# Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps

Claudia Olivetti, *Boston University*

Barbara Petrongolo, *London School of Economics, CEP, CEPR, and IZA*

We analyze gender wage gaps correcting for sample selection induced by nonemployment. We recover wages for the nonemployed using alternative imputation techniques, simply requiring assumptions on the position of imputed wages with respect to the median. We obtain higher median wage gaps on imputed rather than actual wage distributions for several OECD countries. However, this difference is small in the United States, the United Kingdom, and most central and northern EU countries and becomes sizable in southern EU countries, where gender employment gaps are high. Selection correction explains nearly half of the observed negative correlation between wage and employment gaps.

## I. Introduction

There is substantial international variation in gender pay gaps, from around 30 log points in the United States and the United Kingdom to

We wish to thank Ivan Fernandez-Val, Larry Katz, Kevin Lang, Alan Manning, and Steve Pischke for their very helpful suggestions. We also acknowledge comments from seminars at several institutions, as well as from presentations at the Bank of Portugal Annual Conference 2005, the SOLE/EALE Conference 2005, the Conference in Honor of Reuben Gronau 2005, and the National Bureau of Economic Research Summer Institute 2006. Olivetti aknowledges the Radcliffe Institute for Advanced

[*Journal of Labor Economics*, 2008, vol. 26, no. 4]

0734-306X/2008/2604-0006$10.00

between 10 and 25 log points in a number of central and northern European countries, down to an average below 10 log points in southern Europe. International differences in overall wage dispersion are typically found to play a role in explaining the variation in gender pay gaps (Blau and Kahn 1996, 2003). The idea is that a given level of dissimilarities between the characteristics of working men and women translates into a higher gender wage gap the higher the overall level of wage inequality. However, the Organization for Economic Cooperation and Development (OECD 2002, chart 2.7) shows that, while differences in the wage structure do explain an important portion of the international variation in gender wage gaps, the inequality-adjusted wage gap in southern Europe remains substantially lower than such gaps in the rest of Europe and the United States.

In this article we argue that, in addition to differences in wage inequality, and therefore in the returns associated to characteristics of working men and women, a significant portion of the international variation in gender wage gaps may be explained by differences in characteristics themselves, whether observed or unobserved. This idea is supported by the striking international variation in employment gaps, ranging from 10 percentage points in the United States, the United Kingdom, and Scandinavian countries to 15–25 points in northern and central Europe, up to 30–40 points in southern Europe and Ireland (see fig. 1). If selection into employment is nonrandom, it makes sense to worry about the way in which selection may affect the resulting gender wage gap. In particular, if women who are employed tend to have relatively high-wage characteristics, low female employment rates may become consistent with low gender wage gaps simply because low-wage women would not feature in the observed wage distribution. This idea could thus be well suited to explain the negative correlation between gender wage and employment gaps that we observe in the data.

Different patterns of employment selection across countries may in turn stem from a number of factors. First, there may be international differences in labor supply behavior and in particular in the role of household composition and/or social norms in affecting participation. Second, labor demand mechanisms, including social attitudes toward female employment and their potential effects on employer choices, may be at work, affecting both the arrival rate and the level of wage offers of the two genders. Finally, institutional differences in labor markets regarding unionization and minimum wages may truncate the wage distribution at different points in different countries, affecting both the composition of

Studies for financial support during the early stages of the project. Petrongolo aknowledges the ESRC for financial support to the Centre for Economic Performance.
Contact the corresponding author, Barbara Petrongolo, at b.petrongolo@lse.ac.uk.

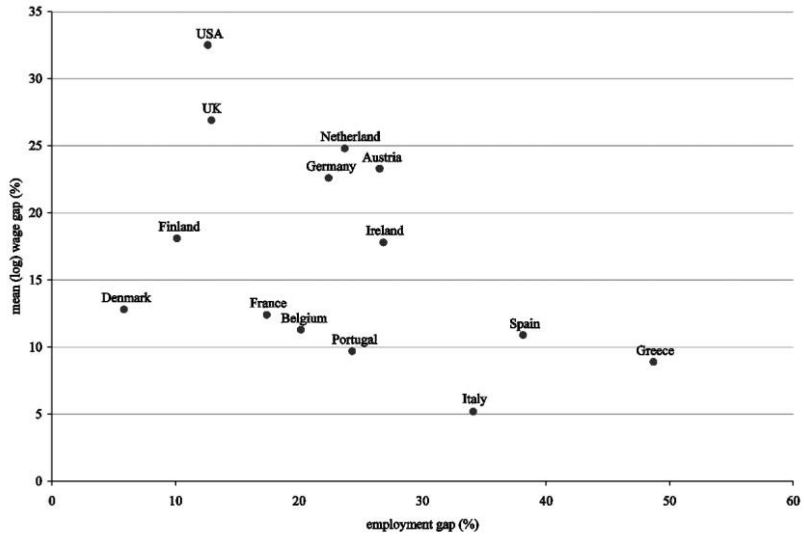


FIG. 1.—Gender gaps in mean (log) hourly wages and in employment, 1999. Coefficient of correlation: −0.474.

employment and the observed wage distribution. In this article we will be agnostic as regards the separate role of these factors in shaping gender gaps, and aim at recovering alternative measures of selection-corrected gender wage gaps.

Although there exist substantial literatures on gender wage gaps, on the one hand, and gender employment, unemployment and participation gaps, on the other hand,[1] to our knowledge the variation in both quantities and prices in the labor market has not been simultaneously exploited to understand important differences in gender gaps across countries. In this article we claim that the international variation in gender employment gaps can indeed shed some light on well-known cross-country differences in gender wage gaps. We will explore this view by estimating selection-corrected wage gaps.

To analyze gender wage gaps across countries, allowing for sample selection induced by nonemployment, we recover information on wages for those not working in a given year using alternative imputation tech-

[1] See Altonji and Blank (1999) for an overall survey on both employment and gender gaps for the United States, Blau and Kahn (2003) for international comparisons of gender wage gaps, and Azmat, Güell, and Manning (2006) for international comparisons of unemployment gaps.

niques.[2] Our approach is closely related to that of Johnson, Kitamura, and Neal (2000) and Neal (2004), and it simply requires assumptions on the position of the imputed wage observations with respect to the median. Importantly, it does not require assumptions on the actual level of missing wages, as is typically required in the matching approach, nor does it require the arbitrary exclusion restrictions often invoked in two-stage Heckman sample selection correction models.

We then estimate median wage gaps on the sample of employed workers and on a sample enlarged with wage imputation for the nonemployed. The impact of selection into work on estimated wage gaps is assessed by comparing estimates obtained under alternative sample inclusion rules. The attractive feature of median regressions is that the results are only affected by the position of wage observations with respect to the median and not by specific values of imputed wages. If missing wage observations are correctly imputed on the side of the median where they belong, then median regressions retrieve the true parameters of interest.

Imputation can be performed in several ways, and our alternative imputation methods will address slightly different economic mechanisms of selection. First, we use panel data and, for all those not in work in some base year, we search backward and forward to recover wage observations from the nearest wave in the sample. This implicitly assumes that an individual's position with respect to the base-year median can be signaled by her wage from the nearest wave. As imputation is simply driven by wages observed in other waves, we are in practice allowing for selection on unobservables. Estimates based on this procedure tell what level of the gender wage gap we would observe if the nonemployed earned "similar" wages to those earned when they were employed, where "similar" here means on the same side of the base-year median.

While this imputation method arguably uses the minimum set of potentially arbitrary assumptions, it cannot provide wage information on individuals who never work during the sample period. In order to recover wages also for those never observed in work, we use observable characteristics of the nonemployed to make educated guesses concerning their position with respect to the median. In this case we are allowing for selection on observable characteristics only, assuming that the nonemployed would earn wages "similar" to the wages of the employed with matching characteristics, where again "similar" means on the same side of the base-year median. Having done this, earlier or later wage obser-

[2] We do not attempt to provide a structural model of wage determination that would in principle characterize general equilibrium effects of sample selection but would do so at the cost of making assumptions on production technologies involving male and female work. We are simply trying to estimate the gender wage gap correcting for sample selection.

vations for those with imputed wages in the base year can shed light on the goodness of our imputation methods.

We next use probability models for assigning individuals on either side of the median of the wage distribution. To do this we fit a probit model for the probability of an employed worker's belonging above their gender-specific median, based on education and labor market experience, and obtain predicted probabilities for the nonemployed. We then construct an imputed sample using such predicted probabilities as sampling weights.

We complete our set of results by estimating bounds to the distribution of wages (see Manski 1994), using either the actual or the imputed wage distribution in turn. Bounds computed using the observed wage distribution are interesting because they show that all our wage gap estimates based on imputation do fall within these bounds. When the imputed wage distribution is used, the increase in the proportion of individuals with a wage (actual or imputed) allows us to tighten the bounds, as predicted by the theory.

In our study we use panel data sets that are as comparable as possible across countries, namely, the Panel Study of Income Dynamics (PSID) for the United States and the European Community Household Panel Survey (ECHPS) for Europe. We consider the period 1994–2001, which is the longest time span for which data are available for all countries. Our estimates on these data deliver higher median wage gaps on imputed rather than actual wage distributions for most countries and across alternative imputation methods. This implies, as one would have expected, that women tend on average to be more positively selected into work than men. However, the difference between actual and potential wage gaps is small in the United States, the United Kingdom, and most central and northern European countries, and it becomes sizable in southern Europe, where the gender employment gap is highest. Under our most conservative correction, sample selection into employment explains nearly half of the observed negative correlation between gender wage and employment gaps. In particular, in Spain, Italy, Portugal, and Greece, the median wage gap on the imputed wage distribution reaches levels closely comparable to those of the United States and of other central and northern European countries.

Our results thus show that, while the raw wage gap is much higher in Anglo Saxon countries than in southern Europe, the reason is probably not to be found in more equal pay treatment for women in the latter group of countries but mainly in a different process of selection into employment. Female participation rates in Catholic countries and Greece are low and and are concentrated among high-wage women. Having corrected for lower participation rates, the wage gap there widens to levels similar to those of other European countries and the United States.

The article is organized as follows. Section II discusses the related literature. Section III describes the data sets used and presents descriptive evidence on gender gaps. Section IV describes our imputation methodologies. Section V estimates raw median gender wage gaps on actual and imputed wage distributions, to illustrate how alternative sample selection rules affect the estimated gaps. Conclusions are brought together in Section VI.

## II. Related Work

The importance of selectivity biases in making wage comparisons has long been recognized since seminal work by Gronau (1974) and Heckman (1974, 1979, 1980). The current literature contains a number of country-level studies that estimate selection-corrected wage gaps across genders or ethnic groups, based on a variety of correction methodologies. Among studies that are more closely related to our article, Neal (2004) estimates the gap in potential earnings between black and white women in the United States by fitting median regressions on imputed wage distributions, using alternative methods of wage imputation for women nonemployed in 1990. He finds that the gap between potential earnings of white and black women is at least 60% higher than the gap in actual earnings, thus revealing that black women are more positively selected into work. Using both wage imputation and matching techniques, Chandra (2003) finds that the wage gap between black and white U.S. males is also understated, due to selective withdrawal of black men from the labor force during the 1970s and 1980s.[3]

Turning to gender wage gaps, Blau and Kahn (2006) study changes in the U.S. gender wage gap between 1979 and 1998 and find that sample selection implies that the 1980s gains in women's relative wage offers were overstated and that selection may also explain part of the slowdown in convergence between male and female wages in the 1990s. Their approach is based on wage imputation for those not in work, along the lines of Neal (2004). Mulligan and Rubinstein (2008) also argue that the narrowing of the gender wage gap in the United States during 1975–2001 may be a direct impact of progressive selection into employment of high-wage women, in turn attracted by widening within-gender wage dispersion. Correction for selection into work is implemented here using a two-stage Heckman (1979) selection model. The authors show that, while in the 1970s the gender selection bias was negative, that is, nonemployed women had higher earnings potential than working women, it became positive in the mid 1980s.[4]

[3] See also Blau and Beller (1992) and Juhn (2003) for earlier use of matching techniques in the study of selection-corrected race gaps.

[4] Earlier studies that discuss the importance of changing characteristics of the female workforce in explaining the dynamics of the gender wage gap in the United States include O'Neil (1985), Smith and Ward (1989), and Goldin (1990).

Related work on European countries includes Beblo et al. (2003), Albrecht, van Vuuren, and Vroman (2004), and Blundell et al. (2007). Blundell et al. examine changes in the distribution of wages in the United Kingdom during the period 1978–2000. They allow for the impact of nonrandom selection into work by using bounds to the latent wage distribution according to the procedure proposed by Manski (1994). Bounds are first constructed based on the worst-case scenario and then progressively tightened using restrictions motivated by economic theory. Features of the resulting wage distribution are then analyzed, including overall wage inequality, returns to education, and gender wage gaps. Albrecht et al. estimate gender wage gaps in the Netherlands having corrected for selection of women into market work according to the Buchinsky's (1998) semiparametric method for quantile regressions, and they find evidence of strong positive selection into full-time employment. Finally, Beblo et al. show selection-corrected wage gaps for Germany using both the Heckman (1979) and the Lewbel (2007) two-stage selection models. They find that correction for selection has an ambiguous impact on gender wage gaps in Germany, depending on the method used.

Interestingly, most studies find that correction for selection has important consequences for our assessment of gender wage gaps. At the same time, none of these studies use data for southern European countries, where employment rates of women are lowest and thus the selection issue should be most relevant. In this article we use data for the United States and for a representative group of European countries to investigate how nonrandom selection into work may affect international comparisons of gender wage gaps.

## III. Data

### A. The PSID

Our analysis for the United States is based on the Michigan Panel Study of Income Dynamics (PSID). This is a longitudinal survey of a representative sample of U.S. individuals and their households. It has been ongoing since 1968. The data were collected annually through 1997 and every other year after 1997. In order to ensure consistency with European data, we use six waves from the PSID, from 1994 to 2001. We restrict our analysis to individuals aged 25–54, having excluded the self-employed, full-time students, and individuals in the armed forces.[5]

[5] The exclusion of self-employed individuals may require some justification in so far as the incidence of self-employment varies importantly across genders and countries, as well as the associated earnings gap. However, the available definition of income for the self-employed is not comparable to the one we are using for the employees, and the number of observations for the self-employed is very limited for European countries. Both these factors prevent us from including the self-employed in our analysis.

The wage concept that we use throughout the analysis is the gross hourly wage. This is given by annual labor income divided by annual hours worked in the calendar year before the interview date. Employed workers are defined as those with positive hours worked in the previous year.

The characteristics that we exploit for wage imputation for the nonemployed are human capital variables, spouse income, and nonemployment status, that is, unemployed versus out of the labor force. Human capital is proxied by education and work experience controls. Ethnic origin is not included here as information on ethnicity is not available for the European sample. We consider three broad educational categories: less than high school, high school completed, and college completed. They include individuals who have completed less than 12 years of schooling, between 12 and 15 years of schooling, and at least 16 years of schooling, respectively. This categorization of the years of schooling variable is chosen for consistency with the definition of education in the ECHPS, which does not provide information on completed years of schooling but only on recognized qualifications.

Information on work experience refers to years of actual labor market experience (either full- or part-time) since the age of 18. When individuals first join the PSID sample as a head or a wife (or cohabitor), they are asked how many years they worked since age 18 and how many of these years involved full-time work. These two questions are also asked retrospectively in 1974 and 1985, irrespective of the year in which respondents had joined the sample. The answers to these questions are used to construct a measure for actual work experience, following the procedure of Blau and Kahn (2006). Given the initial values reported, we update work experience information for the years of interest using the longitudinal work history file from the PSID. For example, in order to construct the years of actual experience in 1994 for an individual who was in the survey in 1985, we add to the number of years of experience reported in 1985 the number of years between 1985 and 1994 during which they worked a positive number of hours.[6] This procedure allows us to construct the full work experience in each year until 1997. As the survey became biannual after 1997, there is no information on the number of hours worked by individuals between 1997 and 1998 and between 1999 and 2000. We fill missing work experience information for 1998 following again Blau and Kahn (2006). In particular, we use the 1999 sample to estimate logit models for positive hours in the previous year and in the year preceding the 1997 survey, separately for males and females. The

[6] The measure of actual experience used here includes both full-time and part-time work experience, as this is better comparable to the measure of experience available from the ECHPS.

explanatory variables are race, schooling, experience, a marital status indicator, and variables for the number of children aged 0–2, 3–5, 6–10, and 11–15 who are living in the household at the time of the interview. Work experience in the missing year is obtained as the average of the predicted values in the 1999 logit and the 1997 logit. We repeat the same steps for filling missing work experience information in 2000.

Spouse income is constructed as the sum of total labor and business income in unincorporated enterprises both for spouses and cohabitors of respondents. Finally, the reason for nonemployment, that is, unemployment versus inactivity, is obtained from self-reported information on employment status.

## B. The ECHPS

Data for European countries are drawn from the European Community Household Panel Survey (ECHPS).[7] This is an unbalanced household-based panel survey, containing annual information on a few thousand households per country during the period 1994–2001.[8] The ECHPS has the advantage that it asks a consistent set of questions across the 15 member states of the preenlargement European Union. The employment section of the survey contains information on the jobs held by members of selected households, including wages and hours of work. The household section allows us to obtain information on the family composition of respondents. We exclude Sweden and Luxembourg from our country set as wage information is unavailable for Sweden in all waves and unavailable for Luxembourg after 1996.

As for the United States, we restrict our analysis of wages to individuals aged 25–54 as of the survey date, and we exclude the self-employed, those in full-time education, and those in the military. The definitions of variables used replicates quite closely those used for the United States.

Hourly wages are computed as gross weekly wages divided by weekly usual working hours. The education categories used are less than upper secondary high school, upper secondary school completed, and higher education. These correspond to ISCED 0–2, 3, and 5–7, respectively. Unfortunately, no information on actual experience is available in the

[7] Previous work using ECHPS data for international comparisons of gender gaps include the OECD (2002) survey and Arulampalam, Booth, and Bryan (2007), who study the variation in gender pay gaps across quantiles of the wage distribution in 11 EU countries.

[8] The initial sample sizes are as follows: Austria: 3,380; Belgium: 3,490; Denmark: 3,482; Finland: 4,139; France: 7,344; Germany: 11,175; Greece: 5,523; Ireland: 4,048; Italy: 7,115; Luxembourg: 1,011; Netherlands: 5,187; Portugal: 4,881; Spain: 7,206; Sweden: 5,891; United Kingdom: 10,905. These figures are the number of households included in the first wave for each country, which corresponds to 1995 for Austria, 1996 for Finland, 1997 for Sweden, and 1994 for all other countries.

ECHPS, and we use a measure of potential work experience, obtained as the current age of respondents minus the age at which they started their working life. Spouse income is computed as the sum of labor and nonlabor annual income for spouses or cohabitors of respondents. Finally, unemployment status is determined using self-reported information on the main activity status. (Descriptive statistics for both the U.S. and the EU samples are reported in table A1 in the online appendix.)

### C. Descriptive Evidence on Gender Gaps

Figure 1 plots raw gender gaps in log gross hourly wages and employment rates for all countries in our sample. All estimates refer to 1999, which will be the base year in our analysis. At the risk of some oversimplification, one can classify countries into three broad categories according to their levels of gender wage gaps. In the United States and the United Kingdom, men's hourly wages are between 27 and 33 log points higher than women's hourly wages. Next, in northern and central Europe, the gender wage gap in hourly wages is between 11 and 25 log points, from a minimum of 11 log points in Belgium to a maximum of 25 log points in the Netherlands. Finally, in southern European countries, the gender wage gap is on average below 9 log points, from 5 in Italy to 11 in Spain. Such gaps in hourly wages display a roughly negative correlation with gaps in employment to population ratios. Employment gaps range from less than 13 percentage points, in the United States, the United Kingdom, and Scandinavia,[9] to 17–27 points in northern and central Europe, to 34–49 percentage points in southern Europe. The coefficient of correlation between the two series is $-0.474$ and is significant at the 10% level.

Such negative correlation between wage and employment gaps may reveal significant sample selection effects in observed wage distributions. If the probability of an individual being at work is positively affected by the level of her potential wage offers, and this mechanism is stronger for women than for men, then high gender employment gaps become consistent with relatively low gender wage gaps simply because low-wage women are relatively less likely than men to feature in observed wage distributions.

A simple and intuitive way to illustrate the role of sample selection consists of making alternative conjectures about the potential wages of the nonemployed, as a fraction of observed wages for the employed, as suggested by Smith and Welch (1986, 123). For this purpose we divide the population into three education groups: low, middle, and high, as defined in Section III. True wages for each gender $g$ (= male, female) can be expressed as $W_g = \sum_j \delta_{jg} W_{jg}$, where $\delta_{jg}$ is the population share of edu-

[9] Similar to other Scandinavian countries, the employment gap in Sweden over the same sample period is 5.2 percentage points.

cation group $j$ for gender $g$ and $W_{jg}$ is the associated true wage. In turn, $W_{jg}$ is a weighted average of actual wages for the employed and potential wages for the nonemployed. Assuming that the nonemployed would earn a wage that is equal to a proportion $\gamma$ of the wage of the employed, $W_{jg}$ can then be expressed as

$$W_{jg} = \tilde{W}_{jg}[\gamma + n_{jg}(1 - \gamma)], \tag{1}$$

where $n_{jg}$ is the employment rate of education group $j$ and gender $g$ and $\tilde{W}_{jg}$ is their observed average wage. The reason for first computing (1) by education and then aggregating over education groups is that gender employment gaps vary widely by education. Specifically, they everywhere decline with educational levels, but if anything they do so more strongly in southern Europe than elsewhere (see Olivetti and Petrongolo 2006, table 1A).

The parameter $\gamma$ represents the type and extent of sample selection into employment. In particular, values of $\gamma < 1$ (respectively $> 1$) indicate positive (respectively negative) sample selection. For a given $\gamma$, the role of selection is magnified by a lower employment rate, $n_{jg}$. Denoting by $w$ the log of potential wages, the gender wage gap for education group $j$ is $w_{jmale} - w_{j\text{female}}$. This decreases with $\gamma$ if women have lower employment rates than men, and it increases with the gender employment gap if there is positive sample selection ($\gamma < 1$).

We can now assess the difference between observed and potential wage gaps across alternative values of $\gamma$ after aggregating (1) across education groups. This is shown in table 1 for $\gamma$ = 0.7, 0.5, and 0.3. Column 1 reports for reference the mean wage gap on the 1999 employed sample, as also pictured in figure 1, together with its correlation with the employment gap and its coefficient of variation.[10] Columns 2–4 report the mean wage gap, having corrected for sample selection using (1). Gender wage gaps increase everywhere with lower values of $\gamma$, and, as expected, they do more so in countries with high gender employment gaps. In other words, the higher the gender employment gap, the stronger the impact of a certain degree of positive sample selection. Selection correction gets rid of the negative correlation between gender wage and employment gaps and reduces the coefficient of variation in wage gaps. It is interesting to

[10] The coefficient of correlation is better suited here to assess cross-country variation than the simple standard deviation as the level of the wage gap is also systematically affected by wage imputation. In col. 1, following Krueger and Summers (1988), we adjust the standard deviation of estimated gender gaps across countries to account for the upward bias induced by the least-squares sampling error, i.e., SD = $(\mathrm{Var}(\hat{b}_c) - \sum_{c=1}^{14}\hat{\sigma}_c^2/14)^{1/2}$, where $\hat{b}_c$ is the estimated wage gap in country $c$, $\hat{\sigma}_c$ is the corresponding standard error, and 14 is the number of countries. To obtain the coefficient of variation, we divide SD by the cross-country mean of the estimated $\hat{b}_c$'s. The same adjustment applies to all coefficients of variation reported in tables 2–4.

**Table 1**
**Mean Wage Gaps under Alternative Values of $\gamma$**

| | Base Sample | $\gamma = .7$ | $\gamma = .5$ | $\gamma = .3$ |
|---|---|---|---|---|
| Country: | | | | |
| United States | .325 | .409 | .434 | .460 |
| United Kingdom | .269 | .309 | .336 | .365 |
| Finland | .181 | .239 | .260 | .283 |
| Denmark | .128 | .167 | .178 | .189 |
| Germany | .226 | .295 | .333 | .373 |
| Netherlands | .248 | .334 | .385 | .440 |
| Belgium | .113 | .202 | .246 | .292 |
| Austria | .233 | .306 | .359 | .416 |
| Ireland | .178 | .311 | .367 | .430 |
| France | .124 | .207 | .260 | .318 |
| Italy | .052 | .223 | .311 | .414 |
| Spain | .109 | .296 | .387 | .494 |
| Portugal | .097 | .218 | .264 | .313 |
| Greece | .089 | .353 | .471 | .612 |
| Correlation | −.474 | .302 | .616 | .806 |
| Coefficient of variation | .462 | .247 | .245 | .273 |

Sources.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

Note.—The symbol $\gamma$ represents the ratio between the potential wages of the nonemployed and the observed wages of the employed. Figures reported in rows 1–14 are gender differences in mean log wages. Wages for each gender are obtained as weighted averages across three education groups. Wages for each education group depend on $\gamma$, as illustrated in eq. (1). Figures in the last two rows provide the cross-country correlation between gender and employment gaps and the coefficient of variation for the gender wage gap, respectively. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education, 1999.

note that such correlation becomes positive because selection correction raises the resulting wage gap disproportionately more in countries with very high employment gaps, most notably southern Europe.

Of course, values of $\gamma$ used here for the relative wages of the nonemployed are hypothetical, and thus they only illustrate the mapping between the extent of sample selection and wage gaps. The rest of this article seeks to retrieve evidence on the wages of the nonemployed. As will become clear in the next section, the identifying assumptions needed to do this are much weaker when one estimates median, rather than mean, wage gaps. The focus in the rest of this article will thus be on median gender pay gaps.

## IV. Methodology

Let $w$ denote the natural logarithm of hourly wages and $F(w|g)$ the cumulative log wage distribution for each gender, where $g = 1$ denotes males and $g = 0$ denotes females. In what follows, our variable of interest is the difference between (log) male and female median wages:

$$D = m(w|g = 1) - m(w|g = 0), \tag{2}$$

where $m()$ is the median function. The (log) wage distribution for each gender is defined by

$$F(w|g) = F(w|g, I = 1)\Pr(I = 1|g) + F(w|g, I = 0)[1 - \Pr(I = 1|g)], \quad (3)$$

where $I = 1$ for the employed and $I = 0$ for the nonemployed.

Estimated moments of the observed wage distribution are based on the $F(w|g, I = 1)$ term alone. If there are systematic differences between $F(w|g, I = 1)$ and $F(w|g, I = 0)$, cross-country variation in $\Pr(I = 1|g)$ may translate into misleading inferences concerning the international variation in the distribution of potential wage offers. This problem typically affects estimates of female wage offer distributions; this is even more the case when one is interested in cross-country comparisons of gender wage gaps, given the cross-country variation in $\Pr(I = 1|g = \text{male}) - \Pr(I = 1|g = \text{female})$, measuring the gender employment gap. But $F(w|g)$, the term of interest, is not identified, because data provide information on $F(w|g, I = 1)$ and $\Pr(I = 1|g)$ but clearly not on $F(w|g, I = 0)$, as wages are only observed for those who are in work.

In particular, using (3), the median log wage for each gender, $m$, is defined by

$$F(m|g, I = 1)\Pr(I = 1|g) + F(m|g, I = 0)[1 - \Pr(I = 1|g)] = \frac{1}{2}. \quad (4)$$

Our goal is to retrieve gender gaps in median (potential) wages, as illustrated in equation (2), with gender medians defined in equation (4). To do this we need to retrieve information on $F(m|g, I = 0)$, representing the probability that nonemployed individuals have potential wages below the median.

It can be shown that knowledge of $F(m|g, I = 0)$ allows us to identify the median wage gap in potential wages using median wage regressions, as a simpler alternative to numerically solving (4). Let us consider the linear wage equation

$$w_i = \beta_0 + \beta_1 g_i + \varepsilon_i, \quad (5)$$

where $w_i$ denotes (log) potential wages, $\beta_0$ is a constant term, $\beta_1$ is the parameter of interest, and $\varepsilon_i$ is an error term such that $m(\varepsilon_i|g_i) = 0$. Denote by $\hat{\beta}$ the hypothetical least absolute deviations (LAD) regression estimator based on potential wages, that is, $\hat{\beta} \equiv \arg\min_\beta \sum_{i=1}^{N} |w_i - \beta_0 - \beta_1 g_i|$, where $\beta \equiv [\beta_0 \beta_1]'$.

However, wages $w_i$ are only observed for the employed; they are missing for the nonemployed. Consider an example in which missing wages fall completely below the median regression line, that is, $w_i < \hat{w}_i = \hat{\beta}_0 + \hat{\beta}_1 g_i$ for the nonemployed ($I_i = 0$), or equivalently, $F(m|g, I = 0) = 1$. One can then define a transformed dependent variable $y_i$ that is equal to $w_i$ for $I_i = 1$ and to some arbitrarily low imputed value $\underline{w}$ (such that $\underline{w}_i < \hat{w}_i$) for $I_i = 0$, and the following result holds (see Bloomfield and Steiger [1983], sec. 2.3, for detail and formal proof):

$$\hat{\beta}_{\text{imputed}} = \arg\min_{\beta} \sum_{i=1}^{N} |y_i - \beta_0 - \beta_1 g_i|$$

$$= \hat{\beta} \equiv \arg\min_{\beta} \sum_{i=1}^{N} |w_i - \beta_0 - \beta_1 g_i|. \tag{6}$$

Condition (6) states that the LAD estimator is not affected by imputation. In other words, obtaining $\hat{\beta}$ using the transformed dependent variable $y_i$ gives the same estimate that one would obtain if potential wages were available for the whole population. Now consider an alternative example in which missing wages fall completely above the median regression line, that is, $w_i > \hat{w}_i$ for $I_i = 0$, or equivalently, $F(m|g, I = 0) = 0$. The result in (6) still holds, having set $y_i$ equal to some arbitrarily high imputed value $\bar{w}$ (such that $\bar{w}_i > \hat{w}_i$ ) for the nonemployed. More in general, the LAD estimator is not affected by imputation when the missing wage observations are imputed "so as to maintain the same sign of the residual" (Bloomfield and Steiger 1983, 52). That is, (6) is valid whenever missing wage observations are imputed on the "correct" side of the median. As a further example, suppose that the potential wages of the nonemployed could be classified into two groups, $L$ and $U$, such that $w_i < \hat{w}_i$ for $i \in L$ and $w_i > \hat{w}_i$ for $i \in U$. One can define $y_i$ as a transformed variable such that $y_i = w_i$ for $I_i = 1$, $y_i = \underline{w}$ for $I_i = 0$ and $i \in L$, and $y_i = \bar{w}$ for $I_i = 0$ and $i \in U$, and LAD inference is still valid.

Using this result, one can estimate median wage gaps, based on wage imputation for the nonemployed that simply requires assumptions on the position of the imputed wage observations with respect to the median of the wage distribution, as done in Johnson et al. (2000) and Neal (2004). The attractive feature of median regressions is that results are only affected by the position of imputed wage observations with respect to the median and not by specific values of imputed wages as it would be in the matching approach. In this article, we will estimate median wage gaps under alternative imputation rules, that is, under alternative conjectures over $F(m|g, I = 0)$. These imputation rules are described in detail below.

### A. Imputation on Wages from Other Waves

We first exploit the panel nature of our data sets and, for all those not in work in some base year $t$, we recover (the real value of) hourly wage observations from the nearest wave in the sample, $t'$, and we use them as imputed wages ($y_i$) for estimating (6). The underlying identifying assumption is that, for a given individual $i$, the latent wage position with respect to her predicted (gender-specific) median when she is nonemployed can be proxied by her wage in the nearest wave in which she is employed. As the position with respect to the median is determined using alternative information on wages, as opposed to measured characteristics, we are allowing for selection on unobservables.

Formally, we will assume

$$F(m|g_i, I_{it} = 0) = F(m|g_i, I_{it'} = 1), \tag{7}$$

where $t$ is our base year and $t'$ is the wave nearest to $t$ in which we have a nonmissing wage observation. In practice, we impute $y_{it} = w_{it'}$ for $I_{it} = 0$. This procedure of imputation makes sense if an individual's position in the wage distribution stays on the same side of the median when the individual is switching employment status. As we estimate median wage gaps, we do not need an assumption of stable rank throughout the whole wage distribution but only with respect to the median. Should the position of individuals in the wage distribution change with employment status, movements that happen within either side of the median do not invalidate this method.

While imputation based on this procedure arguably uses the minimum set of potentially arbitrary assumptions, it has the disadvantage of not providing any wage information on individuals who never worked during the sample period. It is therefore important to understand in which direction this problem may distort, if at all, the resulting median wage gaps. If women are on average less attached to the labor market than men, and if attachment increases with potential wages, then the difference between the median gender wage gap on the imputed and the actual wage distribution tends to be higher the higher the proportion of imputed wage observations in total nonemployment in the base year. Consider, for example, a country with a very persistent female employment status: women who do not work in the base year and are therefore less attached are also unlikely to work at all in the whole sample period. In this case, low wage observations for less-attached women are unlikely to be recovered and the estimated wage gap is relatively low. Proportions of imputed wage observations over the total nonemployed population in 1999 (our base year) are reported in table A2 of the online appendix: the differential between male and female proportions tends to be higher in Germany, Austria, France, and southern Europe than elsewhere. Under reasonable assumptions we should therefore expect the difference between the median wage gap on the imputed and the actual wage distribution to be biased downward relatively more in this set of countries. This, in turn, means that we are being relatively more conservative in assessing the effect of nonrandom employment selection in these countries than elsewhere.

Even so, it would, of course, be preferable to recover wage observations also for those never observed in work during the whole sample period. To do this, we rely on the observed characteristics of the nonemployed.

### B. Imputation on Observables

We use observable characteristics for wage imputation with two methods. With the first method, we make assumptions on the position of missing wages with respect to their gender-specific median, based on a small number of characteristics, summarized into the vector $X_i$. We can illustrate this with a very simple example. Suppose that $X_i$ only includes years of completed education. This implies that we are using information on education for someone who is nonemployed to place them above or below their gender-specific median. We can define a threshold for $X_i$, $\underline{x}$ (say, 11 years of schooling), below which nonemployed individuals would earn below-median wages, and another threshold $\bar{x}$ (say, 16 years), above which individuals would earn above-median wages.

More formally we assume that

$$F(m|g_i, I_i = 0, X_i \leq \underline{x}) = 1; \quad F(m|g_i, I_i = 0, X_i \geq \bar{x}) = 0, \tag{8}$$

where $\underline{x}$ and $\bar{x}$ are low and high values of $X_i$, respectively.[11] In this case, the imputed dependent variable $y_i$ is set equal to $\underline{w}$ for $i$ such that $I_i = 0$ and $X_i \leq \underline{x}$ and is set equal to $\bar{w}$ for $i$ such that $I_i = 0$ and $X_i \geq \bar{x}$. This method for placing individuals with respect to the median follows an educated guess, based on their observable characteristics. However, we can use wage information from other waves in the panel to assess the goodness of such guess, as will be illustrated in Section V.B.

With the second method, we use probability models for imputing missing wage observations. In this case our imputation rule assumes that

$$F(m|g_i, I_i = 0, X_i) = \hat{P}_i, \tag{9}$$

where $\hat{P}_i$ is the predicted probability to belong below the median, based on probit estimates.

We implement this imputation method in two steps. In the first step, we estimate the probability of an individual's wage belonging below the median of the wage distribution, based on a set of observable characteristics. On the employed sample, we define $M_i = 1$ for individuals earning less than their gender-specific median and $M_i = 0$ for the others. We estimate a probit model for $M_i$ for each gender, with explanatory variables $X_i$. Using the probit estimates, we obtain predicted probabilities of having a latent wage below the median, $\hat{P}_i = \Phi(\hat{\gamma} X_i) = \Pr(M_i = 1|X_i)$, for the nonemployed subset, where $\Phi$ is the cumulative distribution function of the standardized normal distribution and $\hat{\gamma}$ is the estimated parameter

[11] All variables in (8) refer to the (same) base year, so time subscripts are omitted.

vector from the probit regression. Predicted probabilities $\hat{P}_i$ are then used in the second step as sampling weights for the nonemployed. That is, we construct an imputed sample in which the employed feature with their observed wage and the nonemployed feature with a wage below median with a weight $\hat{P}_i$ and a wage above median with a weight $1 - \hat{P}_i$. The statistics of interest is the gender wage gap estimated on the imputed sample. The associated variance is obtained by bootstrap to correct for the fact that the weights used are based on probit estimates.

Note that in the first step we need a reference median wage in order to define $M_i$. The readily available candidate would be the median observed wage, but precisely due to selection this may be quite different from the latent median wage, thus potentially delivering biased estimates. In order to attenuate this problem, we also perform repeated imputation on an expanded sample, augmented with wage observations from adjacent waves. This allows us to get a better estimate of the potential median in the first step of our procedure, thereby generating more appropriate estimates of the median wage gap on the final, imputed sample.

### C. Discussion on Imputation Methodology

To ease the interpretation of the results presented in the next section, we discuss here the main differences between alternative imputation methods. The three methods described differ in terms of the underlying identifying assumptions and the resulting imputed samples. The first method, where missing wages are imputed using wage information from other waves, implicitly assumes that an individual's position with respect to the median can be proxied by his or her wage in the nearest wave in the panel. With this procedure, one can recover, at best, individuals who worked at least once during the 8-year sample period. We thus emphasize that this is a fairly conservative imputation procedure, one in which we impute wages for individuals who are relatively weakly attached to the labor market but not for those who are completely unattached and thus have never been observed in work. This procedure has the advantage of restricting imputation to a relatively "realistic" set of potential workers, and thus it is the one we mostly rely upon to make quantitative statements.

In the second and third imputation methods, we assume instead that an individual's position with respect to the median can be proxied by some of his or her observable characteristics. In the second method, we use characteristics to take educated guesses regarding the position of missing wages. Clearly this procedure is more accurate for values of the observables in the tails than in the middle of the distribution. For example, guessing the position with respect to the median for individuals with either college or no education at all is safer than doing it for secondary school graduates, who are thus best left without an imputed wage. In doing this, our imputed sample is typically larger than the one obtained with the first method, although it is still substantially smaller than the existing population. Finally, with the third method, we estimate the probability of belonging above the median for the whole range of our vector of characteristics, thus recovering predicted probabilities and imputed wages for the whole existing population.

Different imputed samples will have an impact on our estimated median wage gaps. In so far as women tend to be more positively selected into employment than men, the larger the imputed sample with respect to the actual sample of employed workers, the larger the estimated correction for selection.

Having said this, it is important to stress that, with all three imputation methods used, we never impose positive selection ex ante (except in a benchmark example), and thus there is nothing that would tell a priori which way correction for selection is going to affect the results. This is ultimately determined by the wages that the nonemployed earned when they were previously (or later) employed and by their observable characteristics, depending on methods.

Before moving on to the discussion of our estimates, it is worthwhile to motivate our choice of selection correction methodology and to frame it in the context of the existing literature on sample selection. A number of approaches can be used to correct for nonrandom sample selection in wage equations and/or recover the distribution in potential wages. The seminal approach suggested by Heckman (1974, 1979) consists in allowing for selection on unobservables, that is, on variables that do not feature in the wage equation but that are observed in the data.[12] Heckman's two-stage parametric specifications have been used extensively in the literature in order to correct for selectivity bias in female wage equations. More recently, these have been criticized for lack of robustness and distributional assumptions (see Manski 1989). Approaches that circumvent most of the criticism include semiparametric selection correction models that have appeared in the literature since the early 1980s (see Vella [1998] for an extensive survey of both parametric and nonparametric sample selection models). Two-stage nonparametric methods allow us in principle to approximate the bias term by a series expansion of propensity scores from

[12] In this framework, wages of employed and nonemployed would be recovered as

$$E(w|Z_w, I = 1) = Z_w\delta_w + E(\varepsilon_w|\varepsilon_I > -Z_I\delta_I)$$

$$E(w|Z_w, I = 0) = Z_w\delta_w + E(\varepsilon_w|\varepsilon_I < -Z_I\delta_I),$$

respectively, where $Z_w$ and $Z_I$ are the set of covariates included in the wage and selection equations, respectively, with associated parameters $\delta_w$ and $\delta_I$, and $\varepsilon_w$ and $\varepsilon_I$ are the respective error terms.

the selection equation, with the qualification that the term of order zero in the polynomial is not separately identified from the constant term in the wage equation unless some additional information is available (see Buchinsky 1998). Usually the constant term in the wage regression is identified from a subset of workers for which the probability of work is close to one, but in our case this route is not feasible since for no type of women is the probability of working close to one in all countries.

Selection on observed characteristics is instead exploited in the matching approach, which consists in imputing wages for the nonemployed by assigning them the observed wages of the employed with matching characteristics (see Blau and Beller 1992; and Juhn 1992, 2003). The approach of this article is also based on some form of wage imputation for the nonemployed, but it simply requires assumptions on the position of the imputed wage observations with respect to the median of the wage distribution. Importantly, it does not require assumptions on the actual level of missing wages, as is typically required in the matching approach, nor it requires the arbitrary exclusion restrictions that are often invoked in the two-stage Heckman sample selection correction models.

## D. Bounds

As discussed above, each imputation method is based on identifying assumptions that are largely untested. In order to illustrate that results delivered by our imputation methods are reasonable, we also provide "worst case" bounds to the gap in potential median wages that do not require any identifying assumption, as shown by Manski (1994) and Blundell et al. (2007). We will then check that our estimated wage gaps on imputed wage distributions fall into these bounds.

Manski notes that substituting the inequality $0 \leq F(w|g, I = 0) = 1$ into (3) gives bounds for the true cumulative distribution

$$F(w|g, I = 1)\,\Pr(I = 1|g) \leq F(w|g)$$
$$\leq F(w|g, I = 1)\,\Pr(I = 1|g) + [1 - \Pr(I = 1|g)]. \tag{10}$$

If one is interested in the median of $F()$, denoted by $m$, (10) implies that

$$F(m|g, I = 1)\,\Pr(I = 1|g) \leq \frac{1}{2}$$
$$\leq F(m|g, I = 1)\,\Pr(I = 1|g) + [1 - \Pr(I = 1|g)]. \tag{11}$$

These bounds on $F()$ deliver the following worst case bounds on the gender-specific median

$$m^{l}(w|g) \leq m(w|g) \leq m^{u}(w|g) \tag{12}$$

such that

$$\frac{1}{2} = F(m^l|g, I = 1)\,\Pr(I = 1|g) + [1 - \Pr(I = 1|g)], \tag{13}$$

$$\frac{1}{2} = F(m^u|g, I = 1)\,\Pr(I = 1|g). \tag{14}$$

Bounds on the gender-specific median can be obtained by solving (13) and (14) using data on the observed wage distribution and employment rates. Note that conditions (13) and (14) imply that one can only identify bounds for the median if $\Pr(I = 1|g) = 1/2$. Hence, we will not be able to obtain such bounds for the female median wage (and, therefore, for the gender wage gap) in countries where less than 50% of the women have a wage observation.

Having said this, the bounds for the median gender wage gap $D$ defined in (2) are obtained as follows:

$$\begin{aligned} & m^l(w|g = \text{male}) - m^u(w|g = \text{female}) \\ & \leq D \leq m^u(w|g = \text{male}) - m^l(w|g = \text{female}). \end{aligned} \tag{15}$$

## V. Results

### A. Imputation on Wages from Adjacent Waves

In our first set of estimates, an individual's position with respect to the median of the wage distribution in the base year is proxied by the position of his or her wage obtained from the nearest available wave. The kind of imputation made here requires that individuals stay on the same side of their gender median across different waves in the panel (see eq. [7]). Results obtained with this method are reported in table 2.

Column 1 reports the actual wage gap for reference: this is the median wage gap for individuals with an hourly wage in 1999, which is our base year. The wage gaps of column 1 replicate very closely those plotted in figure 1, with the only difference being that figure 1 plotted mean as opposed to median wage gaps.[13] As in figure 1, the United States and the United Kingdom stand out as the countries with the highest wage gaps, followed by central Europe, and finally by Scandinavia and Southern Europe.

In column 2, missing wage observations in 1999 are replaced with the real value of the nearest wage observation in a 2-year window, while in column 3, they are replaced with the real value of the nearest wage obser-

[13] The absence of any important difference between mean and median wage gaps on the observed wage distribution is good news for our approach, based on the recovery of selection-corrected median wage gaps.

**Table 2**
**Median Wage Gaps under Alternative Sample Inclusion Rules: Wage Imputation Based on Wage Observations from Adjacent Waves**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Country: | | | | | |
| United States | .339 | .352 | .361 | .354 | .363 |
| United Kingdom | .256 | .277 | .284 | .291 | .301 |
| Finland | .160 | .194 | .197 | .203 | .206 |
| Denmark | .086 | .100 | .100 | .093 | .093 |
| Germany | .191 | .223 | .214 | .234 | .234 |
| Netherlands | .178 | .193 | .199 | .195 | .199 |
| Belgium | .078 | .099 | .112 | .098 | .112 |
| Austria | .192 | .224 | .234 | .220 | .229 |
| Ireland | .232 | .273 | .284 | .292 | .300 |
| France | .095 | .133 | .152 | .140 | .164 |
| Italy | .059 | .062 | .075 | .070 | .079 |
| Spain | .097 | .153 | .168 | .143 | .157 |
| Portugal | .150 | .168 | .186 | .169 | .185 |
| Greece | .111 | .148 | .184 | .148 | .185 |
| Correlation | −.329 | −.263 | −.181 | −.269 | −.199 |
| Coefficient of variation | .484 | .416 | .382 | .427 | .392 |

Sources.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

Note.—All wage gaps are significant at the 1% level. Figures in the last two rows display the cross-country correlation between the reported gender wage gap and the gender employment gap and the coefficient of variation of the gender wage gap. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education, 1999. Sample inclusion rules by columns: (1) Employed at time of survey in 1999. (2) Wage imputed from other waves when nonemployed (−2, +2 window). (3) Wage imputed from other waves when nonemployed (−5, +2 window). (4) Wage imputed from other waves when nonemployed (−5, +2 window), adjusted for real wage growth by gender and country. (5) Wage imputed from other waves when nonemployed (−5, +2 window), adjusted for real wage growth by gender and country.

vation in the whole sample period, meaning a maximum window of [−5, +2] years. Moving across from column 1 to column 3, gender wage gaps tend to increase as more wage observations are included into the imputed sample. This is indicative of positive sample selection, or, in other words, estimated wage gaps on the observed wage distribution are downward biased due to nonrandom sample selection into employment because low-wage women are less likely to feature in the observed wage distribution.

But there is important cross-country variation in the role of selection. In particular, one can see that the median wage gap remains substantially unaffected or marginally affected in the United States, the United Kingdom, Scandinavia, Germany, and the Netherlands; it increases by around 25% in Austria, Ireland, Italy, and Portugal; by 40% in Belgium; and by more than 60% in France, Spain, and Greece. As expected, gender wage gaps tend to respond more strongly to selection correction in countries with high employment gaps. This can be clearly grasped by looking at the cross-country correlation between employment and wage gaps. In column 1, such correlation is −0.329, and it falls by 45% in column 3. Employment selection thus explains nearly a half of the observed correlation between wage and employment gaps. Another indicative cross-country statistic is the coefficient of variation in the gender wage gap: this falls by 31% between column 1 and column 3; thus selection explains almost one-third of the observed cross-country variation of wage gaps.

For each sample inclusion rule in columns 1–3, one can compute the adjusted employment rate for each gender, that is, the proportion of the adult population that is either working or has an imputed wage. These proportions are reported later in columns 1–3 of table 5. When moving from column 1 to column 3, the fraction of women included increases substantially in most countries, including some countries where the estimated wage gap is not greatly affected by the sample inclusion rules. Moreover, the fraction of men included in the sample also increases, albeit less so than for women. It is thus not simply the lower female employment rate in several countries that drives our findings but also the fact that in some countries selection into work seems to be less correlated to wage characteristics than in others.

The estimates of columns 2 and 3 of table 2 do not control for aggregate wage growth over time. If aggregate wage growth were homogeneous across genders and countries, the estimated wage gaps based on wage observations for other waves in the panel would not be affected. But, if there is a gender differential in wage growth, and if such differential varies across countries, then simply using earlier (later) wage observations would deliver a higher (lower) median wage gap in countries where wage growth for women is lower than for men.[14] We thus estimate real wage growth by regressing log real hourly wages for each country and gender on a linear trend.[15] The resulting coefficients are reported in table A3 in the online appendix. These are then used to adjust real wage observations outside the base year and to re-estimate median wage gaps. The resulting median wage gaps on the imputed wage distribution are reported in columns 4 and 5 of table 2. Despite some differences in real wage growth rates across genders and countries, adjusting estimated median wage gaps does not produce any appreciable change in the results reported in columns 2 and 3, which do not control for real wage growth.

Note finally that, in table 2, we are only recovering wages for a quarter, on average, of nonemployed women in the four southern European countries, as opposed to more than a half in the rest of countries (see table A2 in the online appendix). For men, cross-country differences are less marked, as respective proportions are 57% and 63%. Such differences

[14] Note, however, that, even if real wage growth were homogeneous across genders, imputation based on wage observations from adjacent waves would not be affected only if the proportion of men and women in the sample remained unchanged after imputation.

[15] Of course, for our estimated rates of wage growth to be unbiased, this procedure requires that participation into employment be unaffected by wage growth, which may not be the case.

happen because (non)employment status tends to be relatively more persistent in southern Europe than elsewhere, and this is much more so for women than for men. As noted in Section IV, given that we recover relatively fewer less-attached women in southern Europe, we are being relatively more conservative in assessing the effect of nonrandom employment selection in southern Europe as compared to elsewhere. For this reason it is important to try to recover wage observations also for those never observed in work in any wave of the sample period, as explained in the next subsection.

### B. Imputation on Observables

In table 3, we exploit some observable characteristics of the nonemployed for assigning them on one or the other side of their gender median (eq. [8] gives the formal identifying assumption). Column 1 reports for reference the median wage gap on the base sample, which is the same as the one reported in column 1 of table 2. In column 2, we assume that all those not in work in 1999 would have wage offers below the median for their gender.[16] This is an extreme assumption, and it is the only case in which we impose ex ante positive sample selection. This assumption is clearly violated for countries like Italy, Spain, and Greece, in which more than half of the female sample is not in work in 1999, and thus estimates are not reported for these three countries. However, also for other countries, there are reasons to believe that not all nonemployed individuals would have wage offers below their gender mean. Having said this, estimated median wage gaps increase substantially for most countries, except for Denmark and Finland. The correlation with employment gaps turns positive and quite strong because the wage gap in high employment-gap countries increases disproportionately relative to other countries.

Of course, one cannot know exactly what wages these individuals would have received had they worked in 1999. But we can form an idea of the goodness of this assumption by looking again at wage observations (if any) for these individuals in all other waves in the panel. This allows us to see whether an imputed observation had a wage that was indeed below their predicted gender median at the time he or she was observed working. Specifically, we take all imputed observations in 1999. Among these, we select those who ever worked at some time in the sample period. Out of this subset, we compute the proportion of observations who had wages below the predicted gender median. Such proportions are computed for men and women and are reported in columns headed "M" and "F,"

[16] In practice, whenever we assign someone a wage below the median, we pick $w_i = -5$, this value being lower than the minimum observed (log) wage for all countries and thus lower than the median. Similarly, whenever we assign someone a wage above the median, we pick $\bar{w}_i = 20$.

**Table 3**
**Median Wage Gaps under Alternative Imputation Rules: Wage Imputation Based on Observables—Educated Guesses**

| | 1 | 2 | | | 3 | | | 4 | | | 5 | | | 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Goodness Imputation (%) | | | Goodness Imputation (%) | | | Goodness Imputation (%) | | | Goodness Imputation (%) | | |
| | Wage Gap | Wage Gap | M | F | Wage Gap | M | F | Wage Gap | M | F | Wage Gap | M | F | Wage Gap |
| Country: | | | | | | | | | | | | | | |
| United States | .339 | .432 | .90 | .74 | .335 | 1.00 | .89 | .348 | .67 | .78 | .348 | .82 | .88 | .367 |
| United Kingdom | .256 | .375 | .80 | .63 | .238 | .83 | .81 | .246 | .71 | .56 | .253 | .87 | .89 | .287 |
| Finland | .160 | .191 | .86 | .73 | .179 | .87 | .86 | .161 | .40 | .43 | .160 | .80 | .83 | .188 |
| Denmark | .086 | .122 | .77 | .77 | .100 | .85 | .81 | .076 | .63 | .55 | .092 | 1.00 | .60 | .100 |
| Germany | .191 | .368 | .87 | .53 | .214 | .88 | .74 | .192 | .53 | .54 | .196 | .94 | .86 | .215 |
| Netherlands | .178 | .353 | .67 | .50 | .246 | .67 | .62 | .208 | .67 | .78 | .189 | .79 | .78 | .221 |
| Belgium | .078 | .228 | .68 | .70 | .113 | .70 | .82 | .082 | .80 | .52 | .084 | .69 | 1.00 | .116 |
| Austria | .192 | .360 | .89 | .60 | .210 | .91 | .72 | .192 | .00 | .43 | .208 | .88 | .80 | .234 |
| Ireland | .232 | .591 | .83 | .33 | .213 | .87 | .94 | .237 | .75 | .79 | .241 | .73 | .92 | .291 |
| France | .095 | .357 | .81 | .56 | .134 | .84 | .82 | .114 | .75 | .79 | .102 | .83 | .86 | .161 |
| Italy | .059 | . . . | .75 | . . . | .092 | .78 | .69 | .200 | .93 | .74 | .088 | .83 | .86 | .186 |
| Spain | .097 | . . . | .71 | . . . | .182 | .74 | .65 | .199 | .70 | .78 | .109 | .72 | .96 | .235 |
| Portugal | .150 | .357 | .65 | .52 | .188 | .61 | .69 | .264 | .67 | .71 | .162 | .69 | .69 | .258 |
| Greece | .111 | . . . | .78 | . . . | .175 | .81 | .73 | .349 | .67 | .70 | .160 | .73 | .60 | .362 |
| Correlation | −.329 | .625 | | | −.120 | | | .435 | | | −.200 | | | .395 |
| Coefficient of variation | .484 | .364 | | | .329 | | | .393 | | | .430 | | | .336 |

Sources.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

Note.—All wage gaps are significant at the 1% level. In specification 2 no results are reported for Italy, Spain, and Greece as more than 50% of women in the sample are nonemployed. The goodness of imputation is measured using wage observations (if any) from other waves in the panel, and specifically computing what proportion of these is on the same side of the median that we assumed in imputation. Figures in the last two rows display the cross-country correlation between the reported gender wage gap and the gender employment gap, and the coefficient of variation of the gender wage gap. Sample: aged 25–54, excluding the self-employed, the military and those in full-time education, 1999. Sample inclusion rules by columns: (1) Employed at time of survey in 1999. (2) Impute wage < median when nonemployed. (3) Impute wage < median when unemployed. (4) Impute wage < median when nonemployed and education < upper secondary and experience < 10 years; impute wage > median when nonemployed and education ≥ higher education and experience ≥ 10 years. (5) Impute wage < median when nonemployed and spouse income in bottom quartile. (6) Wage imputed from other waves when nonemployed (−5, +2 window) and sample inclusion rule 4.

respectively. They are fairly high for men, but they are sensibly lower for women, which makes the estimates based on this extreme imputation case a benchmark rather than a plausible measure for the gender wage gap.

In column 3, we impute a wage below the median to all those who are unemployed (as opposed to nonparticipants) in 1999. The unemployed, by definition, are receiving wage offers (if any) below their reservation wage, while the employed have received at least one wage offer above their reservation wage. At constant reservation wages, the unemployed have lower potential wages than the observed wages of the employed and are thus assigned an imputed wage value below the median. This imputation leaves the median wage gap roughly unchanged with respect to the base sample in the United States, the United Kingdom, Scandinavia, Germany, Austria, and Ireland and raises it substantially elsewhere, especially in southern Europe. Also, the proportion of "correctly" imputed observations, computed as for the previous imputation case, is now much higher. Those who do not work because they are unemployed are thus relatively more likely to be overrepresented in the lower half of the wage distribution. Selection now explains 64% of the correlation between wage and employment gaps and 32% of the cross-country variation.

In column 4, we follow standard human capital theory and assume that all those with less than upper secondary education and fewer than 10 years of labor market experience have wage observations below the median for their gender. Those with at least higher education and at least 10 years of labor market experience are instead placed above the median. In the four southern European countries, the gender wage gap increases enormously with respect to the actual wage gap of column 1 and, as a consequence, the correlation with employment gaps turns positive. It is interesting that the proportion of correctly imputed observations is high in Ireland, France, and southern Europe but that is not so much so in other countries, where imputation based on unemployment works better than imputation based on human capital components.

The imputation method of column 5 is implicitly based on the assumption of assortative mating along wage attributes and consists of assigning wages below the median to those whose partners have total income in the bottom quartile of the gender-specific distribution. The assumption is that individuals married to low-productivity spouses also have low productivity, and thus the spouse's wage is taken as a proxy for an individual's potential wage offer. The results are qualitatively similar to those of column 3: the wage gap is mostly affected in southern Europe, but on average it is less affected than in other imputation examples. It would be natural to perform a similar exercise at the top of the distribution by assigning a wage above the median to those whose partner earns income in the top quartile. However, in this case, the proportion of correctly imputed observations was too low to rely on the assumption used for imputation. We have also considered imputation based on low spouse education, obtaining very similar results as with low spouse income. Finally, we considered imputation based on high spouse education, and, similar to imputation based on high spouse income, the goodness of imputation turned out worse.

We also combined imputation methods by using, first, wage observations available from other waves and then imputing the remaining missing ones using education and experience information, as done in column 4. The results, reported in column 6, show again a much higher gender gap in France and southern Europe and not much of a change elsewhere with respect to the base sample of column 1.

Similar to the previous imputation method, we report in columns 4–8 of table 5 the proportion of men and women included in our imputed samples. As expected, we are now able to recover wage information for a higher fraction of the adult population. In column 4, such proportions are generally not equal to 100% because we did not impute wages to those who are employed but have missing information on hourly wages due to nonresponse, as the selection mechanism driving nonresponse is clearly different from that driving nonemployment.

Finally, we report estimates based on a probabilistic, two-step imputation technique, summarized in equation (9). In the first step, we use the 1999 base sample to estimate a probit model for the probability of belonging above the gender-specific median, controlling for education (upper secondary and higher education), experience, and its square.[17] The estimated coefficients for the first-stage probit regression (not reported) conform to standard economic theory: individuals with higher levels of educational attainment and/or of labor market experience are more likely to feature in the top half of the wage distribution. These estimates are used as sampling weights in the second step to construct an imputed sample, on which we estimate the median gender wage gap and the corresponding bootstrapped standard error (with 200 replications).

The results of this exercise are summarized in table 4. Column 1 reports the median wage gap for the base sample, which is the same as the one reported in column 1 of tables 2 and 3. Column 2 reports the estimated median wage gap obtained from the probabilistic model described, having used the observed 1999 median as the reference median for our probit estimates. As fewer than 50% of women are employed in Italy, Spain, and Greece, we cannot credibly estimate a probit model where $M_i = 1$

[17] We also estimated a more general specification that also controls for marital status, the number of children of different ages, and the position of the spouse in their gender-specific distribution of total income. Since the results of the exercise do not vary in any meaningful way across specifications, we only report findings for the human capital specification.

**Table 4**
**Median Wage Gaps under Alternative Imputation Rules: Wage Imputation Based on Observables—Probabilistic Model**

| | Base Sample | Weighted Imputation | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Country: | | | |
| United States | .339 | .359 | .371 |
| United Kingdom | .256 | .264 | .292 |
| Finland | .160 | .179 | .199 |
| Denmark | .086 | .100 | .100 |
| Germany | .191 | .200 | .232 |
| Netherlands | .178 | .229 | .235 |
| Belgium | .078 | .117 | .154 |
| Austria | .192 | .205 | .236 |
| Ireland | .232 | .319 | .341 |
| France | .095 | .182 | .186 |
| Italy | .059 | . . . | .229 |
| Spain | .097 | . . . | .333 |
| Portugal | .150 | .272 | .272 |
| Greece | .111 | . . . | .593 |
| Correlation | −.329 | .291 | .686 |
| Coefficient of variation | .484 | .339 | .427 |

Sources.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

Note.—All wage gaps are significant at the 1% level. In specification 2 no results are reported for Italy, Spain, and Greece as more than 50% of women in the sample are nonemployed. Figures in the last two rows display the cross-country correlation between the reported gender wage gap and the gender employment gap and the coefficient of variation of the gender wage gap. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education, 1999. Sample inclusion rules by columns: (1) Employed at time of survey in 1999. (2) Impute wage < (resp. >) median with probability $\hat{P}_i$(resp. $1 - \hat{P}_i$) if nonemployed. $\hat{P}_i$ is the predicted probability of having a wage below the gender-specific base sample median, as estimated from a probit model including two education dummies, experience, and its square. (3) Impute wage < (resp. >) median with probability $\hat{P}_i$(resp. $1 - \hat{P}_i$) if nonemployed. $\hat{P}_i$ as above, having enlarged the base sample with wage observation from adjacent waves.

for workers earning less than the median for these countries. In column 3, we use as the reference median the one obtained on a wage distribution enlarged with wage imputation from all other waves, and in this case the fraction of missing wages is below 50% for men and women in all countries. If wage imputation is correct, this procedure delivers a reference median that is closer to the latent median than the observed median. Comparing column 1 to columns 2 and 3 shows that the median wage gap on imputed wage distributions increases mildly in most countries down to Austria but rises substantially in Ireland, France, and Portugal and enormously in Italy, Spain, and Greece, which are the countries with the highest employment gaps.

To broadly summarize our findings (a summary is provided in table 5), one could note that, whether one corrects for selection on unobservables (table 2) or observables (tables 3 and 4), our results are qualitatively consistent in identifying a clear role of sample selection in countries with

**Table 5**
**Percentage of Adult Population in Samples for Tables 2–4**

| | No. observations, 1999 | | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | | 8 | | 9 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | F | M | F | M | F | M | F | M | F | M | F | M | F | M | F | M | F | M | F |
| United States | 2,835 | 3,551 | 95.9 | 83.3 | 98.2 | 91.4 | 98.5 | 92.7 | 100.0 | 100.0 | 96.4 | 84.1 | 96.9 | 88.9 | 96.7 | 86.3 | 99.0 | 95.3 | 99.7 | 99.3 |
| United Kingdom | 1,998 | 2,458 | 87.2 | 74.3 | 92.4 | 84.9 | 93.3 | 88.0 | 96.9 | 96.8 | 91.2 | 76.2 | 90.5 | 81.0 | 89.9 | 77.1 | 95.3 | 91.1 | 98.8 | 98.2 |
| Finland | 1,464 | 1,697 | 92.6 | 82.5 | 96.9 | 92.9 | 97.2 | 93.6 | 99.4 | 98.8 | 98.8 | 91.2 | 93.2 | 86.7 | 93.2 | 83.3 | 97.4 | 95.3 | 100.0 | 99.6 |
| Denmark | 996 | 1,054 | 94.0 | 88.1 | 99.5 | 97.4 | 99.5 | 97.7 | 98.3 | 98.1 | 97.5 | 94.1 | 95.0 | 90.5 | 94.6 | 88.9 | 99.7 | 98.0 | 99.9 | 99.6 |
| Germany | 2,822 | 3,064 | 91.2 | 68.8 | 96.7 | 83.1 | 98.1 | 87.0 | 98.9 | 93.4 | 97.6 | 75.2 | 92.2 | 72.0 | 92.6 | 70.7 | 98.4 | 88.4 | 99.5 | 97.0 |
| Netherlands | 2,322 | 2,750 | 94.1 | 70.4 | 96.7 | 81.5 | 97.1 | 84.3 | 99.8 | 99.1 | 96.7 | 82.0 | 95.8 | 77.8 | 95.6 | 72.9 | 98.1 | 89.1 | 99.8 | 99.6 |
| Belgium | 1,170 | 1,382 | 90.3 | 70.2 | 94.1 | 77.9 | 94.9 | 81.5 | 98.7 | 98.2 | 95.8 | 80.2 | 91.6 | 77.1 | 92.8 | 73.5 | 95.7 | 86.6 | 99.0 | 97.5 |
| Austria | 1,296 | 1,378 | 95.8 | 69.3 | 98.5 | 78.7 | 98.7 | 81.6 | 99.8 | 97.8 | 99.2 | 72.2 | 95.8 | 70.9 | 96.6 | 71.4 | 98.7 | 82.7 | 99.8 | 95.8 |
| Ireland | 1,053 | 1,345 | 84.9 | 58.1 | 89.8 | 70.8 | 90.7 | 73.8 | 99.6 | 99.0 | 92.9 | 61.0 | 88.3 | 64.0 | 88.9 | 62.2 | 93.0 | 78.3 | 99.8 | 99.2 |
| France | 2,601 | 2,923 | 73.1 | 55.7 | 92.8 | 75.6 | 94.4 | 80.1 | 85.5 | 90.2 | 79.8 | 64.6 | 75.3 | 63.7 | 74.8 | 58.0 | 95.7 | 84.9 | 98.4 | 97.0 |
| Italy | 3,113 | 3,744 | 79.3 | 45.2 | 89.9 | 54.9 | 91.3 | 57.9 | 94.9 | 96.8 | 91.9 | 56.1 | 83.2 | 64.0 | 81.5 | 49.7 | 93.8 | 74.9 | 98.6 | 96.6 |
| Spain | 2,680 | 3,065 | 83.7 | 45.5 | 91.2 | 59.4 | 92.6 | 62.5 | 99.7 | 99.6 | 93.8 | 56.1 | 87.0 | 60.1 | 86.5 | 47.9 | 94.7 | 73.8 | 99.8 | 99.6 |
| Portugal | 2,019 | 2,251 | 91.5 | 67.2 | 95.1 | 75.5 | 95.9 | 78.1 | 99.7 | 99.3 | 95.0 | 73.5 | 94.6 | 82.0 | 92.6 | 70.1 | 98.4 | 90.2 | 99.7 | 99.1 |
| Greece | 1,379 | 1,938 | 87.1 | 38.4 | 94.5 | 49.0 | 95.1 | 52.1 | 99.9 | 99.4 | 95.9 | 47.1 | 89.7 | 60.6 | 88.5 | 42.1 | 96.4 | 71.1 | 100.0 | 99.3 |

SOURCES.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

NOTE.—Figures in cols. 1–9 represent the proportions of males and females included in the sample across imputation rules of tables 2–4. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education. Sample inclusion rules by column: (1) Employed at time of survey in 1999. (2) Wage imputed from other waves when nonemployed ($-2$, $+2$ window). (3) Wage imputed from other waves when nonemployed ($-5$, $+2$ window). (4) Impute wage < median when nonemployed. (5) Impute wage < median when unemployed. (6) Impute wage < median when nonemployed and education < upper secondary and experience < 10 years; impute wage < median when nonemployed and education $\leq$ higher education and experience $\geq 10$. (7) Impute wage < median when nonemployed and spouse income in bottom quartile. (8) Sample inclusion rules 3 and 6 above. (9) Sample inclusion rule 3 above and wage imputed using probabilistic model (see note to table 4).

high employment gaps, especially in France and southern Europe. Quantitatively, the correction for sample selection is smallest when wage imputation is performed using wage observation from other waves in the panel, and it increases when it is instead performed using observed characteristics of the nonemployed. As argued above, this is mainly due to different sizes of the imputed samples. While only individuals with some degree of labor market attachment feature in the imputed wage distribution in the first case, the use of observed characteristics may in principle allow wage imputation for the whole population, thus including individuals with no labor market attachment at all. Interestingly, the fact that controlling for unobservables does not greatly change the picture obtained when controlling for a small number of observables alone (education, experience, and spouse income) implies that most of the selection role can indeed be captured by a set of observable individual characteristics.[18]

## C. Bounds

Each imputation rule requires assumptions about the position of the nonemployed relative to the median of the potential wage distribution. In order to show that we obtain reasonable estimates for the median wage gap under each specification, we compute bounds following the procedure discussed in Section IV. Table 6 reports "worst case" bounds to the potential distribution for the base sample and for a subset of wage imputation rules.

Column 1 reports bounds using the actual wage distribution to obtain the $F()$ terms in conditions (13) and (14). All estimates for the median wage gap obtained with alternative imputation methods and reported in tables 2–4 lie within the bounds to the potential distribution reported in column 1. Note that, mechanically, the bounds for the gender-specific

[18] We have performed a number of robustness tests and more disaggregate analyses on the results reported in tables 2–4. First, we repeated all estimates using a common set of age weights (obtained from the U.S. 1999 sample) for all countries. Results using such weights were virtually identical to those obtained without weights, and thus variation in the age structure across countries does not seem to explain much of the observed variation in gender pay gaps. Second, for the imputation rules reported in tables 2 and 3, we have repeated our estimates separately for three education groups (less than upper secondary education, upper secondary education, and higher education), and we found that most of the selection occurs between rather than within groups, as median wage gaps disaggregated by education are much less affected by sample inclusion rules than in the aggregate. Finally, we have repeated our estimates separately for three demographic groups: single individuals without children in the household, married or cohabiting without children, and married or cohabiting with children. We found evidence of a strong selection effect in France and southern Europe among those who are married or cohabiting, especially when they have children, and much less evidence of selection among single individuals without children.

**Table 6**
**"Worst Case" Bounds to Median Wage Gaps under Alternative Imputation Rules**

| | 1 | | 2 | | 3 | | 4 | |
|---|---|---|---|---|---|---|---|---|
| | Lower Bound | Upper Bound | Lower Bound | Upper Bound | Lower Bound | Upper Bound | Lower Bound | Upper Bound |
| United States | .153 | .513 | .268 | .397 | .164 | .504 | .218 | .447 |
| United Kingdom | −.042 | .562 | .126 | .373 | .005 | .505 | .045 | .460 |
| Finland | .059 | .255 | .147 | .234 | .131 | .236 | .073 | .236 |
| Denmark | .035 | .155 | .074 | .099 | .064 | .115 | .040 | .150 |
| Germany | −.065 | .521 | .110 | .297 | .013 | .399 | −.025 | .462 |
| Netherlands | −.028 | .406 | .088 | .274 | .072 | .289 | .040 | .322 |
| Belgium | −.184 | .321 | −.061 | .203 | −.070 | .203 | −.110 | .246 |
| Austria | −.030 | .417 | .087 | .288 | .011 | .365 | −.017 | .402 |
| Ireland | −.533 | .815 | −.037 | .494 | −.374 | .680 | −.311 | .639 |
| France | −.668 | .830 | −.089 | .265 | −.387 | .565 | −.430 | .623 |
| Italy | . . . | . . . | −.419 | .377 | −.496 | .485 | −.282 | .360 |
| Spain | . . . | . . . | −.510 | .374 | −.718 | .715 | −.577 | .640 |
| Portugal | −.396 | .470 | −.111 | .323 | −.208 | .376 | −.066 | .315 |
| Greece | . . . | . . . | −.940 | .924 | . . . | . . . | −.496 | .566 |

Sources.—Michigan Panel Study of Income Dynamics and European Community Household Panel Survey.

Note.—In specification 2 no results are reported for Italy, Spain, and Greece as more than 50% of women in the sample are nonemployed. Similarly, this is the case in specification 3 for Greece. Sample description: aged 25–54, excluding the self-employed, the military, and those in full-time education, 1999. Sample inclusion rules by columns: (1) Employed at time of survey in 1999. (2) Wage imputed from other waves when nonemployed (−5, +2 window). (3) Impute wage < median when unemployed. (4) Impute wage < median when nonemployed and education < upper secondary education and experience < 10 years; impute wage > median when nonemployed and education ≥ higher education and experience ≥ 10 years.

median are tighter the higher the employment rate for that gender. Since variation in male employment rates is low relative to variation in female rates, cross-country differences in the tightness of the bounds mostly stem from differences in women's employment selection across countries. Bounds for the median gender wage gap are thus much tighter for the United States, the United Kingdom, and countries all the way down to Austria than they are for Ireland, France, and Portugal—for which they are so large as to be completely uninformative. Indeed, we cannot even obtain bounds to the median wage gap for Italy, Spain, and Greece on the base sample because less than 50% of women are employed.

A restriction typically used to tighten such bounds is that of stochastic dominance (see Blundell et al. 2007), which assumes various forms of positive selection into employment. As this is precisely something that our article is assessing, we cannot use it as an identifying assumption. But we can instead compute bounds after wage imputation, that is, using imputed wage distributions to compute the $F()$ terms in (13) and (14). This procedure has the advantage of tightening the bounds without assuming positive sample selection ex ante. The estimated bounds are reported in columns 2–4 of table 6. In column 2, the wage distribution used is one in which missing wage observations are replaced by observed wages in the nearest available wave (as in col. 2 of table 2). In column 3, missing wage observations are imputed below the median if an individual is unemployed (as in col. 3 of table 3). In column 4, they are imputed using education and experience levels of the nonemployed (as in col. 4 of table 3). As employment rates are higher in columns 2–4 than in column 1, bounds do become tighter. However, they still remain relatively large in southern Europe, where employment rates remain relatively low even after wage imputation.

## VI. Conclusions

Gender wage gaps in the United States and the United Kingdom are much higher than in other European countries, and especially so with respect to France and southern Europe. Although at first glance this fact may suggest evidence of a more equal pay treatment across genders in the latter group of countries, appearances can be deceptive.

In this article, we note that gender wage gaps across countries are negatively correlated with gender employment gaps, and we illustrate the importance of nonrandom selection into work in understanding the observed international variation in gender wage gaps. To do this, we perform wage imputation for those not in work by simply making assumptions on the position of the imputed wage observations with respect to the median. Imputation is performed according to different methodologies based on observable or unobservable characteristics of missing wage observations.

We find higher median wage gaps on imputed rather than actual wage distributions for most countries in the sample, meaning that, as one would have expected, women tend on average to be more positively selected into work than men. However, this difference is small in the United States, the United Kingdom, and a number of central and northern European countries, and it is sizable in France and southern Europe, that is, in countries in which the gender employment gap is particularly high. Our (most conservative) estimates suggest that correction for employment selection explains about 45% of the observed negative correlation between wage and employment gaps. In Italy, Spain, Portugal, and Greece, the median wage gap on the imputed wage distribution ranges between 20 and 30 log points across specifications. These levels are closely comparable to those of the United States and of other European countries.

Another interesting result is that we obtain qualitatively similar estimates whether we impute missing wages using available wage information from other waves in the panel or whether we use observable characteristics of the nonemployed. This implies that employment selection mostly takes place along a small number of measurable characteristics.

Our analysis identifies directions for future work. We argue that gender employment gaps are key to understanding cross-country differences in gender wage gaps. Employment gaps may, in turn, be driven by supply or demand forces, or both. In recent work, Fernández and Fogli (2005) and Fortin (2005, 2006) emphasize the role of "soft variables," such as cultural beliefs about gender roles and family values and individual attitudes toward greed, ambition, and altruism, as important determinants of women's employment decisions as well as of gender wage differentials. These "fuzzy" variables may also shape employers' beliefs about women's labor force attachment and thus the demand for female labor. In addition, labor market and financial institutions, as well as the sectoral composition of the economy, may play an important role in the determination of gender employment gaps. Disentangling supply and demand factors that drive cross-country differences in female employment is thus the next step for understanding existing variation in gender pay gaps.

## References

Albrecht, James, Aico van Vuuren, and Susan Vroman. 2004. Decomposing the gender wage gap in the Netherlands with sample selection adjustment. IZA Discussion Paper no. 1400, Institute for the Study of Labor, Bonn.

Altonji, Joseph G., and Rebecca M. Blank. 1999. Race and gender in the labor market. In *Handbook of labor economics*, vol. 3C, ed. Orley Ashenfelter and David Card, 3141–3259. Amsterdam: North-Holland.

Arulampalam, Wiji, Alison L. Booth, and Mark L. Bryan. 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review* 60, no. 2: 163–86.

Azmat, Ghazala, Maia Güell, and Alan Manning. 2006. Gender gaps in unemployment rates in OECD countries. *Journal of Labor Economics* 24, no. 1:1–37.

Beblo, Miriam, Denis Beninger, Anja Heinze, and François Laisney. 2003. Measuring selectivity-corrected gender wage gaps in the EU. ZEW Discussion Paper no. 03-74, Centre for European Economic Research, Mannheim.

Blau, Francine D., and Andrea H. Beller. 1992. Black-white earnings over the 1970s and 1980s: Gender differences in trends. *Review of Economics and Statistics* 74, no. 2:276–86.

Blau, Francine D., and Lawrence M. Kahn. 1996. Wage structure and gender earnings differentials: An international comparison. *Economica* 63 (May): S29–S62.

———. 2003. Understanding international differences in the gender pay gap. *Journal of Labor Economics* 21, no. 1:106–44.

———. 2006. The US gender pay gap in the 1990s: Slowing convergence? *Industrial and Labor Relations Review* 60, no. 1:45–66.

Bloomfield, Peter, and W. L. Steiger. 1983. *Least absolute deviations: Theory, applications and algorithms*. Boston: Birkhauser.

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, no. 2:323–63.

Buchinsky, Moshe. 1998. The dynamics of changes in the female wage distribution in the USA: A quantile regression approach. *Journal of Applied Econometrics* 13, no. 1:1–30.

Chandra, Amitabh. 2003. Is the convergence in the racial wage gap illusory? NBER Working Paper no. 9476, National Bureau of Economic Research, Cambridge, MA.

Fernández, Raquel, and Alessandra Fogli. 2005. Culture: An empirical investigation of beliefs, work, and fertility. NBER Working Paper no. 11268, National Bureau of Economic Research, Cambridge, MA.

Fortin, Nicole M. 2005. Gender role attitudes and the labour market outcomes of women across OECD countries. *Oxford Review of Economic Policy* 21, no. 3:416–38.

———. 2006. Greed, altruism, and the gender wage gap. Unpublished manuscript, Department of Economics, University of British Columbia.

Goldin, Claudia. 1990. *Understanding the gender wage gap: An economic history of American women*. Oxford: Oxford University Press.

Gronau, Reuben. 1974. Wage comparison: A selectivity bias. *Journal of Political Economy* 82, no. 6:1119–43.

Heckman, James J. 1974. Shadow prices, market wages, and labor supply. *Econometrica* 42, no. 4:679–94.

———. 1979. Sample selection bias as a specification error. *Econometrica* 47, no. 1:153–63.

———. 1980. Addendum to sample selection bias as a specification error. In *Evaluation studies review annual*, vol. 5, ed. Ernst W. Stromsdorfer and George Farkas. Beverly Hills, CA: Sage.

Johnson, William R., Yuichi Kitamura, and Derek Neal. 2000. Evaluating a simple method for estimating black-white gaps in median wages. *American Economic Review Papers and Proceedings* 90:339–43.

Juhn, Chinhui. 1992. Decline of male labor market participation: The role of declining market opportunities. *Quarterly Journal of Economics* 107, no. 1:79–122.

———. 2003. Labor market dropouts and trends in the wages of black and white men. *Industrial and Labor Relations Review* 56, no. 4:643–62.

Krueger, Alan B., and Lawrence H. Summers. 1988. Efficiency wages and the inter-industry wage structure. *Econometrica* 56, no. 2:259–93.

Lewbel, Arthur. 2007. Endogenous selection or treatment model estimation. *Journal of Econometrics* 127, no. 2:777–806.
Manski, Charles F. 1989. Anatomy of the selection problem. *Journal of Human Resources* 24, no. 3:343–60.
———. 1994. The selection problem. In *Advances in econometrics, Sixth World Congress*, vol. 1, ed. Christopher A. Sims. Cambridge: Cambridge University Press.
Mulligan, Casey B., and Yona Rubinstein. 2008. Selection, investment, and women's relative wages over time. *Quarterly Journal of Economics* 123, no. 3:1061–1110.
Neal, Derek. 2004. The measured black-white wage gap among women is too small. *Journal of Political Economy* 112 (February): S1–S28.
OECD (Organization for Economic Cooperation and Development). 2002. *Employment outlook*. Paris: OECD.
Olivetti, Claudia, and Barbara Petrongolo. 2006. Unequal pay or unequal employment? A cross-country analysis of gender gaps. CEPR Discussion Paper 5506, Centre for Economic Policy Research, London.
O'Neill, June. 1985. The trend in the male-female wage gap in the United States. *Journal of Labor Economics* 3, no. 1, pt. 2:S91–S116.
Smith, James P., and Michael Ward. 1989. Women in the labor market and in the family. *Journal of Economic Perspective* 3 (Winter): 9–24.
Smith, James P., and Finis R. Welch. 1986. *Closing the gap: Forty years of economic progress for blacks*. Santa Monica, CA: RAND Corporation, R-3330-DOL.
Vella, Francis. 1998. Estimating models with sample selection bias: A survey. *Journal of Human Resources* 33, no. 1:127–69.