)
Sample 3 (Adds nearby wave wages)						
1981	0.555	0.014	0.540	0.597	0.013	0.584
	(0.018)	(0.008)	(0.017)	(0.022)	(0.008)	(0.021)
1990	0.406	0.025	0.381	0.435	0.022	0.413
	(0.020)	(0.010)	(0.018)	(0.026)	(0.011)	(0.024)
1999	0.316	0.001	0.315	0.298	-0.005	0.303
	(0.021)	(0.011)	(0.019)	(0.024)	(0.011)	(0.022)
2011	0.239	-0.031	0.270	0.242	-0.031	0.273
	(0.023)	(0.013)	(0.020)	(0.026)	(0.012)	(0.025)
2015	0.211	-0.042	0.253	0.234	-0.049	0.283
	(0.024)	(0.013)	(0.022)	(0.026)	(0.015)	(0.026)
Convergence (2015 – 1981)	-0.344	-0.056	-0.287	-0.363	-0.062	-0.301
	(0.029)	(0.016)	(0.027)	(0.034)	(0.017)	(0.034)
Sample 4 (Adds imputed wages)						
1981	0.569	0.017	0.551	0.603	0.017	0.586
	(0.016)	(0.008)	(0.015)	(0.023)	(0.008)	(0.022)
1990	0.416	0.030	0.386	0.446	0.027	0.419
	(0.019)	(0.010)	(0.017)	(0.026)	(0.010)	(0.024)
1999	0.321	0.005	0.316	0.306	-0.001	0.308
	(0.020)	(0.010)	(0.018)	(0.023)	(0.010)	(0.021)
2011	0.249	-0.024	0.273	0.245	-0.023	0.267
	(0.021)	(0.012)	(0.018)	(0.027)	(0.011)	(0.026)
2015	0.219	-0.038	0.257	0.241	-0.042	0.283
	(0.021)	(0.012)	(0.020)	(0.026)	(0.013)	(0.025)
Convergence (2015 – 1981)	-0.350	-0.055	-0.294	-0.362	-0.059	-0.303
	(0.027)	(0.014)	(0.025)	(0.034)	(0.015)	(0.033)

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. Bootstrapped standard errors based on 599 draws in parentheses. Table 3A controls for actual experience in the wage equation, while Table 3B controls for potential experience (both in addition to our other human capital specification controls). Note that unconditional quantile coefficients do not precisely go through the sample median, but that the medians are similar to the raw figures reported in Table 1. Decomposition covariates correspond to those in our human capital specification. Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using our human capital specification control set.

Table 4: Selection-Corrected Mean Regression and Decomposition

	Uncorrected Sample Gap	*Corrected Gap (Col. 4 + Col. 5)*	*Selectivity Bias (Col. 2 – Col. 1)*	*Explained*	*Unexplained*
	(1)	(2)	(3)	(4)	(5)
Sample 1 (Full-time, 26+ weeks worked)					
1981	0.446	0.467	0.021	0.123	0.344
	(0.020)	(0.119)	(0.119)	(0.019)	(0.122)
1990	0.279	0.384	0.105	0.063	0.322
	(0.020)	(0.049)	(0.044)	(0.016)	(0.055)
1999	0.245	0.374	0.130	0.036	0.338
	(0.021)	(0.046)	(0.042)	(0.019)	(0.053)
2011	0.208	0.450	0.242	-0.013	0.463
	(0.023)	(0.058)	(0.053)	(0.016)	(0.061)
2015	0.165	0.323	0.159	-0.025	0.348
	(0.025)	(0.158)	(0.155)	(0.017)	(0.158)
Convergence (2015 – 1981)	-0.282	-0.143	0.138	-0.147	0.004
	(0.032)	(0.198)	(0.195)	(0.025)	(0.200)
Sample 2 (Adds part-time and part-year)					
1981	0.486	0.489	0.004	0.168	0.321
	(0.020)	(0.044)	(0.042)	(0.021)	(0.051)
1990	0.353	0.440	0.086	0.102	0.337
	(0.020)	(0.043)	(0.039)	(0.016)	(0.047)
1999	0.287	0.342	0.055	0.072	0.270
	(0.022)	(0.058)	(0.055)	(0.026)	(0.073)
2011	0.228	0.355	0.128	0.015	0.340
	(0.023)	(0.056)	(0.052)	(0.022)	(0.061)
2015	0.197	0.299	0.102	-0.002	0.301
	(0.023)	(0.095)	(0.090)	(0.019)	(0.096)
Convergence (2015 – 1981)	-0.288	-0.190	0.098	-0.170	-0.020
	(0.030)	(0.105)	(0.099)	(0.028)	(0.109)
Sample 3 (Adds nearby wave wages)					
1981	0.555	0.524	-0.030	0.198	0.327
	(0.019)	(0.033)	(0.029)	(0.018)	(0.037)
1990	0.406	0.354	-0.052	0.154	0.199
	(0.019)	(0.021)	(0.013)	(0.015)	(0.023)
1999	0.316	0.333	0.017	0.095	0.238
	(0.022)	(0.033)	(0.023)	(0.024)	(0.047)
2011	0.239	0.294	0.055	0.033	0.261
	(0.023)	(0.067)	(0.064)	(0.022)	(0.076)
2015	0.211	0.194	-0.017	0.028	0.166
	(0.023)	(0.038)	(0.034)	(0.018)	(0.041)
Convergence (2015 – 1981)	-0.344	-0.330	0.014	-0.170	-0.160
	(0.030)	(0.050)	(0.045)	(0.025)	(0.055)

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. Bootstrapped standard errors based on 299 draws in parentheses. Selection-corrected results at the mean are obtained using the Heckman selection correction and are estimated with maximum likelihood estimation. In the case that the MLE fails to converge, we instead use a Heckman Two Step procedure (this only occurs for Sample 3 men in 1981 and 1999). The selection corrected population gap is the predicted average log hourly wage difference for the entire population, obtained from the selection-corrected wage equation. The selection corrected sample gap is computed analogously on each wage sample. Selectivity bias equals the difference between the selection corrected sample gap and the uncorrected sample gap (i.e., column (4)=column (3)-column (1)), and the sum of the explained and the unexplained gaps equals the selectivity corrected sample gap (i.e., column (5) + column (6)=column (3)). Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time worker and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data is available in those years; and Sample 4 imputes remaining missing wages using education, experience, and marital status. The excluded instruments used for selection correction are an indicator for the presence of one child under 6 and an indicator for the presence of two or more children under 6, applied to men and women. Wage equation variables are those listed as the human capital specification in the text.

Table 5: Selection-Corrected Median Regression and Decomposition

	Uncorrected Sample Gap	*Corrected Gap (Col. 4 + Col. 5)*	*Selectivity Bias (Col. 2 – Col. 1)*	*Explained*	*Unexplained*
	(1)	(2)	(3)	(4)	(5)
Sample 1 (Full-time, 26+ weeks worked)					
1981	0.475	0.779	0.304	0.171	0.608
	(0.037)	(0.040)	(0.043)	(0.018)	(0.036)
1990	0.292	0.414	0.122	0.108	0.306
	(0.027)	(0.024)	(0.021)	(0.014)	(0.020)
1999	0.235	0.481	0.246	0.093	0.388
	(0.028)	(0.027)	(0.029)	(0.015)	(0.026)
2011	0.203	0.443	0.240	0.031	0.412
	(0.029)	(0.036)	(0.032)	(0.021)	(0.035)
2015	0.173	0.405	0.231	-0.029	0.433
	(0.029)	(0.034)	(0.030)	(0.021)	(0.030)
Convergence (2015 – 1981)	-0.302	-0.374	-0.073	-0.200	-0.175
	(0.047)	(0.052)	(0.052)	(0.028)	(0.047)
Sample 2 (Adds part-time and part-year)					
1981	0.515	0.655	0.140	0.191	0.464
	(0.025)	(0.023)	(0.019)	(0.018)	(0.024)
1990	0.362	0.436	0.074	0.139	0.297
	(0.024)	(0.022)	(0.017)	(0.016)	(0.019)
1999	0.284	0.407	0.123	0.101	0.307
	(0.025)	(0.023)	(0.021)	(0.016)	(0.025)
2011	0.224	0.390	0.166	0.037	0.354
	(0.029)	(0.030)	(0.025)	(0.021)	(0.030)
2015	0.205	0.281	0.076	0.005	0.276
	(0.027)	(0.029)	(0.024)	(0.020)	(0.027)
Convergence (2015 – 1981)	-0.310	-0.374	-0.064	-0.186	-0.188
	(0.037)	(0.037)	(0.031)	(0.027)	(0.036)
Sample 3 (Adds nearby wave wages)					
1981	0.583	0.603	0.020	0.184	0.420
	(0.023)	(0.019)	(0.013)	(0.016)	(0.021)
1990	0.418	0.438	0.020	0.158	0.279
	(0.024)	(0.020)	(0.015)	(0.014)	(0.018)
1999	0.320	0.372	0.052	0.108	0.264
	(0.025)	(0.020)	(0.018)	(0.015)	(0.020)
2011	0.234	0.301	0.068	0.034	0.268
	(0.026)	(0.024)	(0.019)	(0.017)	(0.026)
2015	0.222	0.258	0.036	0.020	0.238
	(0.026)	(0.023)	(0.017)	(0.019)	(0.021)
Convergence (2015 – 1981)	-0.361	-0.346	0.015	-0.164	-0.181
	(0.035)	(0.030)	(0.021)	(0.025)	(0.030)

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. All results are unconditional on covariates, obtained via simulating the sample wage distribution based on conditional regression estimates. We present an uncorrected gap computed analogously, accounting for a slight difference between column 1 and results presented in Table 3. Selectivity bias equals the difference between the selection corrected gap and the uncorrected sample gap (i.e., column (3) = column (2) - column (1)), and the sum of the explained and the unexplained gaps equals the selectivity corrected gap (i.e., column (4) + column (5) = column (2)). Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using education, experience, and marital status. The excluded instruments used for selection correction are an indicator for the presence of one child under 6 and an indicator for the presence of two or more children under 6, applied to men and women. Wage equation variables are those listed as the human capital specification in the text.

Table 6: Bounding the Median Gender Wage Gap

	Summary			Log Gap Bounds							
				Worst Case		FOSD		Exclusion Restriction		FOSD + Exclusion	
	Male Coverage	*Female Coverage*	*Log Gap*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*	*Lower Bound*	*Upper Bound*
Sample 1 Base											
1981	0.845	0.426	0.476	-0.898	2.375	0.362	2.281	-0.681	1.751	0.435	1.669
				[-0.922, 2.403]		[0.322, 2.300]		[-0.696, 1.767]		[0.385, 1.685]	
1990	0.859	0.533	0.304	-0.549	1.214	0.184	1.122	-0.253	0.867	0.290	0.802
				[-0.737, 1.462]		[0.142, 1.363]		[-0.317, 0.974]		[0.220, 0.908]	
1999	0.861	0.601	0.213	-0.398	0.840	0.114	0.711	-0.230	0.641	0.179	0.584
				[-0.472, 0.895]		[0.063, 0.769]		[-0.296, 0.716]		[0.142, 0.651]	
2011	0.744	0.551	0.187	-0.840	1.176	-0.065	0.905	-0.524	0.873	0.029	0.720
				[-0.967, 1.316]		[-0.129, 1.035]		[-0.620, 1.004]		[-0.010, 0.830]	
2015	0.787	0.589	0.179	-0.637	0.936	-0.021	0.724	-0.420	0.720	0.112	0.611
				[-0.728, 1.038]		[-0.079, 0.812]		[-0.510, 0.811]		[0.051, 0.702]	
Convergence Bounds				-3.012	1.833	-2.302	0.363	-2.17	1.401	-1.557	0.176
Confidence Region				[-3.107, 1.914]		[-2.368, 0.457]		[-2.239, 1.458]		[-1.605, 0.269]	
Sample 2 Base											
1981	0.906	0.599	0.534	-0.051	1.044	0.464	0.990	0.110	0.897	0.502	0.856
				[-0.102, 1.116]		[0.420, 1.063]		[0.071, 0.948]		[0.463, 0.906]	
1990	0.906	0.716	0.369	-0.001	0.762	0.292	0.698	0.101	0.635	0.415	0.583
				[-0.060, 0.811]		[0.251, 0.748]		[0.043, 0.674]		[0.338, 0.633]	
1999	0.909	0.768	0.260	-0.029	0.555	0.201	0.474	0.057	0.445	0.247	0.412
				[-0.086, 0.600]		[0.157, 0.522]		[0.008, 0.493]		[0.210, 0.461]	
2011	0.816	0.705	0.220	-0.284	0.699	0.051	0.550	-0.145	0.567	0.129	0.486
				[-0.343, 0.760]		[-0.010, 0.602]		[-0.213, 0.636]		[0.090, 0.556]	
2015	0.841	0.742	0.208	-0.216	0.605	0.058	0.466	-0.093	0.487	0.179	0.427
				[-0.264, 0.686]		[0.017, 0.522]		[-0.139, 0.557]		[0.128, 0.489]	
Convergence Bounds				-1.26	0.656	-0.933	0.002	-0.99	0.378	-0.677	-0.074
Confidence Region				[-1.353, 0.736]		[-1.021, 0.075]		[-1.053, 0.444]		[-0.736, -0.013]	
Sample 3 Base											
1981	0.978	0.818	0.597	0.422	0.763	0.583	0.742	0.442	0.708	0.697	0.702
				[0.379, 0.801]		[0.550, 0.788]		[0.408, 0.748]		[0.642, 0.746]	
1990	0.981	0.891	0.435	0.313	0.547	0.418	0.536	0.457	0.482	0.574	0.480
				[0.267, 0.595]		[0.369, 0.578]		[0.392, 0.528]		[0.516, 0.523]	
1999	0.974	0.905	0.299	0.195	0.385	0.280	0.371	0.221	0.356	0.362	0.353
				[0.150, 0.432]		[0.240, 0.405]		[0.191, 0.388]		[0.280, 0.384]	
2011	0.937	0.864	0.243	0.040	0.400	0.183	0.353	0.131	0.352	0.276	0.333
				[-0.006, 0.446]		[0.141, 0.396]		[0.075, 0.404]		[0.203, 0.379]	
2015	0.928	0.874	0.234	0.061	0.393	0.165	0.336	0.181	0.346	0.253	0.332
				[0.005, 0.446]		[0.110, 0.390]		[0.123, 0.377]		[0.204, 0.368]	
Convergence Bounds				-0.702	-0.029	-0.577	-0.246	-0.527	-0.097	-0.449	-0.365
Confidence Region				[-0.773, 0.035]		[-0.645, -0.188]		[-0.601, -0.067]		[-0.520, -0.310]	

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. 95 percent confidence regions in brackets. The median gap is the log difference of male and female wages. The gap bounds are computed as: 1) for the upper bound, the male median upper bound minus the female median lower bound, and 2) for the lower bound, the male median lower bound minus the female median upper bound. First-order stochastic dominance (FOSD) bounds assume FOSD of the observed wage distribution relative to the unobserved wage distribution. The excluded instruments are an indicator for the presence of one child under 6 and an indicator for the presence of two or more children under 6, applied to men and women. Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using education, experience, and marital status.

Table 7: Bounding the Median Gender Wage Gap by Education and Experience Groups Using Sample 3

	Summary			*FOSD Bounds*		*Summary*			*FOSD Bounds*	
	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound
	Panel A. Low Edu., Low Exp.					*Panel B. Low Edu., High Exp.*				
1981	0.929	0.693	0.553	0.475	0.812	0.963	0.891	0.489	0.468	0.567
				[0.392, 0.934]					[0.422, 0.620]	
1990	0.914	0.744	0.352	0.262	0.606	0.976	0.938	0.353	0.341	0.388
				[0.175, 0.687]					[0.273, 0.444]	
1999	0.882	0.746	0.448	0.405	0.614	0.975	0.921	0.246	0.230	0.317
				[0.239, 0.741]					[0.164, 0.382]	
2011	0.796	0.699	0.370	0.227	0.632	0.927	0.864	0.178	0.132	0.281
				[0.069, 0.791]					[0.068, 0.378]	
2015	0.797	0.715	0.207	0.056	0.461	0.892	0.855	0.237	0.130	0.344
				[-0.074, 0.573]					[0.032, 0.436]	
Convergence Bounds				-0.756	-0.014				-0.437	-0.124
Confidence Region				[-0.944, 0.121]					[-0.560, -0.021]	
	Panel C. High Edu., Low Exp.					*Panel D. High Edu., High Exp.*				
1981	0.998	0.860	0.320	0.318	0.458	1.000	0.936	0.441	0.441	0.478
				[0.242, 0.542]					[0.363, 0.539]	
1990	0.992	0.886	0.311	0.306	0.408	0.995	0.974	0.307	0.303	0.317
				[0.205, 0.518]					[0.239, 0.391]	
1999	0.987	0.896	0.218	0.206	0.290	0.983	0.957	0.324	0.315	0.353
				[0.123, 0.367]					[0.254, 0.411]	
2011	0.960	0.870	0.142	0.110	0.253	0.970	0.918	0.293	0.255	0.352
				[0.054, 0.337]					[0.183, 0.430]	
2015	0.963	0.885	0.125	0.106	0.217	0.967	0.926	0.265	0.254	0.334
				[0.047, 0.304]					[0.183, 0.409]	
Convergence Bounds				-0.352	-0.101				-0.223	-0.107
Confidence Region				[-0.451, 0.015]					[-0.320, -0.002]	

Notes: See notes for Table 6. All results correspond to FOSD bounding from a base of Sample 3. Low education corresponds to no college education, while high education corresponds to at least some college. Low experience corresponds to less than or equal to eight years of full-time experiences, while high experience corresponds to more than eight years of full-time experience.

Table 8: Identification at Infinity Gender Wage Gap Estimates and Decomposition

	Male Count	Female Count	OLS (Mean)			Unconditional (Median)		
	Male Count	*Female Count*	*Gap*	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Panel A: Infinity Sample Based on Wave-Specific Probit								
1981	1598	185	0.305	-0.070	0.375	0.296	-0.061	0.357
			(0.048)	(0.036)	(0.047)	(0.071)	(0.034)	(0.067)
1990	2062	381	0.067	-0.140	0.207	0.085	-0.132	0.217
			(0.065)	(0.052)	(0.037)	(0.071)	(0.052)	(0.049)
1999	1887	603	0.091	-0.085	0.176	0.034	-0.104	0.138
			(0.041)	(0.037)	(0.039)	(0.050)	(0.041)	(0.051)
2011	826	179	-0.097	-0.307	0.210	-0.123	-0.347	0.224
			(0.065)	(0.063)	(0.065)	(0.079)	(0.075)	(0.095)
2015	1165	352	-0.062	-0.293	0.231	-0.055	-0.309	0.254
			(0.057)	(0.051)	(0.057)	(0.067)	(0.056)	(0.070)
Convergence (2015 – 1981)			-0.367	-0.223	-0.144	-0.351	-0.249	-0.103
			(0.075)	(0.062)	(0.074)	(0.098)	(0.066)	(0.097)
Panel B: Infinity Sample Based on 1981 Probit								
1981	1598	185	0.305	-0.070	0.375	0.296	-0.061	0.357
			(0.048)	(0.036)	(0.047)	(0.071)	(0.034)	(0.067)
1990	2056	534	0.120	-0.071	0.192	0.115	-0.063	0.178
			(0.047)	(0.034)	(0.033)	(0.054)	(0.036)	(0.042)
1999	1885	623	0.098	-0.075	0.173	0.042	-0.093	0.135
			(0.042)	(0.034)	(0.038)	(0.051)	(0.038)	(0.050)
2011	1437	587	0.093	-0.136	0.229	0.072	-0.131	0.202
			(0.051)	(0.047)	(0.041)	(0.065)	(0.050)	(0.054)
2015	1503	554	0.080	-0.163	0.243	0.089	-0.170	0.259
			(0.052)	(0.046)	(0.046)	(0.059)	(0.050)	(0.055)
Convergence (2015 – 1981)			-0.225	-0.093	-0.131	-0.207	-0.110	-0.097
			(0.071)	(0.058)	(0.066)	(0.092)	(0.060)	(0.087)
Panel C: Main Sample 1 Results (Whole Sample)								
1981	2207	1413	0.446	0.128	0.318	0.476	0.105	0.371
			(0.020)	(0.015)	(0.020)	(0.026)	(0.016)	(0.028)
1990	2619	1999	0.279	0.083	0.196	0.304	0.085	0.220
			(0.021)	(0.013)	(0.018)	(0.025)	(0.014)	(0.025)
1999	2384	2087	0.245	0.061	0.184	0.213	0.060	0.153
			(0.022)	(0.015)	(0.021)	(0.028)	(0.016)	(0.028)
2011	2038	2070	0.208	0.004	0.204	0.186	-0.005	0.191
			(0.024)	(0.016)	(0.022)	(0.025)	(0.016)	(0.025)
2015	2192	2154	0.165	-0.019	0.184	0.178	-0.029	0.206
			(0.024)	(0.018)	(0.024)	(0.025)	(0.019)	(0.027)
Convergence (2015 – 1981)			-0.282	-0.147	-0.134	-0.299	-0.133	-0.165
			(0.031)	(0.023)	(0.031)	(0.036)	(0.025)	(0.039)

Notes: See notes for Table 8. In Panel A the employment probability probit is estimated separately for each wave. For Panel B, this probability is calculated using the 1981 PSID wave, with estimated probability coefficients then applied to other waves. Wage equation and employment probability equation variables are those listed as the human capital specification in the text. Decompositions computed with female dummy method. Note that unconditional quantile coefficients do not precisely go through the sample median, but that the predicted medians are very similar to sample medians.

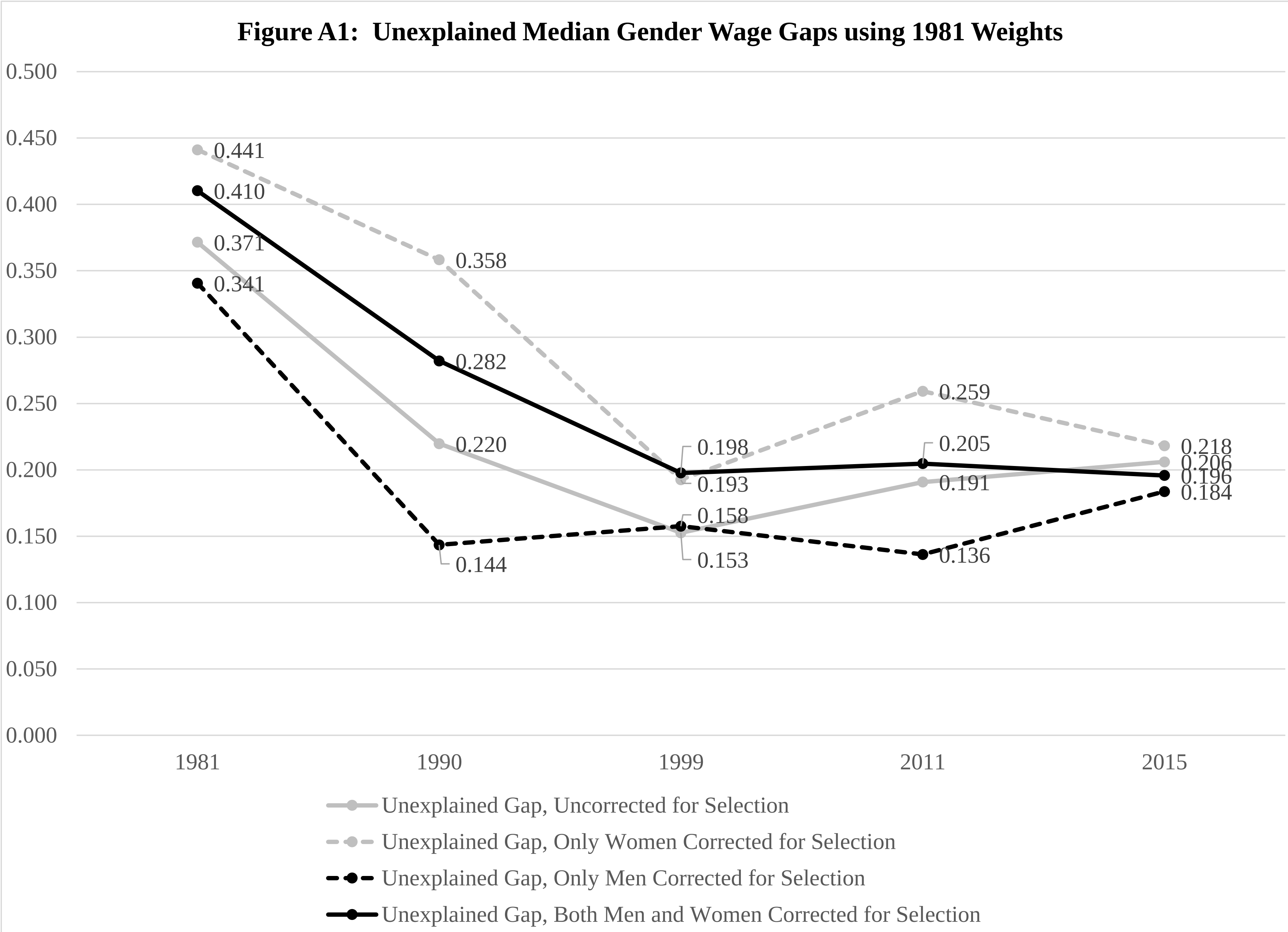


Figure A1: Unexplained Median Gender Wage Gaps using 1981 Weights
0.500
0.450
0.400
0.350
0.300
0.250
0.200
0.150
0.100
0.050
0.000
1981
1990
1999
2011
2015
0.441
0.410
0.371
0.341
0.358
0.282
0.220
0.144
0.198
0.193
0.158
0.153
0.259
0.205
0.191
0.136
0.218
0.206
0.196
0.184
Unexplained Gap, Uncorrected for Selection
Unexplained Gap, Only Women Corrected for Selection
Unexplained Gap, Only Men Corrected for Selection
Unexplained Gap, Both Men and Women Corrected for Selection

Notes for Figure A1

Entries are based on the unconditional median regression models of Table 3A. After estimating each year's regressions, we reweighted each year's data, so that the weighted frequency of Sample 1 observations in Sample 4 was set to the 1981 level. Using these reweighted samples, the entries in Figure A1 are defined the same way as the entries in Figure 1. The Unexplained Gap, Uncorrected for Selection is the unexplained gap comparing men in Sample 1 and women in Sample 1 (this is identical to Figure 1 since it uses only Sample 1 individuals). The Unexplained Gap, Only Women Corrected for Selection is the unexplained gap comparing men in Sample 1 with women in the reweighted Sample 4. The Unexplained Gap, Only Men Corrected for Selection is the unexplained gap comparing men in the reweighted Sample 4 with women in Sample 1. The Unexplained Gap, Both Men and Women Corrected for Selection is the unexplained gap comparing men and women both in the reweighted Sample 4.

Table A1: Selected Findings on Selection-Corrected Trends in the Gender Pay Gap

Paper	Technique	Data	Finding on Selection-Adjusted Gender Pay Gap
Blau and Beller (1988)	Heckman-style correction, using non-labor income and household age distribution as IVs.	March CPS	Mean wage gap fell by 23% (raw) and 20% (unexplained) between 1971 and 1981 for whites (Table 5). Mean gap fell by 13% (raw) and 4% (unexplained) for Blacks during same period (Table 5).
Blau and Kahn (2006)	Median regression combined with replacing unobserved wages with nearby year wages and assigning remaining unobserved wages relative to the median based on observables.	PSID	Median raw wage gap fell by 0.24 log points between 1979 and 1998 (Table 3). Median unexplained gap fell by 0.15 log points during the same period (Table 3).
Blundell, et al. (2007)	Bounding the population wage distribution by sex and using assumptions to shrink bounds.	UK Family Expenditure Survey	Median raw wage gap among 25-year-olds with no college education fell by 0.23 to 0.28 log points (lower and upper bounds) between 1978 and 1998 (Figure 14). Bounds cross zero for 40-year-olds with college education (Figure 14). Bounds for other groups suggest convergence but 95% confidence intervals cross zero.
Mulligan and Rubinstein (2008)	Heckman-style correction, using number of young children in household and its interaction with marital status as IVs.	March CPS	Mean unexplained wage gap did not fall between the 1975-1979 period and the 1995-1999 period (Table 1).
Mulligan and Rubinstein (2008)	Identification at infinity, analyzing sample with ≥80% employment probability.	March CPS	Mean unexplained wage gap did not fall between the 1975-1979 period and the 1995-1999 period (Figure V).
Machado (2017)	Identification from a sample of "always employed" women (work regardless of IV value), using presence of young children as IV.	March CPS	Mean raw wage gap within identification sample fell 0.26 log points between the 1976-1980 period and the 2001-2005 period (Table 5), and there was similar convergence within four separate education level groups (Table 4).
Maasoumi and Wang (2019)	Heckman-style correction and quantile regression extension, using number of young children in household as IV.	March CPS	Mean raw wage gap (smoothed) fell by 0.16 log points between 1976 and 2013 (Figure 2). Median raw wage gap (smoothed) fell by 0.18 log points over same period (Figure 2).

See reference list for full citations. In the findings column, we approximated numerical figures from our best reading of a graph when a similar figure was not available in a table.

Table A2: Unconditional Quantile Estimates of the Gender Wage Gap, Selected Quantiles

	10th Percentile			25th Percentile			75th Percentile			90th Percentile		
	Gap	Explained	Unexplained	Gap	Explained	Unexplained	Gap	Explained	Unexplained	Gap	Explained	Unexplained
Sample 1												
1981	0.395	0.154	0.240	0.479	0.139	0.340	0.415	0.110	0.304	0.438	0.150	0.288
1990	0.252	0.082	0.170	0.246	0.100	0.145	0.277	0.077	0.199	0.273	0.065	0.208
1999	0.228	0.083	0.145	0.251	0.051	0.200	0.243	0.057	0.185	0.288	0.057	0.231
2011	0.211	0.013	0.198	0.207	0.015	0.191	0.238	-0.011	0.248	0.233	-0.039	0.271
2015	0.092	0.031	0.061	0.146	0.000	0.146	0.186	-0.056	0.242	0.216	-0.069	0.285
Convergence (2015 – 1981)	-0.303	-0.123	-0.180	-0.333	-0.139	-0.194	-0.228	-0.166	-0.062	-0.222	-0.219	-0.003
Sample 2												
1981	0.455	0.238	0.217	0.521	0.198	0.323	0.438	0.129	0.309	0.414	0.158	0.256
1990	0.366	0.145	0.221	0.366	0.145	0.220	0.302	0.102	0.200	0.298	0.084	0.214
1999	0.316	0.124	0.192	0.317	0.081	0.236	0.266	0.070	0.195	0.302	0.052	0.251
2011	0.210	0.071	0.139	0.213	0.032	0.181	0.241	-0.011	0.253	0.256	-0.063	0.318
2015	0.162	0.073	0.089	0.178	0.036	0.141	0.208	-0.037	0.245	0.242	-0.056	0.297
Convergence (2015 – 1981)	-0.293	-0.165	-0.128	-0.343	-0.162	-0.181	-0.230	-0.165	-0.065	-0.172	-0.214	0.042
Sample 3												
1981	0.561	0.274	0.287	0.578	0.270	0.308	0.507	0.139	0.368	0.475	0.174	0.301
1990	0.438	0.211	0.227	0.459	0.189	0.270	0.348	0.124	0.224	0.311	0.108	0.203
1999	0.344	0.204	0.140	0.357	0.123	0.234	0.298	0.079	0.218	0.304	0.056	0.248
2011	0.237	0.108	0.129	0.235	0.059	0.176	0.252	0.003	0.248	0.281	-0.047	0.328
2015	0.200	0.103	0.097	0.202	0.054	0.148	0.218	-0.011	0.229	0.212	-0.025	0.236
Convergence (2015 – 1981)	-0.361	-0.171	-0.191	-0.376	-0.216	-0.159	-0.289	-0.151	-0.138	-0.263	-0.199	-0.064
Sample 4												
1981	0.617	0.286	0.332	0.586	0.303	0.283	0.519	0.169	0.349	0.508	0.203	0.305
1990	0.487	0.234	0.253	0.481	0.215	0.266	0.373	0.153	0.220	0.333	0.137	0.197
1999	0.351	0.222	0.128	0.356	0.134	0.222	0.304	0.096	0.208	0.315	0.070	0.245
2011	0.261	0.135	0.126	0.236	0.071	0.165	0.252	0.029	0.223	0.292	-0.020	0.313
2015	0.214	0.120	0.094	0.193	0.065	0.128	0.229	0.006	0.223	0.213	-0.011	0.224
Convergence (2015 – 1981)	-0.403	-0.166	-0.237	-0.393	-0.238	-0.155	-0.289	-0.163	-0.126	-0.295	-0.214	-0.081

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. Note that unconditional quantile coefficients do not precisely go through the sample quantiles, but that the quantiles are similar to the raw figures. Decomposition covariates correspond to those in our human capital specification. Sample 1 is full-time workers with 26 or more weeks of employment; Sample 2 adds part-time workers and workers with less than 26 weeks of employment but at least 100 hours; Sample 3 assigns remaining missing wages using adjacent PSID waves if wage data are available in those years; and Sample 4 imputes remaining missing wages using our human capital specification control set.

Table A3: Alternative Sample 4 Imputations I

	OLS (Mean)			Unconditional (Median)		
	Gap	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Panel A: Baseline						
1981	0.584	0.228	0.356	0.619	0.209	0.410
1990	0.433	0.187	0.246	0.462	0.188	0.274
1999	0.328	0.123	0.205	0.312	0.115	0.197
2011	0.256	0.053	0.203	0.251	0.043	0.208
2015	0.217	0.040	0.177	0.242	0.036	0.206
Panel B: Additional Labor Force Participation Variables						
1981	0.584	0.228	0.356	0.618	0.209	0.409
1990	0.433	0.187	0.246	0.462	0.188	0.274
1999	0.328	0.123	0.205	0.310	0.115	0.196
2011	0.254	0.053	0.201	0.251	0.043	0.208
2015	0.217	0.040	0.177	0.241	0.035	0.206
Panel C: Sample 3 Entrants						
1981	0.607	0.225	0.382	0.642	0.207	0.435
1990	0.445	0.188	0.257	0.482	0.189	0.293
1999	0.338	0.130	0.208	0.322	0.118	0.204
2011	0.261	0.057	0.204	0.261	0.047	0.214
2015	0.223	0.040	0.183	0.248	0.037	0.211
Panel D: Donut (≤ 20, ≥ 80)						
1981	0.576	0.235	0.340	0.605	0.192	0.413
1990	0.444	0.195	0.249	0.462	0.185	0.276
1999	0.333	0.140	0.194	0.308	0.110	0.198
2011	0.258	0.061	0.197	0.249	0.038	0.211
2015	0.218	0.054	0.164	0.239	0.034	0.205
Panel E: Random Decile Wages						
1981	0.581	0.233	0.348	0.621	0.209	0.412
1990	0.433	0.188	0.244	0.462	0.188	0.274
1999	0.325	0.127	0.199	0.311	0.115	0.197
2011	0.254	0.055	0.198	0.251	0.043	0.208
2015	0.216	0.042	0.174	0.242	0.037	0.205

Notes: This table provides results for sensitivity tests of our primary Sample 4 decomposition of the gender wage gap. Panel A assigns wages to Sample 4 entrants using our baseline specification. Panel B adds additional labor force supply variables to the wage decile assignment probit, including the number of children under six and marital status. Panel C estimates the original ordered probit model using only Sample 3 entrants. Panel D only imputes wages for individuals with a predicted probability of being below or above the median of less than or equal to 20 percent or greater than or equal to 80 percent, respectively. Panel E assigns within-decile wages based on a uniform random draw of a wage in each decile (probability weights remain the same). See Table 3 notes for additional sample descriptions.

Table A4: Alternative Sample 4 Imputations II

	OLS (Mean)			Unconditional (Median)		
	Gap	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Panel A: 10 Quantiles (Baseline)						
1981	0.584	0.228	0.356	0.619	0.209	0.410
1990	0.433	0.187	0.246	0.462	0.188	0.274
1999	0.328	0.123	0.205	0.312	0.115	0.197
2011	0.256	0.053	0.203	0.251	0.043	0.208
2015	0.217	0.040	0.177	0.242	0.036	0.206
Panel B: 8 Quantiles						
1981	0.582	0.228	0.355	0.621	0.209	0.411
1990	0.432	0.187	0.245	0.462	0.188	0.274
1999	0.329	0.123	0.205	0.311	0.115	0.196
2011	0.254	0.053	0.201	0.251	0.043	0.207
2015	0.217	0.039	0.178	0.242	0.035	0.207
Panel C: 4 Quantiles						
1981	0.578	0.226	0.352	0.618	0.210	0.408
1990	0.429	0.185	0.243	0.462	0.189	0.273
1999	0.327	0.121	0.206	0.311	0.116	0.196
2011	0.252	0.052	0.200	0.251	0.045	0.206
2015	0.219	0.037	0.182	0.243	0.035	0.207
Panel D: 2 Quantiles						
1981	0.567	0.223	0.344	0.621	0.213	0.408
1990	0.421	0.183	0.238	0.462	0.191	0.271
1999	0.324	0.117	0.207	0.312	0.118	0.194
2011	0.248	0.048	0.200	0.250	0.047	0.204
2015	0.214	0.034	0.180	0.242	0.037	0.205

Notes: This table provides results for sensitivity tests of our primary Sample 4 decomposition of the gender wage gap. Each panel reestimates our wage assignment probit using various partitions of the wage distribution. Panel A reproduces our primary specification that partitions the distribution into deciles. Panel B partitions into octiles, Panel C into quartiles, and Panel D into halves. See Table 3 notes for additional sample descriptions.

Table A5: Probit Error Rates

	Men Above Median, Imputation	Men Above Median, Real	Difference	Women Above Median, Imputation	Women Above Median, Real	Difference
Panel A: S1 Model, S2 Entrant Results (Probability Weights)						
1981	0.363	0.375	-0.011	0.422	0.387	0.035
1990	0.432	0.318	0.113	0.431	0.290	0.142
1999	0.452	0.495	-0.043	0.433	0.375	0.057
2011	0.392	0.308	0.084	0.450	0.381	0.069
2015	0.341	0.298	0.042	0.419	0.331	0.088
Panel B: S2 Model, S3 Entrant Results (Probability Weights)						
1981	0.354	0.243	0.112	0.431	0.295	0.137
1990	0.351	0.153	0.199	0.378	0.234	0.144
1999	0.394	0.359	0.035	0.405	0.315	0.090
2011	0.380	0.265	0.115	0.411	0.314	0.097
2015	0.391	0.227	0.164	0.368	0.260	0.109
Panel C: S2 Entrant Model, S3 Entrant Results (Probability Weights)						
1981	0.394	0.243	0.152	0.384	0.295	0.090
1990	0.273	0.153	0.121	0.251	0.234	0.017
1999	0.445	0.359	0.086	0.364	0.315	0.049
2011	0.330	0.265	0.066	0.343	0.314	0.029
2015	0.314	0.227	0.087	0.311	0.260	0.051

Notes: This table provides results for sensitivity tests of our primary Sample 4 decomposition of the gender wage gap. Specifically, this table compares the fraction of observations with an above median wage ("Above, Real") to that predicted by an ordered probit ("Above, Imputation"). "Above, Imputation" is the mean probability of being above the median (specifically, the mean sum of probability weights for deciles above the median). The ordered probit model is estimated on three different samples: Sample 1, Sample 2, and Sample 2 entrants. The results from the probit model are compared against two different samples: Sample 2 entrants and Sample 3 entrants. See text for more details. See Table 3 notes for additional sample descriptions.

Table A6: Alternative Sample 4 Imputations III

	OLS (Mean)			Unconditional (Median)		
	Gap	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Panel A: 10 Quantiles (Baseline)						
1981	0.584	0.228	0.356	0.619	0.209	0.410
1990	0.433	0.187	0.246	0.462	0.188	0.274
1999	0.328	0.123	0.205	0.312	0.115	0.197
2011	0.256	0.053	0.203	0.251	0.043	0.208
2015	0.217	0.040	0.177	0.242	0.036	0.206
Panel B: Error Adjusted Estimates						
1981	0.605	0.230	0.376	0.651	0.210	0.441
1990	0.444	0.188	0.255	0.477	0.192	0.285
1999	0.337	0.127	0.210	0.320	0.116	0.203
2011	0.264	0.054	0.209	0.263	0.045	0.218
2015	0.219	0.041	0.178	0.243	0.038	0.205

Notes: This table provides results for sensitivity tests of our primary Sample 4 decomposition of the gender wage gap. Panel A reproduces our baseline specification. In Panel B, we adjust the Sample 4 entrant probability weights based on the error rates reported in Table A5, Panel B, which corresponds to computing errors by predicting Sample 3 entrant wages using Sample 2. See the text for more details. See Table 3 notes for additional sample descriptions.

Table A7A: Bounding the Median Gender Wage Gap by Education and Experience Groups, Low Education

	Summary			*Worst Case*		*FOSD*		*Excl. Rest.*		*Excl. Rest. + FOSD*		*Summary*			*Worst Case*		*FOSD*		*Excl. Rest.*		*Excl. Rest. + FOSD*	
	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
	Panel A. Low Edu., Low Exp.											*Panel B. Low Edu., High Exp.*										
Sample 1 Base																						
1981	0.643	0.214	0.555	-0.972	2.336	0.122	2.005	-0.694	1.503	0.375	1.333	0.837	0.615	0.435	-0.016	0.909	0.343	0.785	0.025	0.883	0.366	0.759
				[-1.194, 2.424]		[-0.098, 2.034]		[-0.799, 1.605]		[0.209, 1.402]					[-0.121, 1.042]		[0.261, 0.915]		[-0.065, 0.946]		[0.333, 0.836]	
1990	0.602	0.263	0.248	-1.065	2.143	-0.216	1.652	-0.780	1.443	0.050	1.130	0.870	0.693	0.291	0.000	0.666	0.221	0.571	0.065	0.595	0.235	0.501
				[-1.406, 2.340]		[-0.495, 1.753]		[-0.962, 1.548]		[-0.077, 1.231]					[-0.143, 0.769]		[0.124, 0.664]		[-0.016, 0.676]		[0.200, 0.595]	
1999	0.689	0.317	0.376	-1.194	2.119	0.114	1.863	-0.222	1.427	0.486	1.352	0.862	0.685	0.217	-0.159	0.542	0.120	0.480	-0.065	0.467	0.202	0.454
				[-1.410, 2.229]		[-0.138, 1.962]		[-0.338, 1.494]		[0.325, 1.401]					[-0.253, 0.661]		[0.058, 0.588]		[-0.134, 0.561]		[0.149, 0.549]	
2011	0.423	0.289	0.447	-2.846	3.086	-1.594	1.994	-1.172	1.860	0.035	1.418	0.716	0.611	0.187	-0.479	0.905	-0.067	0.625	-0.300	0.680	0.190	0.562
				[-2.846, 3.086]		[-1.661, 2.100]		[-2.271, 2.352]		[-1.305, 1.514]					[-0.593, 1.053]		[-0.148, 0.764]		[-0.370, 0.816]		[0.030, 0.675]	
2015	0.548	0.325	0.240	-1.363	2.537	-0.345	1.863	-0.743	1.341	0.091	1.132	0.741	0.577	0.193	-0.481	0.945	-0.001	0.735	-0.415	0.816	0.058	0.677
				[-2.170, 2.936]		[-1.213, 1.898]		[-0.841, 1.520]		[-0.063, 1.193]					[-0.786, 1.161]		[-0.122, 0.954]		[-0.503, 0.900]		[-0.012, 0.771]	
Convergence Bounds				-3.699	3.509	-2.350	1.741	-2.246	2.035	-1.241	0.757				-1.390	0.961	-0.785	0.393	-1.297	0.791	-0.701	0.311
Confidence Region				[-4.505, 3.954]		[-3.095, 1.937]		[-2.344, 2.124]		[-1.345, 0.899]					[-1.731, 1.201]		[-0.968, 0.607]		[-1.348, 0.909]		[-0.776, 0.386]	
Sample 2 Base																						
1981	0.758	0.410	0.584	-1.260	2.259	0.323	2.088	-0.813	1.882	0.494	1.823	0.890	0.743	0.464	0.228	0.734	0.412	0.662	0.262	0.681	0.507	0.635
				[-1.376, 2.378]		[0.168, 2.117]		[-0.873, 1.965]		[0.366, 1.872]					[0.156, 0.807]		[0.349, 0.738]		[0.200, 0.747]		[0.414, 0.695]	
1990	0.716	0.470	0.277	-1.077	2.182	0.045	1.813	-0.707	1.284	0.281	1.141	0.897	0.813	0.342	0.139	0.544	0.282	0.481	0.155	0.490	0.453	0.441
				[-1.276, 2.274]		[-0.147, 1.923]		[-0.844, 1.424]		[0.113, 1.242]					[0.074, 0.633]		[0.206, 0.561]		[0.119, 0.569]		[0.287, 0.507]	
1999	0.738	0.501	0.407	-2.919	2.113	0.235	1.874	-0.099	0.920	0.513	0.844	0.909	0.827	0.228	0.031	0.396	0.165	0.362	0.146	0.361	0.218	0.352
				[-3.089, 2.200]		[-0.012, 1.973]		[-0.556, 1.352]		[0.377, 1.346]					[-0.047, 0.481]		[0.101, 0.435]		[0.036, 0.406]		[0.167, 0.395]	
2011	0.547	0.466	0.447	-2.416	2.482	-0.509	1.910	-0.813	1.540	0.291	1.297	0.777	0.720	0.196	-0.237	0.609	0.020	0.418	-0.113	0.501	0.125	0.410
				[-3.418, 3.136]		[-1.511, 2.101]		[-1.025, 1.706]		[-0.090, 1.418]					[-0.313, 0.728]		[-0.060, 0.527]		[-0.181, 0.562]		[0.082, 0.478]	
2015	0.650	0.502	0.254	-1.398	2.264	-0.115	1.858	-0.353	0.822	0.206	0.755	0.796	0.703	0.190	-0.190	0.670	0.026	0.494	-0.070	0.360	0.155	0.323
				[-1.519, 2.336]		[-0.268, 1.871]		[-0.639, 1.395]		[0.093, 1.309]					[-0.365, 0.784]		[-0.058, 0.626]		[-0.196, 0.565]		[0.084, 0.543]	
Convergence Bounds				-3.657	3.524	-2.203	1.535	-2.234	1.635	-1.617	0.261				-0.923	0.442	-0.636	0.082	-0.751	0.098	-0.480	-0.183
Confidence Region				[-3.796, 3.670]		[-2.354, 1.666]		[-2.424, 2.089]		[-1.711, 0.813]					[-1.096, 0.586]		[-0.742, 0.221]		[-0.865, 0.320]		[-0.553, 0.032]	
Sample 3 Base																						
1981	0.929	0.693	0.553	0.204	0.903	0.475	0.812	0.359	0.774	0.608	0.774	0.963	0.891	0.489	0.422	0.581	0.468	0.567	0.587	0.567	0.599	0.563
				[0.118, 0.978]		[0.392, 0.934]		[0.231, 0.873]		[0.487, 0.873]					[0.359, 0.642]		[0.422, 0.620]		[0.449, 0.616]		[0.495, 0.609]	
1990	0.914	0.744	0.352	0.021	0.670	0.262	0.606	0.272	0.592	0.483	0.570	0.976	0.938	0.353	0.305	0.399	0.341	0.388	0.462	0.384	0.462	0.384
				[-0.070, 0.766]		[0.175, 0.687]		[0.085, 0.659]		[0.335, 0.654]					[0.238, 0.453]		[0.273, 0.444]		[0.310, 0.415]		[0.358, 0.415]	
1999	0.882	0.746	0.448	0.199	0.699	0.405	0.614	0.406	0.579	0.520	0.557	0.975	0.921	0.246	0.186	0.330	0.230	0.317	0.209	0.246	0.252	0.246
				[-0.023, 0.841]		[0.239, 0.741]		[0.249, 0.666]		[0.406, 0.655]					[0.121, 0.393]		[0.164, 0.382]		[0.154, 0.333]		[0.201, 0.333]	
2011	0.796	0.699	0.370	-0.072	0.847	0.227	0.632	0.311	0.610	0.362	0.518	0.927	0.864	0.178	0.054	0.343	0.132	0.281	0.109	0.153	0.199	0.148
				[-0.277, 1.010]		[0.069, 0.791]		[0.173, 0.730]		[0.299, 0.648]					[-0.007, 0.435]		[0.068, 0.378]		[0.056, 0.321]		[0.134, 0.320]	
2015	0.797	0.715	0.207	-0.217	0.591	0.056	0.461	-0.017	0.226	0.206	0.212	0.892	0.855	0.237	-0.001	0.423	0.130	0.344	0.165	0.219	0.250	0.219
				[-0.372, 0.774]		[-0.074, 0.573]		[-0.128, 0.349]		[0.117, 0.331]					[-0.056, 0.522]		[0.032, 0.436]		[0.060, 0.315]		[0.168, 0.315]	
Convergence Bounds				-1.120	0.388	-0.756	-0.014	-0.791	-0.134	-0.568	-0.395				-0.582	0.001	-0.437	-0.124	-0.402	-0.368	-0.313	-0.380
Confidence Region				[-1.281, 0.606]		[-0.944, 0.121]		[-0.930, 0.041]		[-0.694, -0.249]					[-0.675, 0.112]		[-0.560, -0.021]		[-0.499, -0.251]		[-0.388, -0.288]	

Notes: See notes for Tables 6 and 7.

Table A7B: Bounding the Median Gender Wage Gap by Education and Experience Groups, High Education

	Summary			*Worst Case*		*FOSD*		*Excl. Rest.*		*Excl. Rest. + FOSD*		*Summary*			*Worst Case*		*FOSD*		*Excl. Rest.*		*Excl. Rest. + FOSD*	
	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Male Coverage	Female Coverage	Log Gap	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
	Panel C. High Edu., Low Exp.											*Panel D. High Edu., High Exp.*										
Sample 1 Base																						
1981	0.760	0.375	0.234	-0.904	1.534	0.067	1.328	-0.605	1.393	0.261	1.223	0.940	0.722	0.401	0.158	0.644	0.367	0.619	0.256	0.599	0.413	0.587
				[-0.999, 1.658]		[-0.085, 1.469]		[-0.716, 1.425]		[0.103, 1.311]					[0.069, 0.734]		[0.292, 0.707]		[0.157, 0.660]		[0.331, 0.642]	
1990	0.809	0.408	0.157	-1.034	1.689	0.017	1.524	-0.547	1.053	0.152	1.012	0.909	0.711	0.251	-0.014	0.549	0.191	0.493	0.079	0.458	0.398	0.440
				[-1.166, 1.757]		[-0.147, 1.596]		[-0.700, 1.234]		[0.053, 1.187]					[-0.093, 0.614]		[0.114, 0.559]		[0.007, 0.507]		[0.289, 0.490]	
1999	0.796	0.473	0.124	-1.343	1.637	-0.014	1.560	-0.536	0.671	0.093	0.632	0.900	0.730	0.273	-0.048	0.624	0.194	0.549	0.037	0.533	0.256	0.506
				[-1.391, 1.990]		[-0.127, 1.917]		[-0.753, 0.912]		[0.005, 0.880]					[-0.130, 0.685]		[0.126, 0.604]		[-0.037, 0.614]		[0.211, 0.579]	
2011	0.768	0.482	0.103	-2.107	2.170	-0.103	1.990	-0.705	0.903	0.006	0.786	0.839	0.664	0.281	-0.253	0.817	0.116	0.645	-0.179	0.755	0.175	0.628
				[-2.168, 2.248]		[-0.206, 2.063]		[-0.986, 1.143]		[-0.079, 1.054]					[-0.363, 0.876]		[0.029, 0.728]		[-0.249, 0.788]		[0.128, 0.694]	
2015	0.789	0.527	0.111	-0.901	1.121	-0.059	0.902	-0.539	0.659	0.085	0.601	0.879	0.731	0.235	-0.130	0.588	0.140	0.486	-0.034	0.503	0.205	0.466
				[-1.555, 1.798]		[-0.149, 1.600]		[-0.680, 0.856]		[0.016, 0.777]					[-0.220, 0.689]		[0.059, 0.575]		[-0.124, 0.582]		[0.162, 0.532]	
Convergence Bounds				-2.435	2.026	-1.387	0.835	-1.932	1.265	-1.138	0.340				-0.774	0.430	-0.480	0.119	-0.633	0.247	-0.382	0.053
Confidence Region				[-3.085, 2.721]		[-1.559, 1.521]		[-2.036, 1.422]		[-1.214, 0.547]					[-0.897, 0.576]		[-0.600, 0.242]		[-0.726, 0.388]		[-0.429, 0.147]	
Sample 2 Base																						
1981	0.900	0.597	0.266	-0.230	0.839	0.206	0.736	-0.090	0.629	0.297	0.536	0.965	0.835	0.427	0.314	0.520	0.415	0.511	0.348	0.475	0.441	0.470
				[-0.389, 0.985]		[0.119, 0.913]		[-0.190, 0.748]		[0.188, 0.690]					[0.206, 0.611]		[0.320, 0.594]		[0.288, 0.537]		[0.365, 0.532]	
1990	0.915	0.669	0.255	-0.176	0.682	0.179	0.631	0.080	0.509	0.284	0.489	0.944	0.867	0.283	0.148	0.428	0.254	0.396	0.177	0.348	0.413	0.346
				[-0.302, 0.831]		[0.061, 0.771]		[-0.047, 0.624]		[0.179, 0.622]					[0.081, 0.484]		[0.165, 0.450]		[0.146, 0.425]		[0.308, 0.421]	
1999	0.889	0.703	0.154	-0.226	0.519	0.093	0.446	-0.082	0.352	0.203	0.318	0.938	0.866	0.306	0.144	0.459	0.262	0.410	0.206	0.402	0.310	0.388
				[-0.302, 0.636]		[0.010, 0.584]		[-0.188, 0.470]		[0.085, 0.444]					[0.078, 0.528]		[0.188, 0.485]		[0.151, 0.463]		[0.259, 0.447]	
2011	0.873	0.676	0.094	-0.357	0.547	0.014	0.439	-0.187	0.424	0.069	0.344	0.886	0.804	0.283	0.006	0.582	0.173	0.448	0.090	0.455	0.349	0.401
				[-0.454, 0.641]		[-0.066, 0.548]		[-0.325, 0.532]		[0.008, 0.481]					[-0.083, 0.687]		[0.098, 0.562]		[0.037, 0.556]		[0.214, 0.508]	
2015	0.869	0.732	0.118	-0.245	0.453	0.017	0.359	-0.109	0.337	0.093	0.305	0.905	0.846	0.270	0.077	0.482	0.211	0.412	0.110	0.439	0.310	0.407
				[-0.324, 0.567]		[-0.059, 0.428]		[-0.177, 0.420]		[0.041, 0.393]					[0.001, 0.567]		[0.132, 0.473]		[0.066, 0.492]		[0.230, 0.448]	
Convergence Bounds				-1.084	0.683	-0.719	0.153	-0.737	0.427	-0.443	0.007				-0.443	0.168	-0.300	-0.004	-0.365	0.090	-0.159	-0.034
Confidence Region				[-1.259, 0.897]		[-0.907, 0.269]		[-0.859, 0.558]		[-0.590, 0.126]					[-0.566, 0.318]		[-0.416, 0.109]		[-0.419, 0.164]		[-0.255, 0.042]	
Sample 3 Base																						
1981	0.998	0.860	0.320	0.201	0.458	0.318	0.458	0.239	0.347	0.429	0.347	1.000	0.936	0.441	0.407	0.478	0.441	0.478	0.441	0.377	0.462	0.377
				[0.116, 0.551]		[0.242, 0.542]		[0.163, 0.470]		[0.315, 0.470]					[0.324, 0.539]		[0.363, 0.539]		[0.355, 0.489]		[0.400, 0.489]	
1990	0.992	0.886	0.311	0.200	0.432	0.306	0.408	0.248	0.392	0.330	0.392	0.995	0.974	0.307	0.286	0.319	0.303	0.317	0.397	0.307	0.408	0.307
				[0.089, 0.526]		[0.205, 0.518]		[0.146, 0.459]		[0.242, 0.459]					[0.209, 0.394]		[0.239, 0.391]		[0.291, 0.362]		[0.327, 0.362]	
1999	0.987	0.896	0.218	0.113	0.298	0.206	0.290	0.146	0.228	0.307	0.228	0.983	0.957	0.324	0.280	0.363	0.315	0.353	0.321	0.192	0.346	0.191
				[0.032, 0.375]		[0.123, 0.367]		[0.082, 0.323]		[0.204, 0.323]					[0.210, 0.426]		[0.254, 0.411]		[0.263, 0.320]		[0.288, 0.318]	
2011	0.960	0.870	0.142	-0.024	0.279	0.110	0.253	0.059	0.242	0.143	0.230	0.970	0.918	0.293	0.194	0.377	0.255	0.352	0.227	0.162	0.334	0.139
				[-0.091, 0.366]		[0.054, 0.337]		[-0.016, 0.315]		[0.102, 0.296]					[0.121, 0.463]		[0.183, 0.430]		[0.182, 0.343]		[0.242, 0.318]	
2015	0.963	0.885	0.125	0.027	0.243	0.106	0.217	0.089	0.218	0.153	0.210	0.967	0.926	0.265	0.200	0.359	0.254	0.334	0.333	0.323	0.335	0.323
				[-0.040, 0.330]		[0.047, 0.304]		[0.028, 0.288]		[0.110, 0.279]					[0.124, 0.431]		[0.183, 0.409]		[0.237, 0.384]		[0.283, 0.384]	
Convergence Bounds				-0.431	0.042	-0.352	-0.101	-0.258	-0.021	-0.194	-0.220				-0.278	-0.048	-0.223	-0.107	-0.044	-0.118	-0.042	-0.139
Confidence Region				[-0.537, 0.176]		[-0.451, 0.015]		[-0.376, 0.072]		[-0.308, -0.121]					[-0.375, 0.072]		[-0.320, -0.002]		[-0.188, -0.027]		[-0.148, -0.062]	

Notes: See notes for Tables 6 and 7.

Table A8: Identification at Infinity Gender Wage Gap Estimates and Decomposition, Female Dummy Method

			OLS (Mean)			Unconditional (Median)		
	Male Count	*Female Count*	*Gap*	*Explained*	*Unexplained*	*Gap*	*Explained*	*Unexplained*
Panel A: Infinity Sample Based on Wave-Specific Probit								
1981	1598	185	0.305	-0.070	0.375	0.302	-0.064	0.366
1990	2062	381	0.067	-0.139	0.206	0.059	-0.130	0.189
1999	1887	603	0.091	-0.085	0.176	0.046	-0.104	0.150
2011	826	179	-0.097	-0.306	0.209	-0.147	-0.353	0.206
2015	1165	352	-0.062	-0.302	0.239	-0.081	-0.314	0.233
Panel B: Infinity Sample Based on 1981 Probit								
1981	1598	185	0.305	-0.070	0.375	0.302	-0.064	0.366
1990	2056	534	0.120	-0.066	0.186	0.078	-0.059	0.137
1999	1885	623	0.098	-0.071	0.169	0.044	-0.089	0.133
2011	1437	587	0.093	-0.140	0.233	0.073	-0.131	0.204
2015	1503	554	0.080	-0.157	0.237	0.076	-0.161	0.237

Notes: Data from the 1981, 1990, 1999, 2011, and 2015 waves of the PSID and includes non-immigrant individuals aged 25 to 54. The sample is limited to Sample 1 observations (full-time workers with 26 or more weeks of employment) with probability of full-time year-round employment equal to or above 0.8. In Panel A the employment probability probit is estimated separately for each wave. For Panel B, this probability is calculated using the 1981 PSID wave, with estimated probability coefficients then applied to other waves. Wage equation and employment probability equation variables are those listed as the human capital specification in the text. Decompositions computed with female dummy method. Note that unconditional quantile coefficients do not precisely go through the sample median, but that the predicted medians are very similar to sample medians.