Motherhood and the Gender Productivity Gap[*]

Yana Gallen[†]

January 26, 2018

Abstract

Using Danish matched employer-employee data, this paper estimates the relative productivity of men and women and finds that the gender "productivity gap" is 8 percent, implying that just under two thirds of the residual wage gap can be accounted for by productivity differences between men and women. I measure the productivity gap by estimating the efficiency units lost in a firm-level production function if a worker is female, holding other explanatory covariates such as age, education, experience, occupation, and hours worked constant. Both mothers and non-mothers are paid less than men, but the (low) relative pay of mothers is completely explained by productivity differences. In contrast, women without children are estimated to be as productive as men but are paid less. The decoupling of pay and productivity for women without children happens during their prime-child bearing years. These estimates are robust to a variety of specifications for the impact of observables on productivity, and robust to accounting for endogenous sorting of women into less productive firms using a control-function approach. This paper also provides estimates of the productivity gap across industries and occupations, finding the same general patterns for mothers compared to women without children within these subgroups.

JEL Classification: J71, J31, J24
Keywords: Discrimination, Wage Gap, Labor Productivity

---

[*]Thank you to my advisor, Matthew Notowidigdo, and my committee members Lori Beaman and Jon Guryan for their very helpful advice and comments. I also thank David Berger, Stéphane Bonhomme, Matthias Doepke, Martin Eichenbaum, Trevor Gallen, Brent Hickman, Rune Vammen Lesner, John List, Lee Lockwood, Casey Mulligan, Seema Jayachandran, Rune Vejlin and participants at the Northwestern Applied Micro Reading Group and Applied Micro Seminar, ASU, the Chicago Federal Reserve, Conference on Labour Market Models and their Implications, Cornell, CeMENT, Harris School, H2D2 conference, DePaul, Duke, Nordic Data Workshop in Oslo, Purdue University, UIC, and UIUC LER for helpful comments and discussions. Thank you to Henning Bunzel and LMDG, Department of Economics and Business, Aarhus University for access to the data. LMDG is a Dale T. Mortensen Visiting Niels Bohr professorship project sponsored by the Danish National Research Foundation. All mistakes are my own.

[†]Correspondence: University of Chicago Harris School of Public Policy, 1155 E 60th St, Chicago, IL 60637. Tel.: (773) 834-2784. Email: yana@uchicago.edu.

1 Introduction

What can explain the gender wage gap? Natural first candidates are differences in the characteristics of men and women which affect productivity: the occupations they choose, the experience they have, the education they have. Wage regressions which control for these productivity-affecting characteristics do reduce the pay gap, but not completely. What can explain the residual gender wage gap? One possibility is discrimination, but a less explored possibility is that women are being paid less than men because they are less productive for unobservable reasons. In this paper, I will describe how much (or little) of the difference in earnings for men and women can be explained by differences in their productivity. In Danish administrative data, in addition to having a rich set of worker characteristics, it is possible to directly study how firm output varies with the gender of employees. I find that about 8 percentage points of the 12 percent residual pay gap can be explained by productivity differences between men and women.

To measure the productivity gap, I estimate a firm-level production function that takes labor, material goods, and capital as inputs and treats male and female labor units as perfect substitutes. The gender productivity gap is the efficiency units lost if a worker is female, holding other explanatory covariates such as age, education, experience, and hours worked constant. I use Danish data which matches worker characteristics with their firm's accounting information in order to estimate the gender productivity gap. Whether discrimination against women in the labor market exists, and if so whether it takes the form of unequal pay for equal work, or simply unequal work is central to the debate of how to remedy the wage gap. Nonetheless, productivity differences are not often studied[1] as sources of the gender pay gap. This is in part because high-quality data on revenue and inputs linked to employee characteristics is rarely available.

I find a non-negligible productivity gap in Denmark, which implies that on average, discrimination does not take the form "unequal pay for equal work." However, this average masks differences between women over the lifecycle, particularly differences for mothers compared to non-mothers. Motherhood is associated with large wage and earnings penalties (Kleven et al. [2015], Angelov et al. [2016]). This paper is the first to link parenthood by gender to productivity measures.

For mothers, I find that the earnings gap is somewhat smaller than the productivity gap, suggesting that there is little or no discrimination (in the form of uncompensated output) against

[1] Hellerstein et al. [1999] leads the exceptions, which I will discuss in detail in the next section

mothers. This is consistent with the literature suggesting that the wage gap occurs only for women with children who work fewer and more flexible hours than their male counterparts (see for example Goldin [2014], Gicheva [2013], and Kleven et al. [2015]) and that there may be some output loss associated with these work arrangements. Though there is no evidence of discrimination (in the form of uncompensated productivity) against mothers, there is evidence of discrimination against women without children. While wage gap is smaller for women without children, women without children are actually more productive than similarly aged men. I find that the disparity between wages and productivity for non-mothers happens especially during their prime child-bearing ages. After age 40, there are no meaningful differences in the relative productivity of mothers and non-mothers. Discrimination, then, is largest in the group with a smaller residual pay gap (non-mothers)[2].

The large gap between the productivity and pay of women without children may reflect employer's statistical discrimination: if productivity falls with motherhood but employers cannot lower wages immediately, then employers may offer lower wages to productive women in anticipation of motherhood. The gap between pay and productivity may also reflect employee's own desire to smooth wages over the lifecycle. I do not find a larger gap between pay and productivity for married or cohabiting women compared to single women without children. To the extent that marriage and cohabitation increase the employer's expectations about and employee's probability of childbirth, this test suggests that statistical discrimination is unlikely to be the main driver of the wedge between pay and productivity for prime-age non-mothers.

I present estimates of the productivity gap in the cross-section, over time, by industry, and accounting for selection of workforce composition based on unobservables. Selection is a problem for estimating the true productivity gap if women sort into firms with lower total factor productivity. In this case, the estimate of the relative productivity of men and women will reflect both the true difference in their productivity if they were randomly assigned to firms and the average difference in the TFP of firms where women work relative to firms where men work. To control for sorting, I use a control function approach of Olley and Pakes [1996]. If some component of TFP is known to the firm at the time they make their decision to hire a woman

[2]Here I label differences between average pay and average productivity by gender/parenthood "discrimination," as the wage gap literature often does. These differences between pay and productivity by gender/parenthood could arise from gender differences in worker preferences, for structural reasons, or due to firm-level statistical or taste-based discrimination.

relative to a man, then this portion of TFP will also influence their investment decision. The firm's investment rule will be monotonic in the unobservable (conditional on capital) and can be inverted to approximate the unknown component of TFP which influences hiring decisions. A flexible polynomial in capital and investment approximates the unobserved component of TFP which is correlated with hiring decisions. This control does not change the overall estimate of the relative productivity of men compared to women. Consistent with the small role for selection in estimating relative productivity via the production function, using a wage decomposition as in Card et al. [2016], I also find no evidence that women work in lower wage firms within this subset of relatively large, private sector Danish firms.

The paper proceeds as follows: Section 2 reviews the relevant literature. Section 3 describes the data used in estimation. Section 4 provides the model and estimating equations. Section 5 presents results and Section 6 concludes.

2 Related Literature

Most literature on the gender pay gap has focused on explaining differences in the relative pay of men and women using wage regressions, finding that occupational choice and (more historically) human capital differences between men and women are important drivers of the average difference in pay. Altonji and Blank [1999] provides an overview of the early literature on the gender wage gap, highlighting the role of differences in preferences, comparative advantage, and human capital accumulation in models of gender wage differentials, with discrimination typically playing the role of the residual, unexplained portion of the gender wage gap. They find that after controlling for education and occupational, industry, and job characteristics, the wage gap in 1995 was 22 percent. Blau [1977] argues that occupational choice plays a very large role in the gender wage gap since inter-firm wage schedules for a given occupation are constrained by a sense of inter-office fairness. Blau and Kahn [2017] re-visit the wage gap and find that in recent years, even more than in the 80s, occupational choice explains a large fraction of the pay gap in the US.[3]

Mulligan and Rubinstein [2008] study the changing nature of female selection into the labor force from the 1970s to the 1990s. They find that while the lowest-skilled women entered the

[3] However, using Danish data, Gallen et al. [2017] do not find an increasing role for occupation over time.

full time labor force in the 1970s, the highest-skilled women enter the full time labor force the 1990s, implying that most of the apparent narrowing of the gender wage gap resulted from compositional changes in the female labor force. Overall, the wage gap literature finds that, despite the important role of occupation on wages (see Goldin [2014], for example), the wage gap has persisted over time and a large portion of the gap is unexplained by observables. This paper differs from the wage-gap literature by not using wages at all in estimation. Instead, I estimate the relative output of a firm that hires a man compared to a woman with the same background, controlling for the possible endogeneity in that decision. In the same spirit as this study, Azmat and Ferrer [2017] document the difference between hours billed and new clients brought in to the firm for male vs. female lawyers. They find that the large differences in the earnings of male and female lawyers (particularly mothers) are largely explained by these measures of productivity.

Perhaps more closely related to the method in this paper, Hellerstein et al. [1999] study the relationship between wage gaps and gaps in marginal product for a variety of observable characteristics in the manufacturing industry. They find that with the exception of gender, differences in wages based on observables are equal to differences in marginal productivity. Methodologically, the authors follow a similar path the one I outline at the beginning of the next section. The authors estimate labor as the sum of labor of different types—male/female, black/white, under 35/35-54/55 and over, less than college/college, unskilled/managers/skilled/ administrative, married/single. The authors estimate an unusually large gender wage gap of -0.45 in their data, but find a gender productivity gap of -0.16. Studying the interaction of gender and occupation, the authors note that the finding of gender-discrimination is driven by non-managerial and non-professional worker groups (I find the opposite). Interacting gender and age, the authors find significant evidence of discrimination only for young workers (I find something similar for women without children).

In this paper, I examine the role of childbearing in explaining the productivity gap. This paper is the first to examine the relationship between the effect of children on the productivity of women (vs. men) and compare this to their effect on wages. In addition, I address selection into firms, which could generate bias the estimation of the productivity gap. Another advantage of my study is the breadth of data I am able to use. In particular, manufacturing (the only industry available to Hellerstein, Neumark, and Troske) is a mostly male industry—69% of all workers are male. The nature of work done in manufacturing (much of it involving manual labor) makes it difficult to believe that women and men are doing the same jobs. I am able to study industries where we would not expect stringent gender-based occupational sorting, and in which women make up a large part of the workforce.

Differential sorting between men and women may reflect preferences, or it may reflect a different type of discrimination. Women may prefer working in low-wage firms because these firms allow more flexible hours. Goldin [2014] argues that women prefer flexibility in hours and work in occupations and choose career paths that allow for hours flexibility, losing the monetary compensation associated with long hours and full availability (such as what is required by many high wage jobs in finance and law). Wasserman [2017] finds evidence that women do steer away from occupations with long hours by studying residency choices of medical students before and after a change in maximal shift lengths.

Card et al. [2016] find that women sort into different firms than men. Using firm accounting data, the authors find that about one-fifth of the gender wage-gap in Portugal can be explained by the dual channels of bargaining and sorting. Differential sorting by men relative to women explains most of the difference in firm effects from an AKM decomposition. Though both use administrative data and firm accounting data to study the gender wage gap, this paper studies a portion of the gender wage gap unexplained by Card et al. [2016]. While Card et al. [2016] focuses on firm effects, differencing out individual-level productivity, this paper studies the difference, on average, between male and female productivity and attempts to correct for the endogeneity generated by sorting. The result that women and men sort into different firms suggests that a basic cross-sectional study of the gender productivity gap will yield biased estimates of the relative productivity of men compared to women. Applying the AKM model to the Danish subset of private sector firms used in this analysis, I do not find evidence of sorting by women into lower-pay firms (controlling for individual effects). Nonetheless, I account for this sorting by using panel data with an Olley-Pakes correction for endogeneity of inputs, which I discuss in more detail in the model section of the paper.

In this paper, I focus on identifying one particular form of discrimination: differences in pay unexplained by differences in output. This type of discrimination would occur if, for example, women did not bargain as well as men for raises (Babcock and Laschever [2003], Leibbrandt and List [2015]) or if firms did not pass improvements in productivity on to female employees as much as male employees, as in Card et al. [2016]. Alternative forms of discrimination are certainly possible and important to understand, but they are not the subject of this paper. Another way in which the wage gap may result from discrimination is if women are not offered jobs at high productivity firms, or if women are not invested in or offered promotions despite being equally able to work in more demanding jobs (Thomas [2015], Stearns [2017]). This type of discrimination, often called "mommy tracking" is difficult to distinguish from preferences, but may occur if firms are sufficiently risk averse and the distribution of female productivity differs from that of male productivity. To find evidence of this type of discrimination, one would need to measure potential output of workers in positions which they are not offered. Albrecht et al. [2015] and Albrecht et al. [2003] offer evidence that promotions of women in Sweden are limited due to employers' beliefs that women will have children in the future. "Mommy track" discrimination, both interesting and important, is not addressed here. Instead, I focus exclusively on the link between realized output and pay.

One advantage of the Danish data relative to US or Portuguese MEE data is the availability of information about a worker's family, namely whether or not they have children. A consistent finding in the gender wage gap literature is that the divergence in the pay of women relative to men happens primarily during the childbearing years. Most relevantly for this paper, Kleven et al. [2015] use Danish data to understand the relationship between motherhood and the gender wage gap. While the presence of children can explain 30% of the gender earnings gap in 1980, children can explain 80% of the gap in 2011. The "child-penalty" comes in the form of (roughly equally) lower labor force participation of mothers, fewer hours of work for mothers, and lower wage rates for mothers. Angelov et al. [2016] document similar penalties for motherhood in the Swedish data. Adda et al. [2011] study the relationship between fertility and wages in a dynamic model with human capital accumulation, career choice, and labor supply decisions. Using German administrative data, the authors find that fertility choices shift the earnings profile of women and explain a good portion of the wage gap.

There is a large body of literature documenting the differences between women and men which may explain the gender wage gap, but are more subtle than differences in human capital accumulation, child-rearing, and occupational choice. As reviewed by Niederle and Vesterlund [2011], women have been documented in both the lab and the field to be less competitive than men, conditional on performance.[4] Gneezy et al. [2009] argue that this link between gender and competition is reversed in a matrilineal society, implying that most of the link is driven by cultural rather than biological differences between men and women. Ichino and Moretti [2009] argue that women are more likely than men to be absent in 28-day intervals, suggesting that menstruation may increase female absenteeism. However, Rockoff and Herrmann [2012] dispute this finding.

Small biological difference may turn into large differences in career pathways when mediated by social norms. As discussed by Fryer and Levitt [2010] differences between boys and girls in mathematical ability expand over time, also suggesting a role of culture. Bursztyn et al. [2017] find that single female MBAs change their stated preferences concerning flexible jobs when their answers are observed by male MBA colleagues. Bertrand et al. [2010] find substantial divergence between male and female MBA student outcomes, despite similar initial conditions, driven by female career interruptions from childbirth. Decades ago, Becker [Becker, 1985] hypothesized that differences in demands on or abilities of women in the home production sector translate to differences in career choice. Hersch and Stratton [1997] provide early evidence of the effect of housework on market wages. The importance of finding flexible work arrangements to accommodate childrearing is emphasized in Goldin [2014] and Blau and Kahn [2013].

This line of research links to the gender wage gap largely through the mechanism of occupational choice. Pan [2015] documents the phenomenon of occupational "tipping": when a sufficient fraction of an occupation is female, the wages in that occupation fall. Blau and Kahn [2017] argue that occupational choice has been an increasingly important explanatory factor when decomposing the wage gap over time. In addition, gender preferences can translate to differences in wages directly via compensating differentials: when risk-taking or competitive behavior is rewarded and women shy away from risky jobs, they will on average by paid less than men. Babcock and Laschever [2003] study the gender gap through the lens of salary negotiations. They find that among Master's students at Carnegie Mellon University, female graduates negotiated their starting salary 7% of the time. In contrast, male graduates negotiated their starting salary 57% of the time. The authors argue that a large portion of the gender earnings gap can be linked to a lower propensity by women to ask for raises. This mechanism would imply

[4] This competitiveness factor has been studied extensively in recent years. See for example Bertrand [2011], Croson and Gneezy [2009], Flory et al. [2015], Buser et al. [2014], Markussen et al. [2014], Kamas and Preston [2012], Berge et al. [2015], Zhang [2015], and Reuben et al. [2015].

that the gap in earnings between men and women is much larger than the gap in productivity between men and women. It is precisely this type of mechanism that I test in the paper.

Women's wages may be lower than men's both due to statistical discrimination (if they are indeed less productive than men on average) as described in Aigner and Cain [1977] or due to taste-based discrimination, as described in Becker [1971]. Gunderson [1989] discuss the very different policy implications of statistical and taste-based discrimination. Since his study, a large literature has emerged which finds some support for taste-based discrimination: Weber and Zulehner [2014], Hellerstein et al. [2002], and Heyman et al. [2013] test and find evidence for the theory first proposed by Becker [1971] which notes that discrimination by employers is costly in the long run, since labor markets are relatively competitive. Weber and Zulehner [2014] and Hellerstein et al. [2002] find that discrimination is correlated with slower firm growth and shut-down. Heyman et al. [2013] find that when markets are less competitive, firm takeover is correlated with an increase in the female labor share and a reduction in the gender pay gap.

Discrimination can, of course, operate in a variety of ways. Firms may discriminate on hiring, but conditional on hiring a woman, compensate her in the same way they would compensate a man. Goldin and Rouse [2000] find that when orchestras moved to blind auditions, women were more likely to be hired. On the intensive margin, firms can compensate women at a lower rate than men, either by paying the less to do the same job, or by slotting them into less productive tasks. Even when offering lower wages, discriminating firms may survive if there is a large match-specific or firm-specific component to productivity. Though in Becker [1971] discrimination doesn't result in a wage differential when the marginal employer doesn't discriminate, a large search literature notes that wage differences emerge in frictional labor markets even when small fraction of employers discriminate (see Black [1995], Rosen [1997], Bowlus and Eckstein [2002], and more recently Bond and Lehmann [2015]). Discriminatory firms paying lower wages to women survive if jobs are scarce.

Summarizing the literature on the gender wage gap is both a simple and arduous task—countless studies of the relative wages of men and women have found gaps which have neither disappeared over time nor when considering observable differences between men and women. This paper does not ask what explains changes in the gap over time and it does not ask why women and men differ on observables. Instead, this paper studies the (large) residual that remains in the wage gap when controlling for these observable differences and its relationship to the relative output of men and women, asking how much of the residual wage gap can be explained by differences in the residual marginal product of men compared to women.

3 Data

The data used in this paper are from three primary sources: a relatively new Danish register on employees called eIncome, a more commonly used Danish register on employees called IDA, and a detailed survey of firm accounts, called Regnskabsstatistikken (abbreviated FIRE). eIncome is register data covering all employees working in Denmark, from 2008. The data is reported monthly, by all employers to the Danish Customs and Tax Administration, who pass the data to Statistics Denmark's eIncome Register to be used in calculation of national statistics at the monthly level. The primary advantage of this dataset is that it reports work by all employees in given firm: their occupation, total pay for that month, and total hours worked in the month (as well as the dates of employment).

This dataset is distinct from the commonly used Danish IDA dataset, which is annual and has data on payments and hours based on the worker's status in November of that year. The hugely improved hours variables in eIncome shrink the wage gap considerably. In particular, the gap falls from 16% to 10%. The main improvement is better tracking of workers who are not continuously employed in a firm and enter and exit the data only for a few months. In the IDA dataset, using bracketed hours worked and a November employment measure it is difficult to properly assign total hours worked at a given firm. Bracketing alone accounts for a two percentage point increase in the pay gap. Of course, there is some non-response even in the eIncome dataset—about 15% of the hours data is imputed. All employees are included regardless of hours worked. Main jobs and side jobs are included. Employees who are not residents of Denmark are included. If an employees doesn't have pay for up to 45 days at a job, but subsequently returns to the same employer for pay (for example for training), he or she is included in the data for the months without pay.

The hours worked measure in IDA is based on employee contributions to retirement benefits. The brackets are four bins of weekly hours (0-8, 9-17, 18-27, 27+) or four bins of monthly hours, (0-38, 39-77, 78-116, 117+). The data also measure the fraction of the year worked. There is a large fraction of workers whose hours are not distinguished from one another but may in reality differ substantially. eIncome is not completely immune to this problem, though it is certainly less severe. In eIncome, salaried workers would be listed as working 37 hours a week, unless they clocked in overtime hours. Many likely do not, and work slightly less or slightly more than 37 hours per week. There is no reason to think this is orthogonal to gender—women work fewer hours on the margins we can measure, they may also work fewer hours on the margins we have more difficulty measuring. Indeed US time-use data (the American Time Use Survey) suggests that conditional on working full-time, mothers of older children work about one hour less per day than fathers and mothers of young children work about 40 minutes less per day than fathers[5].

The eIncome register can be linked with data on firm value added using a dataset called FIRE which contains information on firm accounts. The FIRE employer data is the basis for national accounts. As in Baggar, Christiansen, and Mortensen (2014), I follow the methodology for constructing value added and capital stock used in national accounting. The details of this procedure exactly follow Baggar, Christiansen, and Mortensen (2014) and are discussed in the data appendix. FIRE includes information on firms from tax records (such as revenue and the value of capital) and also contains detailed accounting measures from survey. Firms are surveyed based on size. Firms with more than 50 employees are surveyed annually, firms with 20-49 employees are surveyed every other year, firms with 10-19 employees are surveyed every 5th year, and firms with 5-9 employees are surveyed every 10th year.

Firms which are not in the survey in a given year have some of their information imputed into the dataset, though much of the imputation comes from tax records. Detailed information on the cost of intermediate goods, however, is completely imputed for a large fraction of firms in the data. My measure of value added is revenue less the cost of these intermediate inputs so the measurement error generated by using imputed values is on the left hand side and does not systematically bias results. When information is imputed, it is based on industry-level averages weighted by employment and revenue. In the results reported, I use all data which was not completely imputed (that is, data taken from tax records combined with survey results). About 9,000 firms are actually surveyed in each year [6].

To supplement the worker-level information available in the income registers merged with

[5] author's own calculations using ATUS 2010

[6] Restricting only to the set of firms surveyed in detail about their accounts, I estimate the relative productivity of women is 0.94, which is about 1-2 percentage points higher than my baseline estimates and not a statistically significant difference. The cost of using only firms with were actually surveyed is that I would not have power to study differences across industries, across occupation, by age, etc.

FIRE—which is essentially occupation, hours worked, wages, and industry—I add demographic information on workers from IDA. This includes information on the birthdate of all children, year of marriage, education including major, age, and experience (which is constructed as the sum of hours worked in the labor market according to IDA).

3.1 Summary statistics

The earnings gap in Denmark is surprisingly similar to the gap in the US. Table 1 below provides estimates of the earnings gap in the US from Goldin [2014] compared to a similar population in Denmark and compared to my restricted sample of large industries in the FIRE database. The raw earnings gap is smaller in Denmark than in the US but it also is less explained by controlling for hours, education, and occupation. The smaller raw gap is consistent with Blau and Kahn [2003] who find that countries with more compressed wage distributions (such as Denmark) have smaller wage gaps.

The Denmark and US samples are restricted to ages 25-64. In Denmark, the raw gap is 27.7 log points, compared with 32 log points in the US. Controlling for age, hours worked, education, and occupation, the gap falls to 17.2 log points, compared with 19.1 log points in the US. The R-squared from the last wage regression in the US is about twenty percentage points lower than in Denmark. The lower R-squared in the US may reflect noise expected from survey data. Another explanation for the difference in the explanatory power of observables across countries may be that Denmark has a more compressed wage distribution (so there is less wage variation to explain). In addition, unions and collective bargaining determine wages to a far greater extent in Denmark than the US. For a large fraction of workers, wage increases resulting from collective bargaining are determined by tenure and education (see Dahl et al. [2013] for a detailed description of wage bargaining in Denmark). Anecdotally, Denmark has a strong culture of fairness and may prefer pay to be more closely linked to observables relative to performance measures such as effort, for example.

One advantage of the Danish register data compared with the American Community Survey survey in the US is that it provides information on the experience of a worker and also on the firm ID of the worker. Earnings may depend on experience (and women who take time off work to have children may have a different level of experience than men on the same age). Manning and Robinson [2004] argue that this difference in experience explains much of the gender pay gap in the British Household Panel Survey. Earnings may also vary by firm for observationally identical workers. This may reflect differences in non-wage compensation at different firms and in the presence of gender sorting may explain some of the earnings gap. In Table 2 below, I report the results of a regression of log earnings on hours, a quadratic in age, and sequentially add controls for 1. a quadratic in experience and education level dummies, 2. occupation and industry fixed effects (at the 3-digit ISCO level and the 2-digit NACE level, respectively), and 3. the interaction of firm fixed effects and occupation fixed effects.

Adding controls available with the rich Danish data, such as experience and occupation only causes the wage gap to fall slightly. Adding firm and occupation interactions and identifying the earnings gap using differences in the pay of women and men within a firm in a given occupation narrows the earnings gap to just under 12 percent.

I focus my analysis on the six industries (measured at the two digit level) which have the largest number of firm-year observations in the FIRE database: Accommodation and food services, Construction, Manufacturing, Wholesale and retail trade, Other community, social and personal services, and Real estate, renting, and business activities. These make up more than 50 percent of the Danish economy[7] and 98% of non-imputed firm-year observations. Table 3 below provides some summary statistics for the firms in each industry and the dataset overall.

The wage gap varies by industry, ranging from 11%-19%. The fraction of the workforce in a given industry which is male also varies. In construction, a very large fraction of the labor force is male, while in accommodations and food services, fewer than half of workers are male. Notably, this study of productivity differences is focused on industries with relatively more men than average. Because there are no accounting statistics for public sector firms, this large portion of the Danish economy (and place of employment for women, disproportionately) is omitted from the analysis. The potential biases from this omission will be discussed in the next section.

4 Model and Estimation

The goal of this paper is to understand whether differences in the productivity of men and women, conditional on a set of observable characteristics, explains differences in the wages of men and women, conditional on the same set of observable characteristics. To answer this

[7]Measured by 2010 gross value added by industry tables available from Statistics-Denmark

question, I estimate production functions assuming output is affected by the quantity of capital, the quantity of intermediate inputs, and the quantity of labor. Labor will be the sum of male and female labor, where I allow a unit of female labor to be more or less productive than a unit of male labor

$$\mathcal{L} = \beta L^F + L^M$$

Firms take these inputs and produce using some function F.

$$Y_{jt} = A_{jt} F(\mathcal{L}_{jt}, K_{jt}, M_{jt})$$

β measures the labor-preserving tradeoff between men and women: a $\beta < 1$ implies that women are less productive than men and $\beta > 1$ implies that women are more productive than men. Estimating this parameter, β, is the focus of this paper. I discuss the details of this function F in subsection 4.2. First, I discuss the measurement of labor.

4.1 Measuring labor inputs

As first noted in Griliches [1957] and more recently Fox and Smeets [2011], there is a difficulty in measuring the quantity of labor a firm has: an individual with a college degree produces more than an individual with primary school; an executive (likely) produces more than a janitor. When measuring the labor a firm has access to, one must account for the *quality* of that labor. There are a number of possible ways to account for the quality of labor, which I discuss below.

My baseline specification relies on constructing efficiency units of labor from the market-wide male wage equation. In construction of the efficiency units, I assume that a woman with the same characteristics as a man would have the same returns to those characteristics, and attribute any deviation from this to gender-based productivity differences. I control for the quality of various characteristics by running an efficiency units regression on the subsample of male workers in the data of the form

$$\begin{aligned}\ln e_{it}^M = \alpha^M + \gamma_1^M Age_{it} + \gamma_2^M Age_{it}^2 + \gamma_3^M Exp_{it} + \gamma_4^M Exp_{it}^2 + \gamma_5^M HS_{it} + \gamma_6^M Col_{it} + \gamma_7^M BA_{it} + \\ + \sum_{j=1}^{4xx} \delta_j^M H_{it}^j + \sum_{o=1}^{NOCC} \omega_o^M \mathbf{1}\{OCC = o\}_{it} + \sum_{t=2000}^{2010} \phi_t^M Year_t + \varepsilon_{it}\end{aligned} \tag{1}$$

where e_{it}^{M} is male worker i's labor market earnings in year t. Age, Exp, HS, Col, and BA measure a worker's age, hours of experience in the labor market, whether the worker has a high school or less, some college or trade school, or further education, respectively. The efficiency units regression also includes indicators of a worker's occupation at the three digit ISCO level, and year fixed effects. In this baseline specification, the amount of male labor in the firm J at time t is

$$L_{J(t)}^{M} = \sum_{i \in J(t),M} \widehat{e_{it}^{M}}$$

Like male labor in the firm, female labor is also measured using the returns to age, experience, occupation, education, and hours from the male wage equation (1), so that the amount of female labor in firm J at time t is

$$L_{J(t)}^{F} = \sum_{i \in J(t),F} \widehat{e_{it}^{M}}$$

Whether these characteristics are truly paid their marginal product is an interesting and potentially important question. Men in a given firm will on average be older than women that the firm. If younger workers are systematically underpaid, that will affect the interpretation of the gender productivity gap. The literature is not conclusive on this issue[8], but I estimate the gender productivity gap by age bins which should alleviate this concern.

I also consider more detailed estimates of efficiency units—using eIncome data gives a finer measure of hours, and I allow for the interaction of occupation with industry and year in the calculation of efficiency units. Finally, eIncome includes monthly earnings based on both "take-home" pay (the narrow earnings definition) and earnings including benefits such as contributions to retirement accounts, the value of a free full-year of residence, the value of free summer residence, the value of a free pleasure boat, the value of a free TV license, the value of a free phone, the value of a PC, anniversary and severance pay, bonus income, and the value of other employee benefits. I use this broad definition of earnings when constructing efficiency units as a robustness check. If high-skilled men preferred to be paid in, for example, boats (or more realistically in retirement contributions) rather than having money in the bank, then

[8]For example, Hellerstein et al. [1999] find a discrepancy between wage and marginal product only for gender. Dostie [2011] uses more age categories and finds on average concave wage and productivity profiles, where wages do not deviate significantly from productivity. However, Hellerstein and Neumark [2007] finds some evidence that wages are deferred over the life cycle.

my efficiency units calculations would understate the returns to skill specifically for men. More generally, and differences between the wage gaps using the two measures of income are potentially important for understanding the potential role for non-pecuniary benefits in generating wage differences.

An alternative specification of the labor units in a firm follows the Griliches [1957] method which has

$$\begin{aligned}\mathcal{L} = L\left[1+(\phi_F-1)\frac{F}{L}\right]\cdot\left[1+(\phi_R-1)\frac{R}{L}\right]\cdot\left[1+(\phi_p-1)\frac{P}{L}+(\phi_o-1)\frac{O}{L}\right] \\ \cdot\left[1+(\phi_N-1)\frac{N}{L}+(\phi_S-1)\frac{S}{L}+(\phi_C-1)\frac{C}{L}\right]\end{aligned} \tag{2}$$

where F is the number of female workers, R is the number of married workers, G is the number of college workers, P is the number of 35-54 year old workers, O is the number of workers 55 and older, N is the number of unskilled laborers, S is the white collar, technical, and sales workers, and C are the number of high skilled workers. This categorization is (exactly) used in Hellerstein et al. [1999][9] to capture the quality of a firm's labor force. ϕ_F is then equivalent to β as a measure of the productivity difference between a unit of female labor and male labor, accounting for differences in age, marital status, education, and occupation. Both methods assume that workers of different ages, occupations, etc. are perfect substitutes and that observable characteristics factor multiplicatively into productivity. The advantage of (1) is that I can account more flexibly for returns to age, experience, hours worked, and occupation (rather than using large discrete bins), and this flexibility ties estimates directly to traditional estimates of the residual wage gap—I am able to measure the residual productivity gap accounting for the same differences in observables commonly used when measuring the residual wage gap. I discuss this in more detail in the next section. Since there are advantages to both methods, I present estimates of the productivity gap using both (1) to predict the labor units in a firm and estimating the production function using labor as in (2).

Fox and Smeets [2011] directly tackle the problem of measuring labor quality in production functions by experimenting with a variety of approaches in Danish data similar to the data I use in this paper. They find that despite the biases associated with measuring the quality of labor using the wage bill, these estimates preform as well as estimates using Griliches-type

[9] They also had a category for Black workers which I omit in the context of Denmark.

specifications. In principle, any firm-specific innovations to productivity get passed on to workers in a bargaining setting, so using the wage bill to stand in for labor quality correlates covariates and error terms. Empirically, however this does not seem to be a large source of bias. Using the wage-bill method, labor in the firm is given by $\mathcal{L} = \beta W^F + W^M$ where W^F is the sum of earnings of women employed at the firm and W^M is the sum of earnings of men employed at the firm. The benefit of using the firm's wage bill method is that one asks whether a dollar paid to a man is as productive as a dollar paid to a woman, and one does not need to take a stand on how to model differences in returns to observables—they may differ and this is all summarized in wage differences. A discriminatory firm which pays a woman less for the same output would result in a coefficient β greater than 1.

A final possibility for measuring labor quality (also discussed in Fox and Smeets [2011]) is to take an ability-perspective: an AKM-decomposition of worker's wages yields estimates of worker fixed effects and firm fixed effects. The worker fixed effects measure the ability of workers as rewarded by firms but take out everything that is constant within a firm over time (including firm productivity). So long as innovations in firm productivity are not predictive of worker moves (which is not an innocuous assumption), worker effects capture a worker's underlying ability regardless of the firm employing him. Another benefit of using person effects as a stand-in for labor quality is that if women prefer front-loaded pay over their lifecycle, while men prefer back-loaded pay, using the current wage bill or efficiency units regressions from male wages will bias estimates of the productivity gap relative to the pay gap. A life-cycle measure of worker ability gives an average measure of worker productivity. The form of the AKM decomposition used in this paper is

$$\ln e_{it} = \alpha_i + \theta_j + \sum_{t=2000}^{2010} \phi_t Year_t + \Phi_1 HS_{it} + \Phi_2 Col_{it} + \Phi_3 BA_{it} + \epsilon_{it} \tag{3}$$

so that firm J's labor in period t is given by $\mathcal{L}_{J(t)} = \beta \sum_{i \in J(t),F} \hat{\alpha}_i + \sum_{i \in J(t),M} \hat{\alpha}_i$

4.2 Production function

In the baseline, I model firm value added (revenue minus the cost of intermediate inputs) as a translog function of labor and capital:

$$\log(Y_{jt} - M_{jt}) = a_{jt} + \theta_i$$
$$+ \sum_{i \in I} \mathbf{1}\{j \in i\} \cdot \left(\alpha_{1,i}\log(\mathcal{L}_{jt}) + \alpha_{2,i}\log(K_{jt}) + \alpha_{3,i}\log(\mathcal{L}_{jt})^2 + \alpha_{4,i}\log(\mathcal{L}_{jt})\log(K_{jt}) + \alpha_{5,i}\log(K_{jt})^2\right) \tag{4}$$

where $\mathcal{L}_{jt}$ is a measure of the firm's labor force which is the sum of male and female efficiency units as described above, K_{jt} is firm j's value of capital stock, and a_{jt} is the log of revenue TFP, excluding industry fixed effects. I allow $\alpha_1, ..., \alpha_5$ to vary by industry and include industry fixed effects (θ_i) at the NACE 2-digit level.

Profit maximizing firms which take wages as given will set the ratio of the price of labor equal to the ratio of marginal revenue product. In this case, a marginal unit of male or female labor has the same effect on revenue, up to a constant β, so that

$$\frac{w_{jt}^f}{w_{jt}^m} = \beta$$

where w_{jt}^f is the average cost to firm j of hiring an additional unit of female labor at time t and w_{jt}^m is the cost of hiring an additional unit of male labor, controlling for observable differences in the quality of labor which enter the efficiency units calculation (1).

One assumption in the estimation of the relative productivity of female labor, β, is that male and female labor are perfect substitutes. For legal or social reasons, firms may prefer to hire men and women in constant ratios. I relax the assumption of perfect substitutes and estimate at CES aggregation of male and female labor: (4) using

$$\mathcal{L} = (\beta(L^F)^{\frac{\rho-1}{\rho}} + (L^M)^{\frac{\rho-1}{\rho}})^{\frac{\rho}{\rho-1}}$$

where ρ is the elasticity of substitution for labor by gender. As $\rho \longrightarrow \infty$, men and women become perfect substitutes.

4.3 Selection

A long literature discusses the many problems econometricians have faced when estimating parameters of production functions. As noted by Marschak and Andrews [1944], if labor and capital choices were exogenously assigned, rather than chosen by firms based on productivity, then we could simply estimate (4) assuming a_{jt} is a shock process orthogonal to observed labor and capital. However, any unobserved component of TFP which is known to the firm (such as a firm fixed effect) will affect the optimal choice of labor and capital. This biases estimates of the coefficients on labor and capital $\alpha_1, ..., \alpha_5$. The purpose of this paper is not to estimate labor and capital shares in Denmark, but rather to estimate the relative marginal product of men compared with women. For this purpose, endogeneity of input choice is not necessarily a problem. If firms hire a man or woman randomly, then β will not be correlated with productivity (or firm size). In some industries, this may be a reasonable approximation of hiring practices. Overall, however, it will be important to deal with the endogeneity of hiring choices. I borrow the classic Olley and Pakes [1996] control function approach to dealing with this endogeneity.

If some portion of A_{jt} is known to firms at the time they make their labor decisions, the labor share coefficients will be biased. If TFP is also correlated with the decision to hire a man relative to a women, this will bias estimates of β. This would be the case, for example, if a firm which anticipated a change in technology which made it more productive preferred to hire men, perhaps because it believed men were better able to work with new technology or because men were more interested in working with the new technology and only men applied for the new jobs. In both cases, if we can can control for the unobservable known to the firm at the time they make hiring decisions, then we can control for the role of sorting by gender in the estimation of β.

Following Olley and Pakes [1996], I use investment to control for unobservables known to the firm at the time they choose $\mathcal{L}$. The intuition for this control is straightforward: assuming investment has a monotonic relationship with the unobservable component of TFP known to the firm at the time they make their decisions (conditional on capital), then it will be possible to invert the optimal investment rule and use this inverted rule as a control for the unobserved TFP. I describe the assumptions in more detail in Appendix 2.

In this model, a_{jt} has a component which is a shock to the firm after they make labor and investment decisions, and also a known component (ω_t) which is unobservable to the econometrician directly. In other words, we can write $a_{jt} = \omega_{jt} + \varepsilon_{jt}$ where ω_{jt} is known by the firm and affects their optimal labor and investment decision. OP assume that ω_{jt} is a scalar which follows an exogenous first order Markov process—that the distribution $p(\omega_{t+1})$ depends only on the observed ω_{jt}. This assumption allows for simple firm fixed effects $p(\omega_{jt+1}|\omega_{jt}) = p(\omega_{jt+1}|\bar{\omega}_j)$, but is more general [Ackerberg et al., 2007]. Conditional on capital, investment is then increasing in the unobservable ω_{jt} so that we can invert the optimal investment rule and write $\omega_{jt} = \phi(i_{jt}, k_{jt})$. I use a 5th-degree polynomial in investment and capital to represent the inverted investment rule. In Appendix 2 I discuss the assumption necessary for validity of a polynomial in investment and capital as a control function in more detail. Key, of course, to this exercise is the monotonicity of the investment in unobservable productivity. When I estimate an OP version of my main specification, I do so by 2-digit industry since the monotonicity assumption is not plausible when comparing across broad industries.

A natural question is: do men and women select into firms with different productivities? In addition to checking for the role of selection using production function estimation, I check for selection using the method outlined in Card et al. [2016]. I estimate the average firm effects for men and women in the sample of connected firms in the data used in the production function estimation (relatively large, private sector firms). I discuss the details of this procedure in Appendix 2. I find that there is no difference in the average firm effects for men compared to women in this subsample.[10] This suggests that sorting is unlikely to bias estimates of the relative productivity of men compared to women. Confirming this, I find little evidence of systematic sorting within industries using OP. I discuss these results, as well as the main results, below.

5 Results

5.1 Estimates of the productivity gap

Table 4 presents my baseline estimates of the relative productivity of women compared to men via the translog production function in (4). I find that one unit of female labor is equal to about .92 units of male labor. The relative productivity of female labor is closer to .94 when using the 2008+ eIncome sample for estimation. This is driven by the better hours measures available

[10] This contrasts with the role of firm-level sorting in the overall Danish economy. Focusing on an establishment level definition of the firm and including both the private and public sector, Gallen et al. [2017] find that 2.7 percentage points of the log pay gap can be explained by sorting in Denmark in the 2000s.

in that data, rather than a broader definition of earnings (comparing column 2 to column 3 relative to column 3 and column 4). Columns including non-wage benefits are those which use the full definition of income available in eIncome, including payments to retirement accounts and the value of other non-monetary benefits, such as a computer, home, etc. Finally, detailed efficiency units allow for the interaction of occupation with industry and year in the calculation of efficiency units and use the finest level of education major choice. Using a more detailed definition of efficiency units, the productivity gap falls to just four percent, indicating very little difference between a unit of male labor and a unit of female labor. Note, however, when making efficiency units categories too fine, the interpretation of the productivity gap becomes more confounded with differences in returns to (for example) majors between men and women.

While the efficiency units method summarized in (1) is my preferred specification, I also report alternative specifications of efficiency units of labor. In 5, column (1), β is the coefficient on the female wage bill, where total labor is measured as the sum of the male wage bill and β times the female wage bill in the firm. The interpretation of β, then, is the productivity of a dollar spent on female labor relative to a dollar spent on male labor. I find that female labor is more productive per dollar, consistent with the evidence in Table 4 which shows a smaller productivity gap than pay gap. The estimates are three percent smaller when using the eIncome dataset (columns 2 and 3) and do not depend on whether non-wage benefits are included in the definition of the wage bill.

Column 4 of Table 5 is the measure of relative productivity using predicted individual fixed-effects from an AKM decomposition of wages as effective labor. Similar to the wage bill measure, the AKM measure of β can be interpreted as the relative productivity of a female unit of labor measured in average lifetime wages (rather than current period wages) compared to a male unit of labor. Discrimination in the sense of uncompensated productivity is largest using the AKM method. The benefit of the AKM method is that if men and women have different preferences for backloading pay over the lifecycle, then the efficiency wage regressions in (1) may be biased. Since the predicted individual fixed effects in this estimate of the production function does not vary with TFP, it does not have the same mechanical bias that is induced when using the wage bill.[11] Column 5 of Table 5 gives the Griliches estimate of the relative productivity of women as estimated in Hellerstein et al. [1999]. The gap is quite small (5 percent), given the coarse

[11]However, as noted before, Fox and Smeets [2011] argue that the wage bill measure works well for practical purposes, despite this bias.

categories used in estimation: 3 age bins, 4 occupation bins, married vs. single, and male vs. female. All regressions use a translog production function with 2-digit industry fixed effects and industry-specific coefficients on labor and capital.

To deal with the possibility of selection of women into less (or more) productive firms, I take two approaches. First, I decompose log worker earnings into individual and firm fixed effects, controlling for education, age, and year fixed effects. I use the method outlined in Card et al. [2016] to do this, but focus on the connected subset of my sample of private-sector firms. Separately for the male and female sample, I regress

$$\ln e_{it} = \alpha_i + \phi^{G(i)}_{j(i,t)} + X'_{it}\beta^{G(i)} + r_{it}$$

where $\ln e_{it}$ are the log earnings of worker i at date t, $j(i,t)$ is the id of the firm employing worker i at date t, $G(i)$ indicates gender, and X_{it} is a vector of controls which includes year dummies interacted with education dummies and quadratic and cubic terms in age interacted with education dummies. The difference between firm effects $E(\phi^F_{j(i,t)}|m)$ and $E(\phi^F_{j(i,t)}|f)$ summarizes the degree to which men and women work in firms with different average pay. I find only a 0.005 log point difference between these expressions, suggesting the sorting by gender across private sector firms is not a large factor in the gender pay gap. Nonetheless, I can also use the Olley-Pakes control function approach described in Section 4.3 to account for TFP-based sorting by gender. Including a 5th-order polynomial in investment and capital in equation (4) does not change my estimates, but this may be because the assumptions underlying the OP method are not valid when looking across wide industries.

Turning to a by-2-digit industry application of Olley-Pakes does result in a slightly larger role for selection in some sub-industries. Figure 1 gives a histogram of the difference between the productivity gap estimated using a translog production function at the 2-digit industry level[12], and a specification which adds to this production function the OP "control" for unobserved productivity which the firm uses when optimizing its factor choices. The vast majority of my productivity gap estimates do not change when adding the OP control. Interestingly, those that do change move both in the positive and negative direction, suggesting there is so positive sorting of women into more productive firms (this is especially true in manufacturing sub-industries).

[12] I also require that the sub-industry have at least 100 observations in the data

More generally, women tend to sort into less productive firms, but the magnitude of this is small on average and concentrated in a small number of industries.

Implicit in the production function estimated thus far is an assumption that male and female labor are perfect substitutes. This need not be the case. If firms hire male and female labor in some fixed proportion, then total labor is the CES aggregation of male and female labor in the firm (not the sum). Table 6 gives the coefficient on female labor under a CES specification for total labor. There is a substantial fall in the estimated productivity of women relative to men. Intuitively, the first order condition[13] (which is not used in estimation) would imply that β should fall when the fraction of female labor matters to firms: women are more scarce than men in the labor force, yet they are paid less. The wage bill estimate of discrimination also falls. The elasticity of substitution between male and female labor is between 5 and 10, depending on the specification. These fairly large estimates suggest that perfect substitutes is not an unreasonable assumption.

Next, I explore the source of the gap in productivity between men and women. The literature finds that the wage gap increases over a woman's life-cycle, markedly rising when she has children, and falling again only after mid-life (Kleven et al. [2015], Goldin [2014]). If mothers take more time off work to care for children (even in ways not measured by register data on hours worked) then we would expect this group to be driving up the productivity gap. If the productivity gap is instead driven by innate differences between men and women, some other factors correlated with gender, or mis-measurement, it would show up both for mothers and for non-mothers. I find that the productivity gap is driven only by mothers. Women without children are as productive as their male counterparts. I expand on this result in the next section.

5.2 Motherhood

Bertrand et al. [2010] find that in a sample of recent US MBA recipients, the gender gap in career disruptions and female preference for shorter work hours was driven largely by mothers. In Denmark, recent work by Kleven et al. [2015] has argued the much of the Danish wage gap occurs with motherhood. This has changed markedly over time. While the presence of children can explain 40% of the gender earnings gap in 1980, children can explain 80% of the gap in

[13]The FOC is $\beta\left(\frac{L^f_{jt}}{L^m_{jt}}\right)^{-\frac{1}{\rho}} = \frac{w^f_t}{w^m_t}$

2011. The "child-penalty" comes in the form of (roughly equally) lower labor force participation of mothers, fewer hours of work for mothers, and lower wage rates for mothers. In my sample, I consider only mothers who have selected into work and those who are working in industries with good output data, notably excluding the public sector. For these reasons, I find that motherhood explains less of the earnings gap—women with children are paid 85 cents on the dollar and women without children are paid 90 cents on the dollar compared to men without children. Nonetheless, mothers face the largest earnings gap. This paper is the first to study whether motherhood also affects the difference between earnings and productivity.

Wage gaps don't only differ across mothers and non-mothers, however. A literature started with Lundberg and Rose [2002] finds that fathers actually earn *higher* wages than non-fathers, controlling for many correlated factors. Using the PSID, Lundberg and Rose [2002] find a wage gap of 4.2 percent for fathers relative to men without children. Fathers also work more hours than men without children. Approximately the same relationship holds in Denmark for fathers compared with non-fathers. Women earn less as mothers and men earn more as fathers, both in Denmark and the US. This result would be implied by a model of household specialization with human capital accumulation (market and non market specific)—on average men invest in their careers to increase household market income and women invest in household production to increase household non-market output.

Register data makes it possible to incorporate whether or not a worker has a child into the estimates of relative productivity. In Figure 2 I plot the wage gap, measured using a wage regression of log earnings on a quadratic in experience, education dummies, industry, occupation, and hours fixed effects, as well as the interaction of parenthood, gender, and age categories in three year intervals. For each age category, I plot the pay of fathers, mothers, and women without children relative to men without children. As expected, the wage gap is largest for mothers and smallest (negative) for fathers in most age bins. Figure 3 plots the productivity gap for fathers, mothers, and women without children relative to men without children (the productivity analogue of Figure 2). As women move past childbearing age, both the residual productivity and pay of mothers and non-mothers converge[14]. However, when women are in their

[14]Note that there are strong cohort effects which I do not control for in wage regressions. Older mothers and non-mothers in the sample are from a different cohort than younger mothers and non-mothers. This is not a lifecycle analysis because the level of observation for productivity calculations is the firm, not an individual worker

prime childbearing years, mothers are substantially less productive than non-mothers, and non-mothers, fathers, and men without children have approximately the same level of productivity. Women without children are actually more productive than men without children. This is not true of wages: in all age brackets women's wages are lower than men's wages.

One channel through which motherhood may affect productivity is a flexibility penalty described in Goldin [2014]. Mothers require more flexible hours than non-mothers and that flexible job structure may be less productive, especially in traditionally inflexible occupations. Translating ISCO codes to the (to the extent possible) to the 92 professions in the ACS with flexibility measures from O*NET used in Goldin [2014], I categorize each 4 digit ISCO occupation into being either below median flexibility (for example, chief executives), being above median flexibility (for example, pharmacists), or missing flexibly data[15]. I then estimate the production function in (4) using

$$\mathcal{L}_{jt} = \hat{E}_{j,t,M,inflex} + \beta_{F,nf}\hat{E}_{j,t,F,inflex} + \beta_{M,f}\hat{E}_{j,t,M,flex} + \beta_{F,f}\hat{E}_{j,t,F,flex} + \beta_{M,o}\hat{E}_{j,t,M,o} + \beta_{F,f}\hat{E}_{j,t,F,o} \tag{5}$$

where

$$\hat{E}_{j,t,G,O} = \sum_{i \in j(t), G, occ \in \{O\}} \widehat{e_{it}^{M}}$$

for person i working in firm j at time t of gender G (M or F) in occupation occ which is in the set of 46 occupations categorized as Flexible ($flex$), 46 occupations categorized as Not-Flexible ($inflex$), or another occupation (o). The results of this exercise are in Appendix Table 13. To summarize, I find that indeed, conditional on working in jobs requiring long, inflexible hours, women are more productive than men (by nearly 40 percent). Conditional on working in hours-flexible jobs, women are less productive than men, also by about 40 percent. In other jobs, women are less productive than men but only by about 10 percent. These findings are consistent with inflexible hours-jobs being overall less desirable for women, so that conditional on working in those jobs, women are quite productive.

Ignoring selection, the age-decomposition by parenthood in Figure 3 suggests that women's productivity declines substantially when they have children. If wages are sticky down, then employers may have limited scope for adjusting wages in response to this 20% productivity decline. The probability of having a child at age 30 is 13.8% in Denmark. Fertility rates are

[15]Goldin selects occupations paying the highest wage to males

similarly high for all the prime child-bearing years. Suppose, for the purpose of this example, that the length of a wage contract is 4 years. Then employers would want to pay a 28 year old woman 7% less than a man because of risk of childbirth. In other words, taking into account childbearing probabilities, the expected productivity of a 28 year old non-mother over the next four years is seven percent less than her male counterpart's.

Selection may be an issue, of course. Many women leave the private sector when they have children in Denmark (Pertold-Gebicka et al. [2016]). Those who shifted from private to public sector within 2 years of child's birth had about 1/10th a standard deviation less education and made about 10k less in earnings than those who stayed in the private sector, controlling for age and year. This suggests that my estimates are a lower bound of the productivity gap between mothers and non-mothers, since mothers in the private sector are positively selected.

One possible test of whether employer expectations concerning future childbirth are driving the gap between pay and productivity of women without children is whether there is such a gap for women less likely to have children in the near future. When setting wages and making hiring decisions, employers may not observe (and legally may not condition wages on) whether a woman is married or single, but to the extent they do, employers may infer a married women is more likely to have children in the near future than a single woman. The Danish data contains information on whether a person is married or cohabiting with a partner, so I can perform a decomposition of the productivity gap based on marital status as well as gender. Using both the wage bill and the efficiency-units approach suggests that uncompensated productivity is largest for single women without children relative to single men without children and smallest for fathers relative to mothers. This is not consistent with a story in which marriage signals to employers that a woman without children is likely to have children very soon (detailed estimates are provided in Table 7). An alternative possibility is that women prefer relatively smooth wage profiles and choose contracts with smoother wages over their lifecycle than their expected productivity.

5.3 Estimates by industry and occupation

Estimates of the productivity gap vary widely across industries. Figure 4 plots the OP-estimated productivity gap by 2 digit industry, as well as the wage gap by industry. There is a very slightly positive correlation of 0.02 between the series. A regression of the wage gap on the productivity gap gives a coefficient of 0.115. The distribution of relative productivity is wide both within and across industries, with accommodations and food services as well as other services having generally smaller productivity gaps than manufacturing and construction.

At the 2-digit industry level, all standard errors are close to ten percentage points or more. Aggregating up to larger industry groupings and estimating the productivity gap by industry with 2-digit fixed effects, gives substantially more precise estimates of the productivity gap. I report these results (and associated wage gaps) Table 8. Interestingly, the relationship between the wage gap and the productivity gap for mothers is very strong in all industries other than construction: in real estate and renting, other services, wholesale and retail trade, and manufacturing, the wage gap (generally close to 20 percent) is within two percentage points of the productivity gap. In accommodations and food services the gap is closer to three percentage points, while in construction the gap is 7 percentage points.

For women without children compared to men, there is a positive relationship between the pay and productivity gap across industries (the higher the pay gap the higher the productivity gap, excluding construction). The difference between pay and productivity does vary somewhat across industries, ranging from just under twenty percent in real estate and renting, and just over 30 percent in accommodations and food services and construction. Across industries, the difference between pay and productivity of non-mothers is around ten times higher than that for non-mothers.

It is also possible to disaggregate β by occupation: Table 9 reports the results of production function estimation in (4) when labor is given by

$$\mathcal{L}_{jt} = \sum_{o \in \mathbf{O}} \beta^o_{F,c} \hat{E}^o_{j(t),F,c} + \beta^o_{F,nc} \hat{E}^o_{j(t),F,nc} + \beta^o_M \hat{E}^o_{j(t),M}$$

where $\hat{E}_{j(t),\cdot}$ are the sum of efficiency units in firm j estimated *excluding occupation fixed effects* in category $\cdot$, where these categories are female with children, female without children, and male, respectively. The set of occupations $\mathbf{O}$ is management at the highest level, job requires knowledge at the highest level (from school teachers to researchers), job requires knowledge at the medium level (e.g. tech.), office jobs, sales, service, and care jobs, craftsman jobs, blue collar jobs, military, and agriculture, forestry, fishing requiring basic level knowledge, along with large heterogenous categories unknown and other.[16]

There is not a strong relationship between the pay and productivity of either mothers or non-mothers across occupations. The relative productivity of women without children is everywhere higher than that of women with children, and generally higher than the productivity of men in the same occupation. However, women seem to be substantially less productive than men in low skilled jobs, without commensurately low pay. These are jobs in which union contracts have the largest influence on wages and retention, which may explain why women appear to be so dramatically over-compensated. Women without children are in every occupation paid less than than their relative productivity, while women with children are generally (but not everywhere) over paid.

6 Conclusion

This paper presented estimates of the relative productivity of men and women, accounting for age, education, experience, occupational choice, and hours worked. Overall, I find that the productivity of women is about 8 percent lower than men, controlling for age, education, experience, and hours worked. This implies that productivity differences explains just under two-thirds of the residual gender pay gap. This productivity difference may arise from differences in the effort, extra (undocumented) hours worked, or effectiveness of men relative to women.

While on average, the pay gap is quite close to the productivity gap, this is not true over all of the lifecycle. In particular, women without children are estimated to be as productive if not more productive than men without children, but they are paid less than these men. Mothers, on the other hand, are substantially less productive than fathers and are paid commensurate with this productivity gap. The data do not support marriage or cohabitation as a strong predictor of the divergence between pay and productivity, suggesting that employers do not backload wages in anticipation of future childbearing[17], instead workers may prefer smoother wage profiles than productivity profiles over the lifecycle.

The results reported above are generally robust to various different specifications of a firms quantity/quality of labor: the baseline estimate uses an efficiency units approach which predicts

[16]In estimation, β_M for the category "unknown" is restricted to be 1 (so all estimates are productivity relative to men in occupations unknown).

[17]assuming that living with a partner signals that children are more likely to arrive in the coming years

returns to various observables correlated with gender, another method is to use the wage bill to represent labor quality, another method uses person fixed effects from a wage decomposition, and a final method estimates the relative productivity of various observables directly as inputs in the production function. The results are also robust to using more detailed wage and hours measures (which exist only from 2008 onward). My baseline estimates are robust to controlling for the potential role of sorting by women into less productive firms. Finally, the general pattern that women without children are as productive as men, while mothers are substantially less productive holds across industries and occupations.

Like the wage approach, the productivity approach in this paper implies that motherhood is central to the discussion of the gender pay gap. However, while differences in pay between men and women are largest for mothers, I find that differences in pay which cannot be explained by productivity are largest for women without children. The factors driving the gap between the pay and productivity of women without children (preferences, discrimination, occupation sorting) is an interesting avenue for future research.

References

D. Ackerberg, K. Caves, and G. Fraser. Structural Identification of Production Functions. mimeo, UCLA, 2004.

D. Ackerberg, L. Benkard, S. Berry, and A. Pakes. Econometric tools for analyzing market outcomes. In J. Heckman and E. Leamer, editors, *The Handbook of Econometrics*, chapter 63. North-Holland, 2007.

Jerome Adda, Christian Dustmann, and Katrien Stevens. The Career Costs of Children. *IZA Discussion Paper*, (6201), 2011.

Dennis J. Aigner and Glen G. Cain. Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2):175–187, 1977. ISSN 00197939, 2162271X. URL `http://www.jstor.org/stable/2522871`.

James Albrecht, Anders Björklund, and Susan Vroman. Is there a glass ceiling in Sweden. *Journal of Labor Economics*, 21(1):145–177, 2003.

James Albrecht, Peter Skogman Thoursie, and Susan Vroman. Parental leave and the glass ceiling in Sweden. *IFAU working paper 2015:4*, 2015.

Joseph Altonji and Rebecca Blank. Race and gender in the labor market. In O. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics*, volume 3, chapter 48, pages 3143–3259. Elsevier, 1 edition, 1999.

Nikolay Angelov, Per Johansson, and Erica Lindahl. Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579, 2016. doi: 10.1086/684851. URL `https://doi.org/10.1086/684851`.

Ghazala Azmat and Rosa Ferrer. Gender gaps in performance: Evidence from young lawyers. *Journal of Political Economy*, 125(5):1306–1355, 2017.

Linda Babcock and Sara Laschever. *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press, 2003.

Gary Becker. *The Economics of Discrimination*. University of Chicago Press, 1971.

Gary S. Becker. Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, 3(1):S33–S58, 1985. ISSN 0734306X, 15375307. URL http://www.jstor.org/stable/2534997.

Lars Ivar Oppedal Berge, Armando J. Garcia Pires, Kjetil Bjorvatn, and Bertil Tungodden. Competitive in the lab, successful in the field? *Journal of Economic Behavior & Organization*, forthcoming, 2015.

Marianne Bertrand. *New Perspectives on Gender*, volume 4 of *Handbook of Labor Economics*, chapter 17, pages 1543–1590. Elsevier, 2011. URL https://ideas.repec.org/h/eee/labchp/5-17.html.

Marianne Bertrand, Claudia Goldin, and Lawrence F. Katz. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55, July 2010. doi: 10.1257/app.2.3.228. URL http://www.aeaweb.org/articles?id=10.1257/app.2.3.228.

Dan A. Black. Discrimination in an equilibrium search model. *Journal of Labor Economics*, 13(2):309–334, 1995. ISSN 0734306X, 15375307. URL http://www.jstor.org/stable/2535106.

Francine D. Blau. *Equal pay in the office.* Lexington Books, 1977.

Francine D. Blau and Lawrence M. Kahn. Understanding International Differences in the Gender Pay Gap. *Journal of Labor Economics*, 21(1):106–144, 2003.

Francine D. Blau and Lawrence M. Kahn. Female Labor Supply: Why Is the United States Falling Behind? *American Economic Review: Papers & Proceedings*, 103(3):251–256, 2013.

Francine D. Blau and Lawrence M. Kahn. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865, September 2017. URL https://ideas.repec.org/a/aea/jeclit/v55y2017i3p789-865.html.

Stephen Bond and Mans Soderbom. Adjustment Costs and the Identification of Cobb Douglas Production Functions. Economics Papers 2005-W04, Economics Group, Nuffield College, University of Oxford, January 2005.

Timothy N. Bond and Jee-Yeon K. Lehmann. Prejudice and Racial Matches in Employment. *MPRA Paper No. 67494*, 2015.

Audra J. Bowlus and Zvi Eckstein. Discrimination and Skill Differences in an Equilibrium Search Model. *International Economic Review*, 43(4):1309–1345, 2002.

Leonardo Bursztyn, Thomas Fujiwara, and Amanda Pallais. 'Acting Wife' Marriage Market Incentives and Labor Market Investments. *American Economic Review*, 107(11):3288–3319, 2017.

Thomas Buser, Muriel Niederle, and Hessel Oosterbeek. Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics*, 129(3), 2014.

David Card, Ana Rute Cardoso, and Patrick Kline. Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633–686, 2016. doi: 10.1093/qje/qjv038. URL +http://dx.doi.org/10.1093/qje/qjv038.

Rachel Croson and Uri Gneezy. Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–74, June 2009. doi: 10.1257/jel.47.2.448. URL http://www.aeaweb.org/articles?id=10.1257/jel.47.2.448.

Christian Dahl, Daniel le Maire, and Jakob Munch. Wage Dispersion and Decentralization of Wage Bargaining. *Journal of Labor Economics*, 31(3):501 – 533, 2013.

Benoit Dostie. Wages, productivity and aging. *De Economist*, 159(2):139–158, 2011.

Richard Ericson and Ariel Pakes. Markov-perfect industry dynamics: a framework for empirical work. *Review of Economic Studies*, 62 (January):53–82, 1995.

Jeffrey A. Flory, Andreas Leibbrandt, and John A. List. Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1):122–155, 2015. doi: 10.1093/restud/rdu030. URL +http://dx.doi.org/10.1093/restud/rdu030.

Jeremy T. Fox and Valerie Smeets. Does input quality drive measured differences in firm productivity?*. *International Economic Review*, 52(4):961–989, 2011. ISSN 1468-2354. doi: 10.1111/j.1468-2354.2011.00656.x. URL http://dx.doi.org/10.1111/j.1468-2354.2011.00656.x.

Roland G. Fryer and Steven D. Levitt. An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2):210–40, 2010.

Yana Gallen, Rune Lesner, and Rune Vejlin. The Labor Market Gender Gap in Denmark: Sorting Out the Past 30 Years. *IZA working paper no. 10789*, 2017.

Dora Gicheva. Working long hours and early career outcomes in the high-end labor market. *Journal of Labor Economics*, 31(4):785–824, 2013.

Uri Gneezy, Kenneth L. Leonard, and John A. List. Gender differences in competition: evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664, 2009.

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

Claudia Goldin and Cecilia Rouse. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4):715–741, September 2000. doi: 10.1257/aer.90.4.715. URL `http://www.aeaweb.org/articles?id=10.1257/aer.90.4.715`.

Zvi Griliches. Specification bias in estimates of production functions. *Journal of Farm Economics*, 39(1):8–20, 1957. ISSN 10711031. URL `http://www.jstor.org/stable/1233881`.

Morley Gunderson. Male-female wage differentials and policy responses. *Journal of Economic Literature*, 27(1):46–72, 1989. URL `http://EconPapers.repec.org/RePEc:aea:jeclit:v:27:y:1989:i:1:p:46-72`.

Judith K. Hellerstein and David Neumark. Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set. In Berndt and Hulten, editors, *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*. 2007.

Judith K. Hellerstein, David Neumark, and Kenneth R. Troske. Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production. *Journal of Labor Economics*, 17(3):409–446, 1999.

Judith K. Hellerstein, David Neumark, and Kenneth R. Troske. Market Forces and Sex Discrimination. *The Journal of Human Resources*, 37(2):353–380, 2002.

Joni Hersch and Leslie S. Stratton. Housework, fixed effects, and wages of married workers. *The Journal of Human Resources*, 32(2):285–307, 1997. ISSN 0022166X. URL `http://www.jstor.org/stable/146216`.

Fredrik Heyman, Helena Svaleryd, and Jonas Vlachos. Competition, takeovers, and gender discrimination. *ILR Review*, 66(2):409–432, 2013. doi: 10.1177/001979391306600205. URL `https://doi.org/10.1177/001979391306600205`.

Andrea Ichino and Enrico Moretti. Biological gender differences, absenteeism, and the earnings gap. *American Economic Journal: Applied Economics*, 1(1):183–218, January 2009.

Linda Kamas and Anne Preston. The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1):82–97, 2012.

Henrik Jacobsen Kleven, Camille Landais, and Jakob Egholt Søgaar. Parenthood and the Gender Gap: Evidence from Denmark. Working paper, 2015.

Andreas Leibbrandt and John A. List. Do women avoid salary negotiations? evidence from a large-scale natural field experiment. *Management Science*, 61(9):2016–2024, 2015. doi: 10.1287/mnsc.2014.1994. URL `https://doi.org/10.1287/mnsc.2014.1994`.

Shelly Lundberg and Elaina Rose. The Effect of Sons and Daughters on Men's Labor Supply and Wages. *The Review of Economics and Statistics*, 84(2):251–268, 2002.

Alan Manning and Helen Robinson. Something in the way she moves: a fresh look at an old gap. *Oxford Economic Papers*, 56(2):169–188, 2004.

Thomas Markussen, Ernesto Reuben, and Jean-Robert Tyran. Competition, Cooperation and Collective Choice. *Economic Journal*, 124:F163–F195, 2014.

Jacob Marschak and William Andrews. Random Simultaneous Equations and the Theory of Production. *Econometrica*, 12(3/4):143–205, 1944.

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

Muriel Niederle and Lise Vesterlund. Gender and Competition. *Annual Review of Economics*, 3:601–630, 2011.

OECD. *OECD Employment Outlook 2010—Moving Beyond the Job Crisis*. 2010.

G. Steven Olley and Ariel Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry . *Econometrica*, 64(6):1263–1297, 1996.

Jessica Pan. Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics*, 33(2):365–408, 2015.

Barbara Pertold-Gebicka, Filip Pertold, and Nabanita Datta Gupta. Employment Adjustments around Childbirth. *IZA working paper no. 9685*, 2016.

Ernesto Reuben, Paola Sapienza, and Luigi Zingales. Taste for Competition and the Gender Gap Among Young Business Professionals. *NBER Working Paper No. 21695*, 2015.

Jonah E. Rockoff and Mariesa A. Herrmann. Does menstruation explain gender gaps in work absenteeism? *Journal of Human Resources*, Spring, 2012.

Asa Rosen. An Equilibrium Search-Matching Model of Discrimination. *European Economic Review*, 41(8):1589–1613, 1997.

Statistics-Denmark. Production and generation of income (10a3-grouping). www.dst.dk. Data retrieved Aug 6th, 2015.

Jenna Stearns. The Long-Run Effects of Wage Replacement and Job Protection: Evidence from Two Maternity Leave Reforms in Great Britain. mimeo, 2017.

Mallika Thomas. The impact of mandated maternity benefits on the gender differential in promotions: Examining the role of adverse selection. mimeo, 2015.

Melanie Wasserman. Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents. mimeo, 2017.

Andrea Weber and Christine Zulehner. Competition and Gender Prejudice: Are Discriminatory Employers Doomed to Fail? *Journal of the European Economics Association*, 12(2):492–521, 2014.

Y. Jane Zhang. Culture, Institutions, and the Gender Gap in Competitive Inclination: Evidence from the Communist Experiment in China. *SSRN: http://ssrn.com/abstract=2268874*, 2015.

Figures

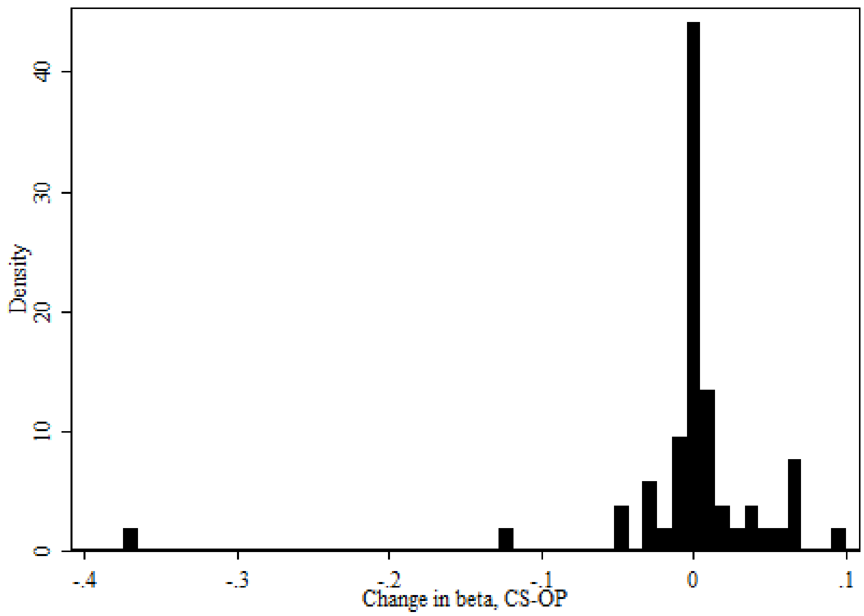

Figure 1: This figure shows the difference between estimates of the productivity gap (by 2-digit NACE industry) with and without the OP control function. Estimates without an OP control function are the coefficients on female labor efficiency units in a translog (in capital and the sum of male and female labor) production function which include detailed industry fixed effects. Estimates with an OP control add a fifth-order polynomial in investment and capital to this production function in order to approximate unobserved productivity known to the firm at the time it makes its labor and investment choices. The difference between the coefficients on female labor in these two production functions captures the role of selection of women into lower TFP-firms in explaining the gap between male and female productivity. I find that in most industries, there in no meaningful selection (captured by a large mass at 0). If there is selection, it is not always negative (though it is more often negative). There are 54 unique sub-industries with at least 100 observations in the data.

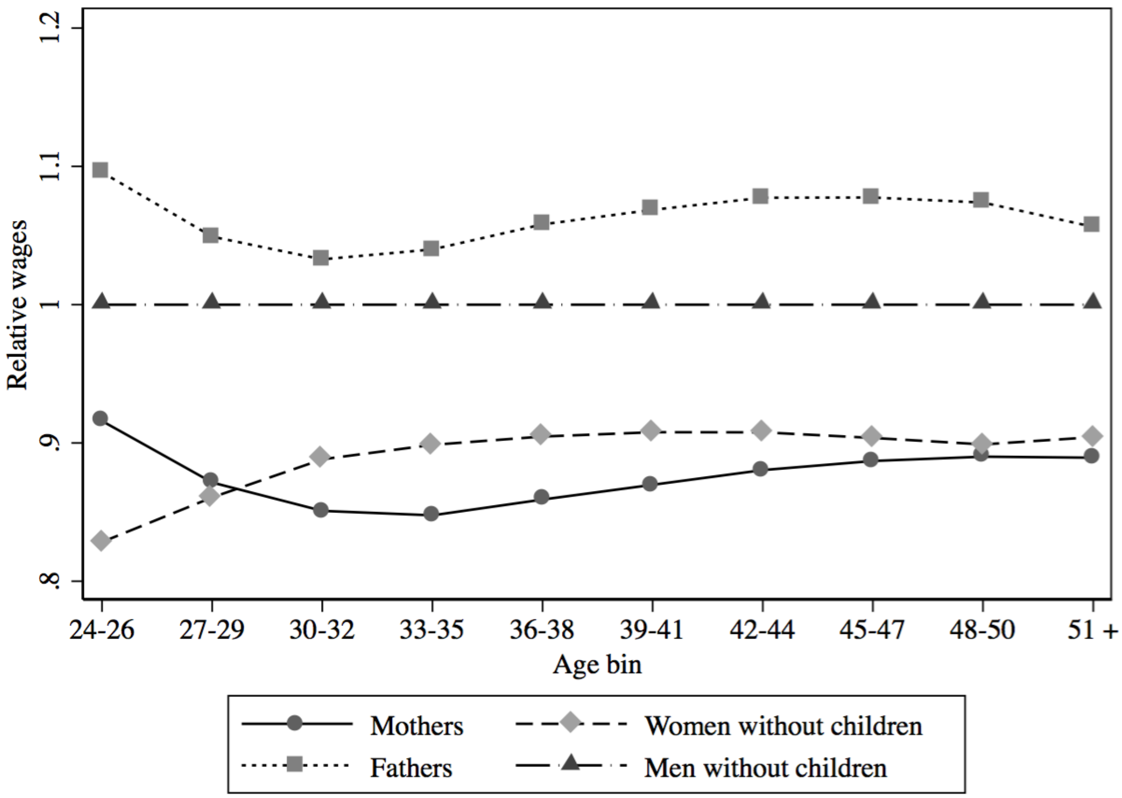

Figure 2: This figure shows the relative wages of women without children compared to men without children of the same age, as well as mothers and fathers compared to men without children of the same age. Relative wages are measured using a wage regression with 2 digit industry fixed effects, 3 education fixed effects, a quadratic in experience, and year fixed effects. For each age category, I normalize the wages relative to those of men without children of the same age.

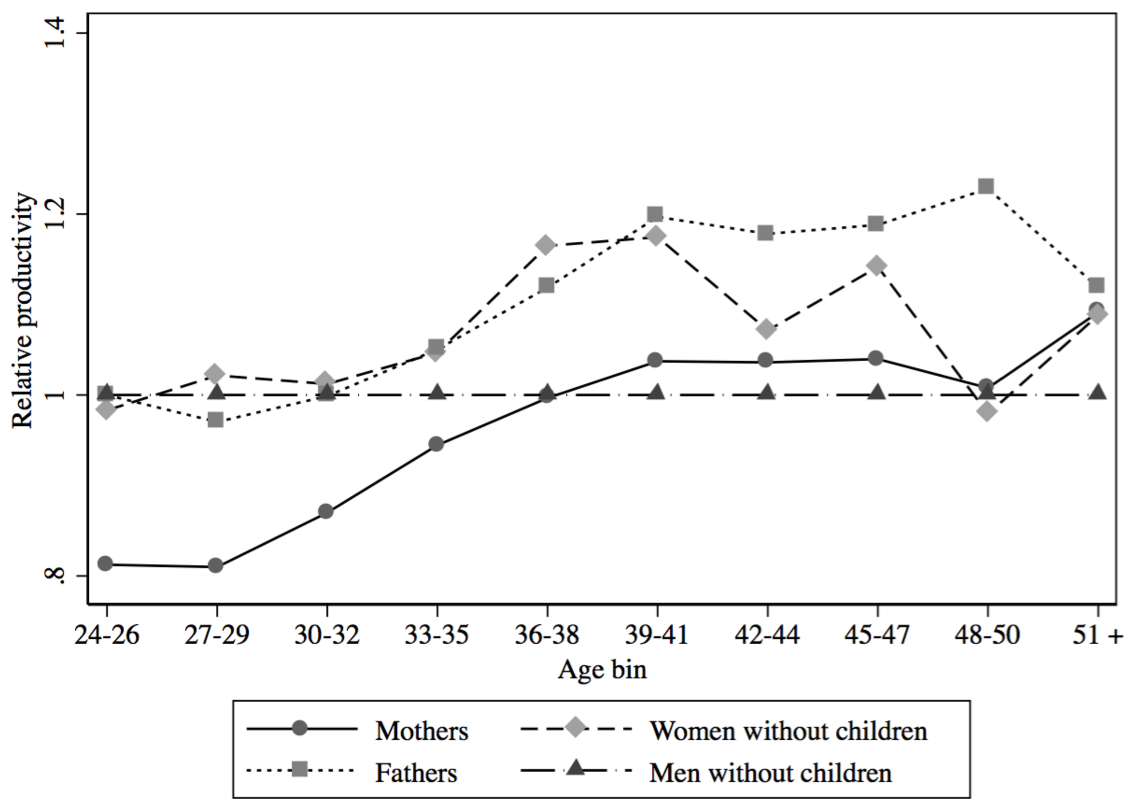

Figure 3: This figure shows the relative productivity of women without children compared to men without children of the same age, as well as mothers and fathers compared to men without children of the same age. Relative productivity is measured using the baseline translog production function with industry specific shares and fixed effects, and the baseline specification for efficiency units but omitting age from the efficiency units. I model efficiency units of the interaction of 12 age bins and 4 gender/parenthood categories as perfect substitutes.

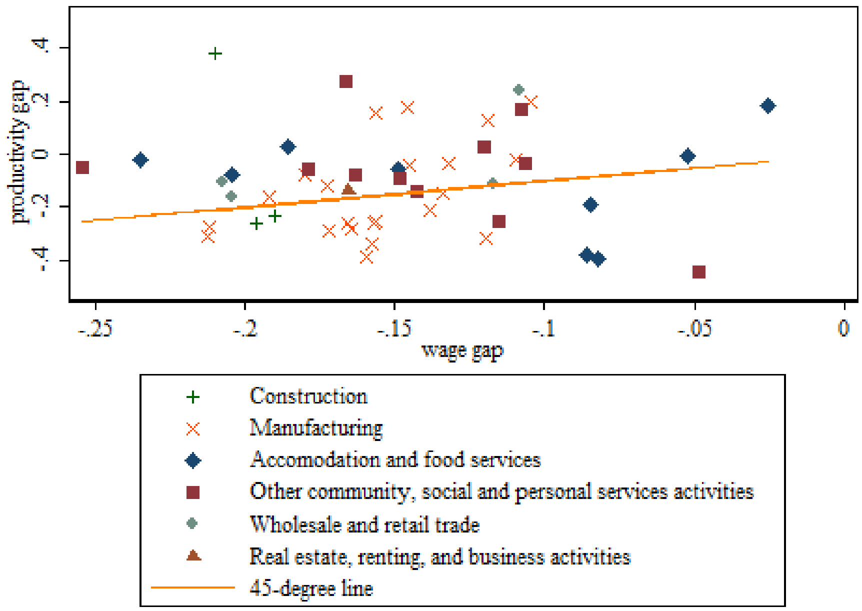

Figure 4: This figure displays a scatter of the wage gap and productivity gap, using an investment control function and translog production function, estimated by industry at the 2-digit NACE-level. Larger industry groups (1 digit level) are represented by different symbols. These are: Accommodations and food services, Construction, Manufacturing, Wholesale and Retail Trade, Other community, social, and personal services, and Real estate, renting, and business activities. The 45 degree line also graphed. The slope of a regression of the productivity gap on the wage gap is 0.11 and clearly, the correlation is very weak at the sub-industry level. Standard errors on the productivity gap when measured at the 2-digit level are large (and variable), ranging between 2 percentage points and thirty percentage points.

Tables

Table 1: The pay gap in Denmark vs. US

Sample	Variables included	Coefficient on female	Standard error	R^2
US	Basic	-0.320	0.0010	0.102
US	Basic, time	-0.196	0.0009	0.353
US	Basic, time, education	-0.245	0.0008	0.475
US	Basic, time, education, occupation	-0.191	0.0010	0.563
Denmark (FIRE)	Basic	-0.277	0.0011	0.095
Denmark (FIRE)	Basic, time	-0.193	0.0006	0.727
Denmark (FIRE)	Basic, time, education	-0.200	0.0006	0.750
Denmark (FIRE)	Basic, time, education, occupation	-0.172	0.0006	0.781

Dependent variable is log earnings. The sample is 2009 to 2011 All regressions include a quadratic in age and year dummies. US regressions also include race. Hours controls are added in the second regressions. Hours are bracketed in Denmark (see see the data appendix) and indicate hours per week and weeks per year in the US. Education indicates primary, high school, or more advanced schooling in Denmark, and similar groups in the US, and is added in the third row. Occupation dummies at the 3 digit level are added in the final row. Goldin's ACS sample includes only individuals ages 25-64. For comparison, I restrict to only these ages in the FIRE sample. The number of observation is 3,291,168 in the US, and 2,879,216 in the restricted FIRE sample.

Table 2: Conditional pay gap 2000-2011

	(1)	(2)	(3)	(4)
Female	-0.1740	-0.1628	-0.1393	-0.1184
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Experience	N	Y	Y	Y
Occupation, Industry FE	N	N	Y	Y
Firm× Occ FE	N	N	N	Y
R-squared	0.8430	0.8442	0.8562	0.8924
N	15613056	15613056	15613056	15613056

Dependent variable is log earnings. All regressions include hours and year controls, a quadratic in age, and education level dummies. Experience indicates a quadratic in experience (measured as hours of employment). Occupation is at the 3 digit ISCO level. Standard errors in parentheses.

Table 3: Cross-industry summary statistics

	Accom./food	Constr.	Manuf.	W/R trade	Other serv.	Real est.
w^f/w^m	0.8808	0.8083	0.8331	0.8418	0.8737	0.8485
fraction men	0.5911	0.8992	0.6919	0.6782	0.5152	0.5935
firm size	6.50	5.65	10.04	7.15	7.18	5.28
N	298370	135808	72000	69215	11799	18920

This table provides summary statistics on variables of interest across industries. w^f/w^m is the average wage gap control for quadratics in age and experience, education level, occupation fixed effects at the 3-digit ISCO level, hours worked, and year. Wage regressions are run by industry. The fraction men are averages measured at the person level. Firm size is measured treating the firm as the unit of observation.

Table 4: Estimates of β (relative female productivity)

	(1)	(2)	(3)	(4)	(5)
β	0.919	0.928	0.941	0.940	0.960
	(0.008)	(0.007)	(0.009)	(0.010)	(0.009)
w^f/w^m	[0.861]	[0.861]	[0.901]	[0.900]	[0.920]
industry-specific shares	N	Y	Y	Y	Y
better hours (2008+)	N	N	Y	Y	Y
including non-wage benefits	N	N	Y	N	Y
detailed efficiency units	N	N	N	N	Y
R^2	0.8489	0.851	0.853	0.8523	0.853
N	852,729	852,729	258,978	258,978	258,978

This table gives estimates of β, the coefficient on female efficiency units in the translog production function regression (4) using (1) to form efficiency units. β can be interpreted as the relative productivity of female labor, controlling for differences in the quality of that labor captured by age, experience, education, and occupation. All regressions include 2-digit industry fixed effects and columns 2-5 allow the coefficients on labor and capital to vary at the 1-digit industry level. Standard errors are bootstrapped (50 samples) at the person level to account for estimation error in forming predicted efficiency units and then, for each estimate of efficiency units, cluster bootstrapped at the firm-level in the production function estimation step. The last row of the table, w^f/w^m is relative female wages, residual of quadratics in age and experience, education level fixed effects, and occupation fixed effects. Columns (3)-(5) use a subset of the data (2008 onward) in estimation because more detailed measures of hours worked are available in that time-period. This does reduce the productivity gap by about two percentage points. Columns (3) and (5) also include non-wage benefits in the estimation of efficiency units. Column (5) estimates efficiency units allowing for the interaction of occupation with industry and year the finest level of education major choice.

Table 5: Alternative measures of β

	(1) Wage Bill	(2) Wage Bill (2008 +)	(3) Wage Bill (2008+ including non-wage benefits)	(4) AKM	(5) Griliches ($\beta = \phi_F$)
β	1.0624	1.0332	1.0350	1.0961	0.9519
	(0.0035)	(0.0055)	(0.0055)	(0.0101)	(0.0060)
R^2	0.8654	0.8724	0.8725	0.8184	0.6412
N	714,254	258,978	258,978	641,916	852,729

This table gives estimates of β, the coefficient on female efficiency units in the translog production function regression (4) using the wage bill (columns 1-3), person fixed effects from an AKM decomposition (column 4), and equation (2) (column 5) to measure effective labor. A coefficient larger than one on wage bill and AKM estimates is consistent with a productivity gap which is smaller than the wage gap, as in Table 4 above. The interpretation of these coefficients is that one dollar paid to female labor is more productive than one dollar paid to male labor. In contrast, column (5), though less than 1 is *also consistent with Table* 4. This is a measure of the relative productivity of a unit of female labor (not a dollar spent on female labor). All regressions use specification (4) to estimate the production function and include 2-digit industry fixed effects and allow the coefficients on labor and capital to vary at the 1-digit industry level. Standard errors are clustered at the firm-level.

Table 6: Imperfect substitutes: relative productivity of female labor (β) and the elasticity of substitution (ρ)

	(1)	(2)	(3)
β	0.868	0.872	1.029
	(0.008)	(0.008)	(0.004)
ρ	5.425	5.496	9.327
	(0.089)	(0.092)	(0.264)
N	852,729	852,729	714,254
industry specific shares	N	Y	Y
wage bill	N	N	Y

This table gives estimates of β, the coefficient on female efficiency units in the translog production function regression (4) allowing total labor to be a CES combination of male and female efficiency units, where ρ is the elasticity of substitution between male and female labor. All regressions include 2-digit industry fixed effects and columns 2 and 3 allow the coefficients on labor and capital to vary at the 1-digit industry level. Standard errors are bootstrapped (50 samples) at the person level to account for estimation error in forming predicted efficiency units and then, for each estimate of efficiency units cluster bootstrapped at the firm-level in the production function estimation step.

Table 7: Productivity-pay gap by gender, marital status, and children

	Productivity -pay gap	Wage bill (β)
female, no children, married or cohabiting	0.149	1.083
	(0.010)	(0.010)
female, no children, single	0.287	1.163
	(0.009)	(0.010)
mother	-0.006	0.947
	(0.005)	(0.006)
male, no children, single	0.115	1.051
	(0.007)	(0.007)
father	-0.126	0.897
	(0.001)	(0.005)
N	852729	714254

This table splits the productivity gap by parenthood, married/cohabiting status, and gender. The first column of estimates gives $\beta^* - 1 - \frac{w^*}{w^{m,nc,married}}$ where $*$ is the category described in the first column from regression (4). Married or cohabiting men are the baseline. The last column does a wage-bill version of this regression (so that a β greater than 1 indicates uncompensated productivity). All regressions include 2-digit industry fixed effects and allow the coefficients on labor and capital to vary at the 1-digit industry level. Standard errors are bootstrapped (50 samples) at the person level to account for estimation error in forming predicted efficiency units and then, for each estimate of efficiency units cluster bootstrapped at the firm-level in the production function estimation step.

Table 8: Gender productivity gap by industry and parenthood

	Accom./food	Constr.	Manuf.	W/R trade	Other serv.	Real est.
female, no children	1.231	1.127	1.043	1.135	1.160	1.075
	(0.015)	(0.042)	(0.038)	(0.031)	(0.074)	(0.082)
	[0.926]	[0.794]	[0.839]	[0.886]	[0.906]	[0.879]
female, children	0.852	0.881	0.846	0.813	0.829	0.834
	(0.009)	(0.021)	(0.020)	(0.017)	(0.047)	(0.045)
	[0.828]	[0.815]	[0.831]	[0.811]	[0.848]	[0.830]
N	298,370	135,808	72,000	69,215	11,799	18,920
R^2	0.770	0.819	0.848	0.704	0.600	0.7470

This table gives estimates of β, the coefficient on the category of efficiency units listed in the first column in a translog production function where the labor of men, women with children, and women without children are treated as perfect substitutes. Standard errors are in parentheses. Relative wages, residual of the same factors which enter efficiency units estimation, are in brackets. Both relative wages and relative productivity are compared to an omitted category of all men. Across industries, relative wages and relative productivity line up nearly perfectly for mothers, but are unrelated for non-mothers.

Table 9: Gender productivity gap by occupation and parenthood

	M	HS	MS	WC	S	C	LS
$\beta^o_{F,nc}/\beta^o_M$	1.045	1.154	1.240	0.992	1.073	1.129	0.744
	(0.111)	(0.036)	(0.030)	(0.019)	(0.022)	(0.036)	(0.067)
	[0.803]	[0.888]	[0.834]	[0.909]	[0.935]	[0.884]	[0.877]
$\beta^o_{F,c}/\beta^o_M$	0.940	0.816	0.867	0.753	0.552	0.907	0.570
	(0.048)	(0.020)	(0.014)	(0.010)	(0.013)	(0.022)	(0.025)
	[0.752]	[0.858]	[0.791]	[0.865]	[0.827]	[0.874]	[0.870]
N individuals	509,790	1,138,811	2,007,428	1,560,069	1,581,954	2,433,454	1,520,541

This table gives the ratio of relative productivity coefficients by occupation. The first row of coefficients is the relative productivity of women without children relative to men and the second row of coefficients is the relative productivity of women with children compared to men in the same occupation (o). o is one of the 11 occupations described below modeled as perfect substitutes in a translog production function. Bootstrapped standard errors are in parentheses. Relative wages, residual of the same factors which enter efficiency units estimation, are in brackets.
Occupation codes: M = Management at the highest level, HS = Job req. knowledge at the highest level (from school teachers to researchers), MS = Job req. knowledge at the medium level (e.g. tech.), WC = Office jobs, S = Sales/Service/Care, C = Craftsman jobs, LS = Blue collar jobs. Occupation data is only available for a person's main job. About 30 thousand person-year observations list the primary job as *military* and another 30 thousand in the category *agriculture, forestry, fishing requiring basic level knowledge*, the estimates for these productivity gaps are in Appendix Table 14, along with large heterogenous categories *unknown* and *other*.

Model Appendix

Assumptions underlying the Olley-Pakes control function approach to the selection problem:

Assumption 1: Factor prices are constant across firms

The assumption that factor prices are constant across firms allows us to infer that firms which choose different levels of investment do so because they predict that their TFP will differ in the next period. If firms face different labor prices, particularly by gender, then β may still biased due to unobservables (factor prices). In Denmark this assumption is not particularly offensive, since wages are set in no small part by collective bargaining and generally are compressed relative to the US.

Assumption 2: Labor is a non-dynamic input

This assumption *would* be unreasonable in countries where it was difficult to re-adjust the labor force every year. Denmark, however, prides itself on a “Flexicurity” system. This is the combination of a very flexible labor market—it’s very easy to fire and hire workers in Denmark—combined with a secure safety net in the case of unemployment. In Denmark and the US, just over 25% of employees are new hires in each year, and about 25% separated from their employer in the same period. In Norway, these rates are closer to 17%. In Italy, they are about 15% [OECD, 2010]. See Appendix Figure 6 for a graph of cross-country separation and hiring data.

Figure 5: This figure is directly replicated from OECD data on relative workforce flexibility, OECD [2010] Figure 2.1, see Annex 3.A1. Country averages of reallocation rates expressed in percentage of total dependent employment and adjusted for industry composition. Austria: 2002-07; Belgium: 2000-07; Canada: 2000-06; the Czech Republic: 2001-07; Denmark: 2000-06; Finland: 2000-07; France: 2000-06; Germany: 2000-06; Greece: 2000-05; Hungary: 2000-05; Iceland: 2002-07; Ireland: 2000-05; Italy: 2000-06; the Netherlands: 2000-07; Norway: 2000-04; Poland: 2004-05; Portugal: 2000-06; the Slovak Republic: 2002-06; Slovenia: 2002-07; Spain: 2000-05; Sweden: 2000-06; Switzerland: 2000-07; Turkey: 2007; the United Kingdom: 2000-07; and the United States: 2000-06.

Assumption 3: Conditional on capital, investment is monotonically increasing in the unobservable ω_{jt}

These assumptions rule out, for example, adjustment costs which differ across firms within an industry. Scalar investment is given by $i_{jt} = i_t(\omega_{jt}, k_{jt})$. Pakes (1994, Theorem 27) shows that when $i > 0$, $i_t(\omega_t, k_t)$ is increasing in ω for every k, so that we can invert the investment rule and write $\omega_{jt} = \phi(i_{jt}, k_{jt})$.[18][19]

Approximating this investment rule with a flexible, higher-order polynomial in k and I yields the equation

$$\log(Y)_{jt} = a_t + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 k_{jt} + \phi(i_{jt}, k_{jt}) + \varepsilon_{jt} \tag{6}$$

where $\phi(i_{jt}, k_{jt})$ is a flexible 3rd degree polynomial in i and k. Since labor does not enter the ϕ polynomial, the labor share and β are identified simply by running this regression.

Ackerberg et al. [2004] (ACF) note that there is a simultaneity problem if investment and labor are truly chosen simultaneously—in this case labor demand can be written $\mathcal{L}(\omega, k)$, problematically. Indeed, if labor can be written as a flexible polynomial in i and k, then there is perfect collinearity between ϕ and inputs in $\mathcal{L}$, making estimated labor coefficients meaningless. ACF suggest a 2-step solution to this problem, as well as a timing assumption which corrects the problem. In the Danish context and with yearly data, this timing is not particularly offensive. More formally:

Assumption 4: Labor is chosen first, then investment is chosen based on an information set correlated but not collinear with the information used to choose labor.

As suggested by Ackerberg et al. [2004] to eliminate the problem posed if i and labor are chosen based on exactly the same information set and factor prices do not vary across firms.[20] In general, all that is needed for identification is different adjustment speeds of various factors (see, Bond and Soderbom [2005]).

To estimate capital share, ψ_2, we can use the knowledge of ψ_1 and β obtained in the first stage to write

$$\log(Y)_{jt} - \psi_1 \log(\mathcal{L})_{jt} = a_{jt} + \omega_{jt} + \varepsilon_{jt}$$

Since ω is a first order Markov process, we can decompose it into its expectation given information at time $t-1$, $g(\omega_{j,t-1})$ and a residual, ξ_{jt}. In addition, we estimate the combination of

[18] Ericson and Pakes [1995] discuss the conditions for this invertibility in equilibrium in more detail.

[19] The general formation also includes firm age as a state variable, but omitting age does not affect the invertibility in equilibrium and simplifies the problem, since the relationship between firm age and productivity is not of interest in this paper.

[20] See Ackerberg et al. [2007] for an extensive discussion of OP and alternatives.

capital effects in the first stage. Let the first stage coefficient on capital be κ_{jt}. We now have

$$\log(Y)_{jt} - \psi_1 \log(\mathcal{L})_{jt} = a_t + \psi_1 k_{jt} + g(\kappa_{j,t-1} - a_{t-1} - \psi_2 k_{j,t-1}) + \xi_{jt} + \varepsilon_{jt}$$

This paper is focused on the estimation of β, which is identified in the first stage in the case of firm entry and exit, measurement error in investment, and lumpy levels of investment [Ackerberg et al., 2007].

Data Appendix

Table 10: Hours and jobs per person-firm-year-month in eIncome

Entries per p-f-y-m	Percent of p-f-y-m observations	25th pctile hours/month	50th pctile hours/month	75th pctile hours/month
1	98.88%	91	157.61	160.33
2	0.98%	6	22	62
3	0.025%	5	17	51
4+	0.0015%	3	8	31

This table describes the distribution of number of separate entries an individual (p) in a given firm (f) in a given year (y) and month (m) has in the eIncome data. The eIncome register is formed from taking monthly payroll statements which include occupation, hours worked, and various compensation breakdowns (take-home pay, adding fringe benefits, adding retirement contributions, etc). Multiple worker observations within a firm in a month (p-f-y-m) may arise because a worker changes occupations/job types in a month or has multiple occupations in a given month in a given firm, or they may arise due to a break in the employment spell in a month (resumed in the same month). There are a total of 135430660 person-firm-year-month observations. 98.88% of these have only one record and virtually all the rest have only two records. These data are used to construct efficiency units when estimates use the 2008+ sample of firms.

Table 11: Efficiency units estimation

More than high school	0.145
	(0.0003)
College	0.434
	(0.0004)
Age	0.061
	(0.0001)
Age^2	-0.001
	(0.0000)
Experience	0.021
	(0.0001)
Experience^2	-0.000
	(0.0000)

This table gives the estimated coefficients on education (omitted category is high school diploma or less), a quadratic in age, and a quadratic in experience. Regressions also include hours bins interacted with the fraction of the year worked, occupation fixed effects at the 3-digit ISCO level, and year fixed effects. Standard errors are in parenthesis.

Table 12: Griliches detail

$\phi_F - 1$ (female)	-0.0482
	(0.006)
$\phi_R - 1$ (married)	0.298
	(0.007)
$\phi_P - 1$ (35-54 year old)	1.038
	(0.015)
$\phi_O - 1$ (55 and older)	0.709
	(0.016)
$\phi_N - 1$ (unskilled laborers)	0.135
	(0.013)
$\phi_S - 1$ (white collar, technical, and sales workers)	0.497
	(0.009)
$\phi_C - 1$ (high skilled workers)	0.565
	(0.016)
α_1	1.192
	(0.011)
α_2	0.421
	(0.006)
α_3	-0.007
	(0.000)
α_4	-0.089
	(0.001)
α_5	0.054
	(0.002)

This table provides details of estimates using specification (2) for constructing $\mathcal{L}$ in the production function (4).

Table 13: Job-flexibility and relative productivity

$\beta_{F,nf}$	1.387
	(0.065)
$\beta_{M,f}$	1.554
	(0.046)
$\beta_{F,f}$	1.166
	(0.052)
$\beta_{M,o}$	1.338
	(0.034)
$\beta_{F,o}$	1.229
	(0.032)

This table provides details of estimates by occupation flexibility, as in equation (5).

Table 14: Gender productivity gap by occupation and parenthood

	Agriculture	Other	Military	Unknown occupation
$\beta^o_{F,nc}/\beta^o_M$	0.756	1.008	0.535	1.180
	(0.174)	(0.036)	(0.300)	(0.012)
$\beta^o_{F,c}/\beta^o_M$	0.519	0.561	-0.130	0.854
	(0.110)	(0.019)	(0.007)	(0.006)
N	37,326	1,720,432	30,207	4,194,407

This table gives the ratio of relative productivity coefficients by occupation for the three omitted occupation categories in Table 9. These are omitted because the sample size is small and/or the categories are not informative. The first row of coefficients is the relative productivity of women without children relative to men and the second row of coefficients is the relative productivity of women with children compared to men in the same occupation (o). o is the occupation listed in the column heading and modeled as perfect substitutes in a translog production function with the occupations in 9. Bootstrapped standard errors are in parentheses.