The Gender Productivity Gap*

(Job Market Paper)

Link to most current version: http://goo.gl/h3hNJi

Yana Gallen†

November 11, 2015

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 12 percent—seventy five percent of the 16 percent residual pay 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, and hours worked constant. To study the mechanisms behind the 4 percent gap in pay that is unexplained by productivity, I use data on parenthood and age. Mothers are paid much lower wages than men, but their estimated productivity gap completely explains their pay gap. In contrast, women without children are estimated to be as productive as men but they are not compensated at the same rate as men. The decoupling of pay and productivity for women without children happens during their prime-child bearing years. I provide estimates of the productivity gap in the cross-section and estimates that account for endogenous sorting of women into less productive firms using a control-function approach inspired by Olley-Pakes. This paper also provides estimates of the gender productivity gap across industries and occupations. Though the results do vary across industries and occupations, the overall estimate of the productivity gap is fairly robust to the specification of the production function.

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

---

*I thank Lori Beaman, David Berger, Stéphane Bonhomme, Matthias Doepke, Martin Eichenbaum, John Guryan, Trevor Gallen, Rune Vammen Lesner, Lee Lockwood, Casey Mulligan, Matthew Notowidigdo, Seema Jayachandran, Rune Vejlin and participants at the Northwestern Applied Micro Reading Group and Applied Micro Seminar, participants at the Nordic Data Workshop in Oslo and participants in the Conference on Labour Market Models and their Implications 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: Department of Economics, 2001 Sheridan Rd, Northwestern University, Evanston, IL 60208. Tel.: (815) 558-8693. Email: yanagallen2015@u.northwestern.edu. Web: sites.northwestern.edu/ype202

1 Introduction

There is a significant gender earnings gap in Denmark which persists when comparing men and women in the same occupation at the same firm. Many economists have run regressions to decompose the wage gap into whatever can be explained by observable differences between men and women and a residual. The residual is often attributed to discrimination [Altonji and Blank, 1999]. 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. Studying private-sector workers in Denmark, I find that about 12 percentage points of the 16 percent 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 laborer is female, holding other explanatory covariates such as age, education, experience, and hours worked constant. I use Danish data which matches worker characteristics with accounting information of their firm in order to estimate the gender productivity gap. A productivity gap of 12 percent has important policy implications—at the very least, it suggests that if discrimination exists in the labor market, it is not in the form of unequal pay for equal work. Nonetheless, productivity differences have not been thoroughly 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.

In the Danish registers, firm accounting data can be linked to not only the gender, age, education, and wage of workers, but it can also be linked to detailed information about the worker's children. I use this data on parenthood to better understand the source of the remaining 4 percent gap in pay that is unexplained by productivity. I allow the relative productivity of women without children to differ from the productivity of mothers.

For mothers, I find that the earnings gap coincides with the productivity gap, suggesting that there is little or no discrimination (in the form of uncompensated output) against mothers. The pay of mothers reflects true differences in their productivity. This is consistent with the literature suggesting that the wage gap occurs only for women with children who work fewer

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

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 flexible work arrangements.

Though there is no evidence of discrimination against mothers, there is evidence of discrimination against women without children. While earnings gap is smaller for women without children (12%), the productivity gap is much smaller (2%). I find that the disparity between wages and productivity for non-mothers happens especially between ages 25-35—prime child-bearing ages. Women of prime child-bearing age who have no children are likely expected by employers to have children in the near future. When they have children, my estimates suggest that their productivity will fall. These estimates are consistent with a model in which employers face sticky wage contracts and offer lower wages to productive women in anticipation of motherhood. The literature which labels residual pay gaps "discrimination" would conclude that discrimination is largest against mothers. I find the opposite: discrimination is largest in the group with a smaller residual pay gap (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 inspired by 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 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 dramatically change the overall estimate of the relative productivity of men compared to women. Overall, a 4 percent difference between the productivity gap and the pay gap remains (without controlling for selection, the difference between pay an productivity gaps is 3 percentage points and the productivity gap is fourteen percent). The results suggest that discrimination in the form of uncompensated output has a role in explaining the pay gap between men and women, but only for women without children.

The paper proceeds as follows: Section 2 reviews the relevant literature. Section 3 describes the data used in estimation. Section 4 describes the earnings gap in Denmark and other institutional characteristics. Section 5 provides the model and estimating equations. Section 6 presents results and Section 7 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. Blau [1977] argues that occupational choice plays a very large role in the gender wage gap since intrafirm wage schedules for a given occupation are constrained by a sense of inter-office fairness. 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.

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

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. 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).

My study makes 2 main advances. First, I address selection into firms, which could generate bias the estimation of the productivity gap. Second, I examine the role of childbearing in explaining 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.

A problem with the cross-sectional evidence on the gender productivity gap used by Hellerstein et al. [1999] is that if women sort into firms based on productivity, the estimates of productivity of women relative to men will be biased and reflect this sorting rather than actual productivity differences between men and women. This sorting is interesting in and of itself. 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).

Card, Cardoso, and Kline (2015) (henceforth CCK) 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 CCK. While CCK 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. I attempt to 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 addition, I study the role of parenthood in worker's productivity and pay. CCK focuses on the role of gender in bargaining and sorting but does not address differences for mothers relative to women without children.

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] or if firms did not pass improvements in productivity on to female employees as much as male employees, as in CCK. 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]). 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. 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.[2] 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.[3] 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. As discussed by Fryer and Levitt [2010] differences between boys and girls in mathematical ability expand over time, also suggesting a role of culture. This line of research links to the gender wage gap largely through the mechanism of occupational choice. 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 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.

[2] For example, **?** discuss women's differential propensity to provide to public goods.

[3] This competitiveness factor has been studied extensively in recent years. See for example Buser et al. [2014], Markussen et al. [2014], Kamas and Preston [2012], Berge et al. [2015], Zhang [2015], and Reuben et al. [2015].

Weber and Zulehner [2014] and Hellerstein et al. [2002] 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. They find that discrimination is correlated with slower firm growth and shut-down. 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. In addition, firms can compensate women at a lower rate than men. 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 fallen away 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 come from two Danish administrative registers. The IDA is a register which contains, from 1980- 2011 the universe of all workers and their wages, earnings, number of children, education, age, gender, occupation, and employer. Information on worker characteristics is then merged with firm-based data (FIRE). The FIRE dataset contains, from 1995-2011, employer reports of revenue for selected employers with more than five employees, as well as detailed information on operating costs, book value of capital, cost of intermediate goods, and many other accounting measures. In order to estimate production functions, I use three basic ingredients from this dataset: 1. revenue, 2. book value of capital, 3. value of material goods used in production.

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. This is not the case with information on total revenue and the cost of capital (since this is a tax-based measure) but it is the case with some information on the cost of intermediate goods. 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 the full dataset. The results are robust to including only surveyed firms. About 9,000 firms are actually surveyed in each year.

I focus my analysis on the five 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, Other services, and Wholesale and retail trade. The total number of firms in my dataset is 39,515. Notably, the category "Other services" includes firms which provide cleaning services and economic consulting firms, so it is quite broad.

My measure of labor uses data from IDA, not FIRE, since firm-records do not have detailed information about workers' experience, age, and gender.[4] In the table below I summarize demographics and earnings of the subset of the IDA population which works in firms in the FIRE

[4]One problem with this is that the occupational classification available for all workers is only available for the primary job. If a worker has multiple jobs and works in different capacities at those jobs (i.e.: in one job he is a manager and at another he is a white-collar worker) this will introduce some noise into the estimation. I suspect this is not a big problem since most workers with multiple jobs are low-skilled and most workers have only one job

dataset. This excludes workers in the public and agricultural sector, for example. However, this covers about 50% of Danish workers after 1999[5]. The observable differences between men and women in the FIRE sample are not negligible. Table 1 provides summary statistics on the FIRE population.

Table 1: Worker characteristics

	Men	Women
Age	37.87	35.32
	(13.16)	(13.24)
Has any children	0.442	0.458
	(0.497)	(0.498)
Number children	0.813	0.815
	(1.052)	(1.018)
Experience (in yrs.)	15.26	11.18
	(11.05)	(9.82)
Higher education	0.162	0.158
	(0.368)	(0.364)
Proportion Managers	0.042	0.016
	(0.210)	(0.126)
Earnings (2008 DKK)	262,499	177,198
	(202,413)	(138,136)
Proportion part-time	0.223	0.366
	(0.417)	(0.482)

Standard deviation in parentheses.

While female and male workers in these industries are of approximately the same age, women have about four years less experience and slightly more children. Men are three times more likely to be managers, and men earn more in a given year. The raw earnings gap without any controls is about forty percent in these industries. In the next section, I discuss the earnings gap in detail over time and residual of observables such as hours worked, education, age, and experience.

[5]Before 1999, FIRE includes only manufacturing and construction

4 The earnings gap

This paper studies the difference in output firms can expect when hiring a man relative to a woman. Since output is measured at the yearly level, rather than comparing to hourly wage differences, I compare the output gap to yearly earnings differences (the earnings gap). Table 2 below displays the result to a regression of 2010 log wages and earnings on an indicator of whether a worker is female and a quadratic in age. The only restriction on the sample is non zero wages/earnings and high quality information on hours worked.

Table 2: Wage vs. earnings gap (2010)

	log(wage)	log(earnings)
Female	-0.1852	-0.2416
	(0.0005)	(0.0011)
Age	0.0766	0.2400
	(0.0001)	(0.0003)
Age^2	-0.0008	-0.0026
	(0.0000)	(0.0000)
R-squared	0.196	0.353
N	2044206	2183859

Dependent variable is log wage in the first column and log earnings in the second column. Regressions include a quadratic in age. Data restricted to workers with high-quality information on hours worked. Data from 2010 only. Standard errors in parentheses.

While the wage gap is a bit smaller than the earnings gap, the magnitude of the difference is small compared to the gap overall. The raw wage gap is about 18 percent, while the raw earnings gap is about 24 percent. This difference reflects the fact that women are more likely to work part time or part of the year (see Appendix figure A1). Blau and Kahn [2013] highlight the difference between female labor force participation in Europe compared to the US and argue that the higher LFP of women in Europe is partially due to family-friendly employment policies. However, this higher female LFP is driven largely by more part-time, low-paying work. Denmark has a much larger proportion of women working part-time than the US, but fewer than similar European countries.

The earnings gap in Denmark is surprisingly similar to the gap in the US. Table 3 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. Since I will later study productivity differences, I need to account for all employees in a firm, which includes workers younger than 25 and older than 64. The Denmark (FIRE) sample includes all workers with positive earnings in the 5 largest industries with accounting data. Including all ages of workers in wage regressions does little to the wage gap but does increase the ability of age to explain earnings differences, so that the R-squared in the Denmark (FIRE) sample is about 20 percentage points larger than the age-restricted Denmark sample. In the sample of workers I will study using firm output data (the selected FIRE sample), the raw gap is 29.9 percent, compared with 32 percent in the US. Controlling for age, hours worked, education, and occupation, the gap falls to 19.6 percent, compared with 19.1 percent in the US.

In the samples with comparable age groups (Denmark and US), the R-squared from wage regressions in the US is about ten 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). 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 4 below, I report the results of a regression of log earnings on hours,

Table 3: 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	Basic	-0.242	0.0006	0.097
Denmark	Basic, time	-0.198	0.0004	0.600
Denmark	Basic, time, education	-0.225	0.0004	0.628
Denmark	Basic, time, education, occupation	-0.214	0.0004	0.636
Denmark (FIRE)	Basic	-0.299	0.0010	0.523
Denmark (FIRE)	Basic, time	-0.190	0.0007	0.785
Denmark (FIRE)	Basic, time, education	-0.200	0.0006	0.796
Denmark (FIRE)	Basic, time, education, occupation	-0.196	0.0007	0.798

Dependent variable is log earnings. The sample is 2009 to 2011 All regressions include a quadratic in age and time dummies. US regressions also include race. Hours controls are added in the second regressions and 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 future comparison, I restrict to age 25-64 in the Denmark sample, but include all ages in the FIRE sample. The number of observation is 3,291,168 in the US, 7,617,221 in Denmark, and 2,879,216 in the restricted FIRE sample.

a quadratic in age, and sequentially add controls for 1. a quadratic in experience and education level dummies, 2. occupation fixed effects, and 3. the interaction of firm and occupation. I restrict to the FIRE 2009-2011 sample for ease of comparison with Table 3 and the US data.

Column 3 of this table is analogous to the estimates of the wage gap that I will provide for comparison to the productivity gap[6] 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 does narrow the earnings gap by about 15% (to 15 percent).

While the raw earnings gap has fallen over time, the residual earnings gap has slightly risen

[6] Though in the remainder of the text, I report women's earnings as a fraction of men's (w^f/w^m) using the full time-period 1995-2011 with year dummies. All future references to w^f will be average female yearly earnings and w^m will be average male yearly earnings. The ratio w^f/w^m will be $1+b$ where b is the coefficient on a female dummy in the wage regression analogous to column 3 of Table 4.

Table 4: Conditional wage gap

	(1)	(2)	(3)	(4)
Female	-0.2002	-0.1917	-0.1821	-0.1524
	(0.0006)	(0.0006)	(0.0006)	(0.0007)
Experience	N	Y	Y	Y
Occupation FE	N	N	Y	Y
Firm× Occ FE	N	N	N	Y
R-squared	0.796	0.797	0.805	0.859
N	2875113	2875113	2875113	2875113

Dependent variable is log earnings. All regressions include hours and year controls, a quadratic in age, and education level dummies as in Denmark (FIRE) sample row 3 of Table 2. Experience indicates a quadratic in experience. Occupation indicates management, high skilled, white collar, or low skilled. Standard errors in parentheses.

over time. The portion of the earnings gap that cannot be explained by observables such as age, experience, education, and occupation has grown since 1995. Figure 1 below plots the raw earnings gap in Denmark, and the residual gap in the FIRE sample over time. The residual earnings gap is measured using a regression analogous to column 3 of Table 4, year-by-year. The raw earnings gap is measured using a regression analogous to column 2 of Table 2, year-by-year.

In general, the earnings gap in Denmark is large, and much of it is explained by differences in hours worked by men and women. The remainder of the gap is difficult to explain. Adding occupation controls and even firm fixed effects, the earnings gap falls to about 15% in the FIRE sample.[7] Overall, the difference in earnings between men and women survives many controls. The purpose of this paper will not be to explain the earnings gap, but rather to understand whether women are being compensated for a lower marginal product relative to men, or if women output more (or less) than their earnings would imply.

To achieve this, I will estimate the substitutability between men and women controlling (using efficiency units) for education, hours, age and experience observables, and bin workers according to their occupation into a CES labor aggregator. The occupation categories of management, high-skilled, white collar, and low-skilled are broad and representative—most firms have

[7] In the overall Danish sample, it is possible to get the earnings gap to about 10% when adding firm fixed effects and industry fixed effects

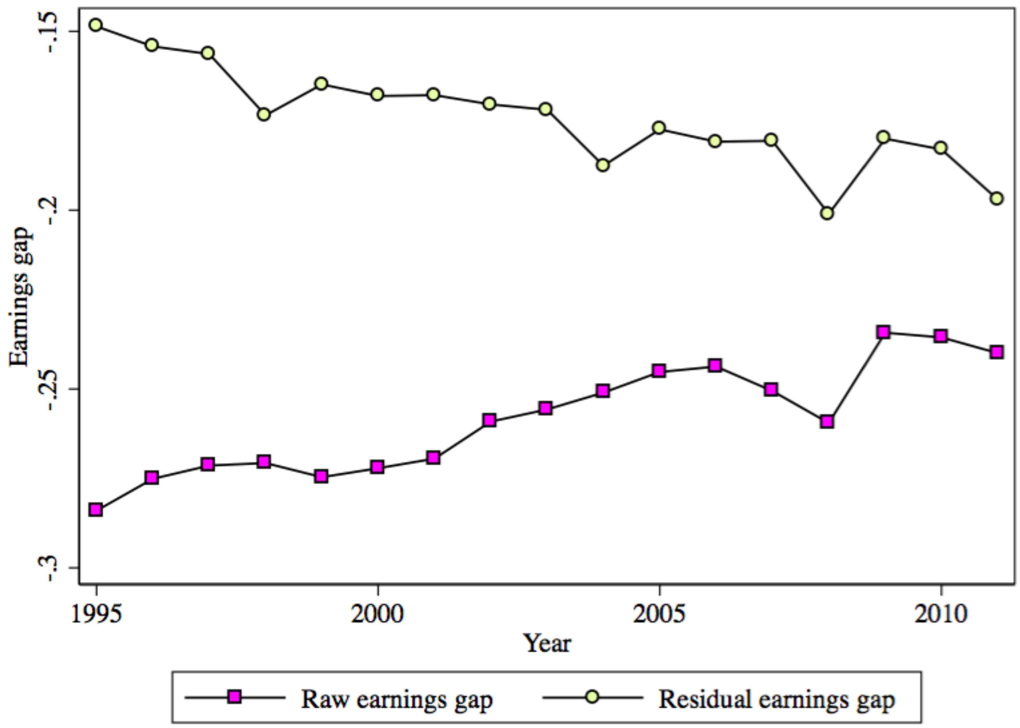

Figure 1: Figure 1 below plots the raw earnings gap in Denmark, and the residual gap in the FIRE sample over time. The residual earnings gap reports the earnings gap residual of hours dummies, a quadratic in age and experience, education level dummies, and occupation dummies.

workers in each category. Overall, the proportion of women in the FIRE sample has changed little over time, moving from 0.35 in 1999[8] to 0.33 in 2011 (falling only during the recession). Figure 2 below plots the proportion of women in each occupation over time. The proportion of female managers has risen steadily during this period, while the proportion of female low-skilled workers has fallen.

5 Model

In this section I present a model of firm-level value-added and its relation to the number of men and women at a firm, as well as the amount of capital purchased by the firm. [9] Value added in

[8]The sample is manufacturing only in 1995, and manufacturing, construction, and wholesale and retail trade only until 1999, slightly skewing the proportions female.

[9]Value added is revenue net of the cost of intermediate goods. The data appendix provides details of how these variables are measured in the data.

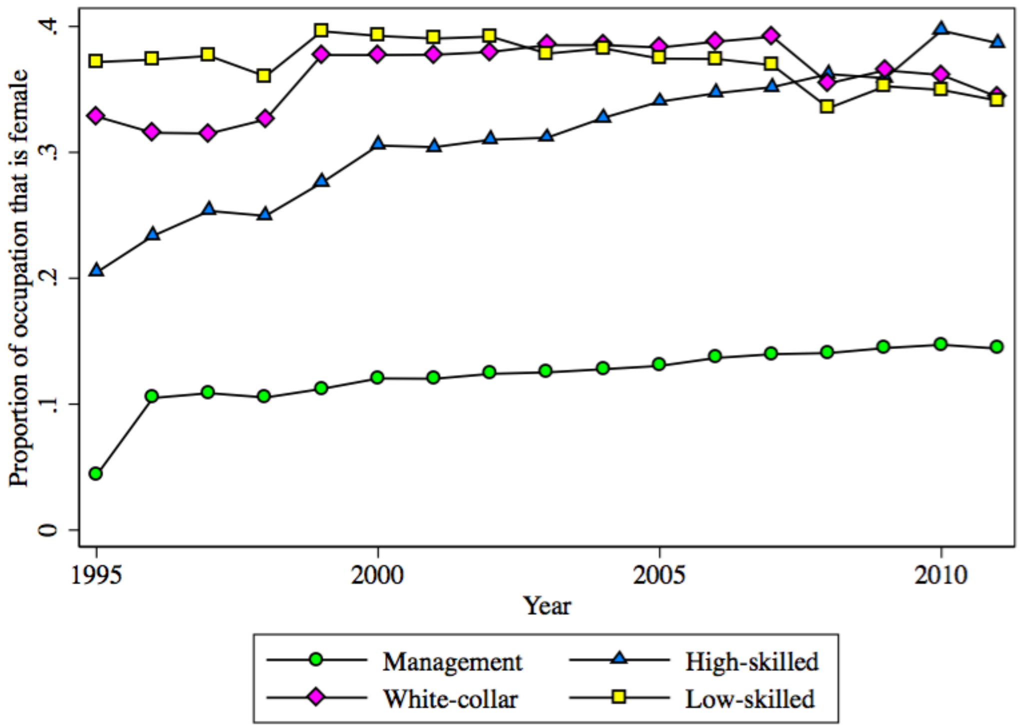

Figure 2: This figure displays the proportion of an occupation that is female over time. The occupation categories are those used in the model estimation and (though broad) are available for all workers in these firms.

firm j in year t is given by equation (1):

$$Y_{jt} = A_{jt} \mathcal{L}_{jt}^{\psi_1} K_{jt}^{\psi_2} \tag{1}$$

where

$$\mathcal{L}_{jt} = \left[\sum_{o=1}^{O} \alpha_o \left(\beta_o \left(L_{ojt}^{f} \right)^{\frac{\rho-1}{\rho}} + \left(L_{ojt}^{m} \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\sigma-1}{\sigma} \frac{\rho}{\rho-1}} \right]^{\frac{\sigma}{\sigma-1}}$$

Here, Y_{jt} is value added (revenue minus the cost of intermediate goods), A_{jt} is productivity, $\mathcal{L}_{jt}$ is the quantity of labor in the firm and K_{jt} measures its stock of capital. Total labor in a firm is a CES aggregation of labor in a variety of occupations (managers, office workers, low-skilled workers, and high-skilled workers, in this paper). Labor comes in two types: male labor (L^m) and female labor (L^f). 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 women are more productive than men. Estimating this parameter, β, is the focus of this paper.

The firm takes wages and rental rates as given and set by the market, then hires labor and invests capital to maximize output. The costs of the firm are given by

$$C_{jt} = \left[\sum_{o=1}^{O} w_{ot}^{f} L_{ojt}^{f} + w_{ot}^{m} L_{ojt}^{m}\right] + r_t K_{jt}$$

because firms take wages as given, the optimal amount of labor for a firm at time t will set marginal product equal to the wage rate. Taking the ratio of wages and the ratio of marginal products for women relative to men, we obtain

$$\frac{VMPL_{ojt}^{f}}{VMPL_{ojt}^{m}} = \beta_o \left(\frac{L_{ojt}^{f}}{L_{ojt}^{m}}\right)^{-\frac{1}{\rho}} = \frac{w_{ot}^{f}}{w_{ot}^{m}}$$

As $\rho \longrightarrow \infty$, $\beta_o \longrightarrow \frac{w_{ot}^{f}}{w_{ot}^{m}}$.[10] In this paper, I will indeed assume that $\rho = \infty$, or that men and women are perfect substitutes, conditional on occupation. The equality of relative factor productivities and their relative prices hinges on two important assumptions: first, factor prices are taken as given by the firm, second, firms choose the relative number of men and women in the firm optimally given this production function. This second assumption implies that in order to test $\beta^o = \frac{w_{ot}^{f}}{w_{ot}^{m}}$, we need to correctly account for the number of men and women at a given firm, especially when there is heterogeneity in the age, experience, and part-time/full-time mix of labor across firms. To do this, I count labor in efficiency units rather than "bodies."

L_{ojt}^{f} is the sum of efficiency-weighted units of labor supplied by women in occupation o at firm j at time t, and L_{ojt}^{m} is the sum of efficiency-weighted units of labor supplied by men in occupation o at firm j at time t. Women are on average more likely to be part-time and have much less experience than men in this sample. Not using efficiency units gives estimates of the output gap closer to -50%, but this, again, is driven by differences in hours worked by men and women, and also the lower experience of women compared to men. To convert personnel roles

[10] Assuming perfect substitutes between men and women, we can also write the log wage ratio in terms of possibly time-varying share variables and the log labor ratio:

$$\log\left(\frac{w_{ot}^{f}}{w_{o't}^{f}}\right) = \log\left(\frac{\alpha_1}{\alpha_2}\frac{\beta^o}{\beta^{o'}}\right) - \frac{1}{\sigma}\log\left(\frac{\beta^o L_{ot}^{f} + L_{ot}^{m}}{\beta^{o'} L_{o't}^{f} + L_{o't}^{m}}\right)$$

for each firm j. This equation, or a time series analogue, is generally used to estimate the elasticity of substitution in the production function, since firm-level output data isn't available (or isn't of interest). Instead, I will estimate σ in the production function using non-linear least squares. This estimation is very imprecise and my parameter of interest (β) isn't sensitive to the value of σ, given σ is not too small. Table 15 in the Appendix shows how my estimates of interest vary when I fix σ at different levels.

to efficiency units, I regress log earnings for males on observable characteristics of age, eduction, hours, and experience.

$$\log(e_{it})^m = \gamma_0 + \gamma_1 X_{it} + u_{it}$$

X_{it} includes dummies for education (high, med, low), a quadratic in age and experience, and dummies for (bracketed) hours worked. I use the estimated coefficients to predict the earnings of all workers and divide by mean predicted earnings (by occupation) to form efficiency units of labor. This normalization is absorbed in factor shares and doesn't affect estimation but makes "counts" of efficiency units closer to body-counts, rather than earnings counts. Since the returns to education, age, experience, and hours worked don't vary greatly by gender, the results are unchanged and robust to estimating efficiency units using women's wages rather than men's.

This paper is focused not on estimating the production function, but on estimating β, the relative productivity of women compared to men. Also of interest is testing the expected relationship between wages and productivity: $\beta = \frac{w^f}{w^m}$. β is estimated from data on firm-level value added, using a nonlinear least squares regression of log value added on log capital and log effective labor, where effective labor is a CES combination of labor from different occupations and men and women are perfect substitutes within an occupation.

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 (1) assuming $\log A_{jt}$ 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 labor share ψ. 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 make two different assumptions about A_{jt}, which I discuss below.

5.1 Cross-sectional

TFP shocks are unknown to firms at the time they make their labor decisions and are uncorrelated over time within firms. Treating shocks to firm productivity as random and unknown by the firm, I estimate a log-version of the Cobb-Douglas specification in equation 1:

$$\log(Y)_{jt} = a_{jt} + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 \log(K)_{jt} \tag{2}$$

I also estimate a translog version of this model of value added with includes also second order terms of log capital and log labor, following Hellerstein et al. [1999]. In particular, I estimate

$$\log(Y)_{jt} = a_{jt} + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 \log(K)_{jt} + \psi_3 (\log(\mathcal{L})_{jt})^2 + \psi_4 \log(K)^2_{jt} + \psi_5 \log(K) \log(\mathcal{L})_{jt} \tag{3}$$

This specification allows for a flexible relationship between labor and capital, but it does not account for the fact that firms may time when they hire a man vs. a women in a way that's correlated with output. To estimate the relationship between β and value added, I use the basic methodology suggested by Olley and Pakes [1996], discussed in the subsection below.

This specification won't allow consistent estimation of labor and capital shares when the firm knows some portion of its A_{jt} shock and uses it when choosing its labor force and how much to invest. However, if β is constant across firms and firms hire men and women randomly, β will still be consistently estimated. If this is not the case, then β will be some average of the actual difference in productivity between men and women if they were randomly hired by firms and the relative productivity of firms with a lot of women (for endogenous reasons) compared to those with a lot of men.

5.2 Endogenous labor composition

If some portion of A_{jt} is known to firms at the time they make their labor decisions, the labor share coefficient will be biased in the regression above. If TFP is also correlated with decisions firms make to hire men relative to 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 they believed men were better able to work with new technology. An alternative would be that men are more interested in working at firms adopting new technology and more men apply for new job openings than women. In both cases, if we 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 below:

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]. Next, two important assumptions are made:

Assumption 1: Factor prices and the depreciation rate 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. The assumption that the depreciation rate is constant across firms allows me to use $k_{t+1} - (1-\delta)k_t$ to represent investment.[11]

Assumption 2: Labor is a non-dynamic input

This assumption *would* be unreasonable in countries where it was difficult to re-adjust the

[11] The polynomial in investment and capital effectively becomes a polynomial in k_{t+1} and k_t. It's not important to correctly estimate δ, but it is important that it not vary across firms.

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 8 for a graph of cross-country separation and hiring data.

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 $\omega_{jt} = \phi(i_{jt}, k_{jt})$.[12][13]

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

$$\log(Y)_{jt} = a_t + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 k_{jt} + \phi(i_{jt}, k_{jt}) + \varepsilon_{jt} \tag{4}$$

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

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

[13] 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.

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.[14]

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 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]. Nonetheless, I restrict the dataset to firms with strictly positive investment, dropping 25% of the sample in order to give estimates from a sample analogous to what is used in the broad literature. This restriction is not necessary for consistent first stage estimates of β and it does not change my first stage results.

6 Results

I focus my analysis on the five 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, Other services, and Wholesale and retail trade. These make up 47 percent of the Danish economy[15]. Table 5 below provides some summary statistics for the firms in each industry and the dataset overall.

Overall the wage gap is about 17 percentage points. This varies by industry markedly, ranging from 22% to 7%. The fraction of the workforce in a given industry which is male also varies. In construction and manufacturing, a very high proportion of the labor force is

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

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

Table 5: Cross-industry summary statistics

	All	Accom./food	Constr.	Manuf.	Other serv.	W/R trade
w^f/w^m	0.8287	0.9331	0.8126	0.8349	0.7850	0.8166
fraction men (eff. units)	0.6934	0.4938	0.8991	0.7146	0.6269	0.6110
fraction men (bodies)	0.6492	0.4464	0.9009	0.6895	0.5462	0.5638
firm size (mean)	22.8227	18.3828	17.8864	36.6928	25.7506	17.9488
firm size (median)	10	10	9	13	11	11
N	527482	39624	116266	128954	50924	191714

Wage regressions and fraction men are averages measured at the person level. Firm size (mean and median) is measured treating the firm as the unit of observation.

male, while in accommodations and food services, less than half of workers are male, even when measured in efficiency units (so accounting for differences in age, education, and hours worked between the genders). The average and median firms in this dataset are slightly larger than in the Danish economy because detailed accounting statistics are kept only for relatively large firms (and there are no detailed statistics for firms with fewer than five employees). 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 later in the section.

First, I present: 1. estimates of the productivity gap using cross-sectional production function estimation, 2. estimates of the productivity gap using an Olley-Pakes correction for endogeneity of inputs and TFP, 3. estimates of the productivity gap for mothers, fathers, and women without children relative to men without children, 4. estimates of the productivity gap by occupation, 5. estimates of the productivity gap increasing the number of occupations in the production function, 6. estimates of the productivity gap by industry.

6.1 Baseline estimates of the productivity gap

As discussed above, estimating β (the productivity of women relative to men) from cross-sectional variation alone is prone to omitted variable bias. Nonetheless, estimates of β using this variation will not markedly differ from estimates using an investment control function. When estimating β using pure cross-section variation, we obtain an estimate of the average ratio of a unit of female effective labor relative to a unit of male effective labor, assuming that women and men are sorted randomly across firms.

Table 6 below provides estimates of the cross-sectional estimate of β using three different assumptions about the production function: 1. Cobb-Douglas, 2. Translog, 3. Translog with industry specific occupation-weights in the CES labor aggregator. All estimates include industry fixed effects and year fixed effects. Column 4 additionally restricts the sample to those firms with positive investment data (the O-P sample). Also included in Table 6 are OP estimates of β, with and without industry specific occupational shares. The first column of Table 6 reports β from the regression

$$\log(Y)_{jt} = a_{ind} + a_t + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 \log(K)_{jt} + \varepsilon_{jt}$$

where $\log(Y)_{jt}$ is log value added, $\log(K)$ is the log value of fixed material inputs[16] and $\log(\mathcal{L})$ is the log of effective labor. Effective labor is a CES combination of labor in four main occupations: management, high-skilled, white-collar, and low-skilled. These quantities are measured in efficiency units, so a low-skilled laborer with a high school degree provides fewer efficiency units than a low-skilled laborer with a college degree since there is a college premium for low skilled workers in wages. Within an occupation category, men and women are perfect substitutes but may have different efficiencies. In particular,

$$\mathcal{L} = \left[\left(\beta L_m^f + L_m^m\right)^{\frac{\sigma-1}{\sigma}} + \alpha_{hs}\left(\beta L_{hs}^f + L_{hs}^m\right)^{\frac{\sigma-1}{\sigma}} + \alpha_{wc}\left(\beta L_{wc}^f + L_{wc}^m\right)^{\frac{\sigma-1}{\sigma}} + \alpha_{ls}\left(\beta L_{ls}^f + L_{ls}^m\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

where the occupation categories are m = management, hs = high-skilled, wc = white collar, and ls = low skilled. β is the coefficient of interest and measures the relative productivity of a women compared to a man of the same background (with the same number of efficiency units). The second column of Table 6 estimates

$$\log(Y)_{jt} = a_{ind} + a_t + \psi_1 \log(\mathcal{L})_{jt} + \psi_2 \log(K)_{jt} + \psi_3 \log(\mathcal{L})^2_{jt} + \psi_4 \log(K)^2_{jt} + \psi_5 \log(K)\log(\mathcal{L})_{jt} + \varepsilon_{jt}$$

where $\mathcal{L}$ is the same as the Cobb Douglas case, so the only difference is a more flexible relationship between value added, effective labor, and capital. In column 3 of Table 6, the α_o shares are allowed to vary by industry.

The overall estimate of β (the productivity gap of women relative to men) in the cross section

[16] see Data Appendix for a detailed description of the variables used to measure value added and the capital stock

is 0.878, while women's wages on average are 82.87% of men's[17]. This suggests that more than half the wage gap can be explained by differences in the relative productivity of men compared to women, in the cross-section. This is similar to findings of Azmat and Ferrer [2015] who study the productivity and pay differences of male and female lawyers, finding that about half the earnings gap is explained by observables and most of the rest of the gap is explained by productivity differences.

Although the average number of men relative to women in the labor force in Denmark overall is 52% [Bank], the industries which answer the firm revenue survey are disproportionally male dominated industries (in particular, the large public sector is mostly female and does not provide Statistics Denmark with revenue figures). Importantly for this exercise, the construction industry is 90% male and manufacturing is more than 70% male. In these industries, it is relatively more likely that the hiring of a woman is a non-random event. Indeed, in these mostly male industries women make up a larger fraction of the total workforce in larger firms. If conditional on workforce size and capital stock, more productive firms have more women then estimates of β will be based in the cross-section.

More generally, cross-sectional analysis treats TFP as a random draw for each firm in each year which is uncorrelated with the proportion of women in the firm. If it is not a random draw each year but is uncorrelated with the proportion of women in the firm then there is no estimation problem. However, as noted in Card et al. [2015], women sort into firms with relatively lower TFP[18]. To correct for TFP unknown to the econometrician but known to firm at the time they make hiring decisions, I follow the method implemented by Olley and Pakes [1996] in their study of the telecommunications industry. In order to control for both endogenous shut down and endogenous factor choices, the authors suggest a control function in investment and capital to proxy for unobservables known to the firm and used by the firm in deciding whether or not to shut down and whether or not to increase the size of their labor force.

The overall estimate of β, restricting β to be the same across occupations and controlling for firm-level unobservables using the Olley-Pakes method is 0.884. This is higher than β when estimated without controlling for selection, so women do seem to sort into less productive firms.

[17]This is when measured using individual data. When taking firm-level averages the pay gap is $-.16$ (about one percentage point smaller).

[18]To the extent that high wage firms are also high productivity firms (which is not obvious in a search model). Abowd et al. [1999] find that high wage firms are indeed high productivity and high profit firms.

Since β only changes by a few percentage points, selection is not a very important factor biasing the estimates of β. Since relative wages are also affected by sorting, the wage gap also reflects this endogeneity. A way to control for this sorting is to compare only men and women working in the same firm in the same year (i.e.: add firm $\times$ time fixed effects to the wage regressions). This gives a wage gap of -0.164, suggesting a small amount of sorting compared to the overall wage gap of -0.171. Notably, this β is an average across industries and occupations. Later in this section, I will provide estimates of β by occupation and across industry. Selection plays a larger role in these breakdowns.

Table 6: Estimates of β

	Cobb-Douglas	Translog	Translog, ind. specific shares	Translog, ind. specific shares O-P sample	O-P	O-P, ind. specific shares
β	0.930	0.909	0.878	0.859	0.894	0.884
	(0.0051)	(0.0039)	(0.0038)	(0.0046)	(0.0059)	(0.0056)
N	527482	527482	527482	269445	269445	269445

Detailed estimates of other parameters in the production function are provided in appendix table A3. Column 1 is the result of a non linear regression of value added on log effective labor units and log capital, with β, the coefficient measuring the substitutability between men and women in labor constrained to be the same across occupations. Column 2 and 3 estimate a higher order version of this regression. Column 3 in addition allows occupation-shares to vary by industry. Column 4 additionally restricts the sample to those firms with positive (non-lumpy) investment. Column 5 adds a control function in investment and capital. Column 6 additionally allows occupation shares to vary by industry.

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.

6.2 Mothers

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, Landais and Søgaard (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 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. 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 79 cents on the dollar and women without children are paid 86 cents on the dollar compared to men without children[19] . 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 this section, I report the results of estimation comparing wage gaps and productivity gaps of fathers, males without children, mothers, and women without children. These groups certainly differ on observables. Table 7 below compares the age, probability of being in management, and probability of working part-time for men and women with and without children in the home.

Women without children also work in different industries than women with children. This is not the case for men. Table 8 below plots the distribution of industries by gender and presence

[19]Controlling for firm×time fixed effects, these numbers are 81 cents and 88 cents, respectively.

Table 7: Summary statistics for parents and singles

	Age	Management	Part-time	N
Men, no children	36.89	0.036	0.266	7761449
	(15.33)	(0.186)	(0.442)	
Fathers	39.60	0.064	0.155	4335573
	(7.59)	(0.246)	(0.363)	
Women, no children	34.01	0.013	0.426	4185166
	(15.58)	(0.114)	(0.494)	
Mothers	37.64	0.022	0.272	2352839
	(6.87)	(0.146)	(0.445)	

Standard deviations in parentheses.

of dependents (children in the home).

Table 8: Industry distribution by presence of children and gender

	Accom./food	Constr.	Manuf.	Other serv.	W/R trade	Total
Men, no children	4.80	19.35	37.16	9.47	29.23	100
Fathers	2.12	19.72	43.29	9.90	24.98	100
Women, no children	11.03	3.54	28.41	14.04	42.99	100
Mothers	4.84	4.73	40.58	16.12	33.73	100
Total	5.58	14.04	37.05	11.44	31.90	100

This table gives the proportion of workers (classified by gender and presence of children in the home) across the 5 industries.

Women without children are twice as likely to work in accommodations and food services compared to women with children. They are about ten percentage points less likely to work in manufacturing and ten percentage points more likely to work in wholesale and retail trade.

To test whether mothers (rather than all women) have a different marginal product than men, effective units of labor now distinguish between mother, fathers, non-parent men, and non-parent women. Effective labor becomes

$$\mathcal{L} = \left[\sum_o \alpha_o^i \left(L_o^{m,nc} + \beta^{f,nc} L_o^{f,nc} + \beta^{f,c} L_o^{f,c} + \beta^{m,c} L_o^{m,c} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{5}$$

where m, ch indicates a male without children, m, c indicates a male with children, f, nc indicates a female without children, and f, c indicates a female with children.

Table 9 provides the results of the O-P nonlinear regression of value added on effective labor and capital, using the specification above for effective labor. This specification allows a separate relative productivity parameter for each category of worker (women without children, women

Table 9: Mothers, fathers, and non-parents

	β^i	$w^i/w^{\text{men, no children}}$	p-value $\beta^i = w^i/w^{\text{men, no children}}$
Women, no children	0.980*	0.880***	0.000
	(0.0091)	(0.0003)	
Mothers	0.785***	0.812***	0.004
	(0.0096)	(0.0004)	
Fathers	1.055***	1.045***	0.396
	(0.0105)	(0.0003)	
N	269445	18054253	

Column 1 of this table gives the productivity of mothers, women without children, and fathers relative to men without children, as estimated with effective labor as in (5). Column 2 gives the wages of mothers, women without children, and fathers relative to men without children, as estimated from wage regressions. Standard errors are in parentheses. *, **, and *** are 10, 5, and 1% significance levels, against a null of $\beta^i = 1$. The p-value from an F-test of whether the productivity gap $1 - \beta^i$ equals the wage gap $1 - w^i/w^{\text{men, no children}}$ is the third column.

with children, and men with children) relative to men without children. There is no evidence of discrimination for mothers or fathers—each is paid in line with their marginal productivity (in the point estimates). There is a large gap in pay for women without children. These women are no less productive than men without children, but their pay is about 12 percent lower than that of men without children.

For mothers, the earnings gap coincides with the output gap, suggesting that the pay gap reflects true differences in productivity for this group. If anything, mothers are slightly overpaid. The difference between their pay and productivity gap is less than three percentage points. This is consistent with Goldin [2014] who argues that mothers require flexible work arrangements and hours and that these types of hours may be less productive than continuously worked, long-duration hours in a variety of occupations. This result is also consistent with Azmat and Ferrer [2015] who provide direct evidence of differences in hours worked between men and women by studying hours billed by lawyers: female lawyers bill fewer hours than male lawyers and also draw fewer new clients to their firm. These gender differences among lawyers are most pronounced for mothers of young children.

Using the five industries in my data-set, I can estimate the relative productivity of mothers and non-mothers by industry. Figure 3 below displays the wage gap and the productivity gap for women without children relative to men and for mothers relative to men, by industry. Across industries, women without children are always more productive than women with children, and they are paid more. In all industries women without children are paid less than their productivity would imply. Mothers are not always paid less than their productivity and generally their wage gap is closer to their productivity gap.

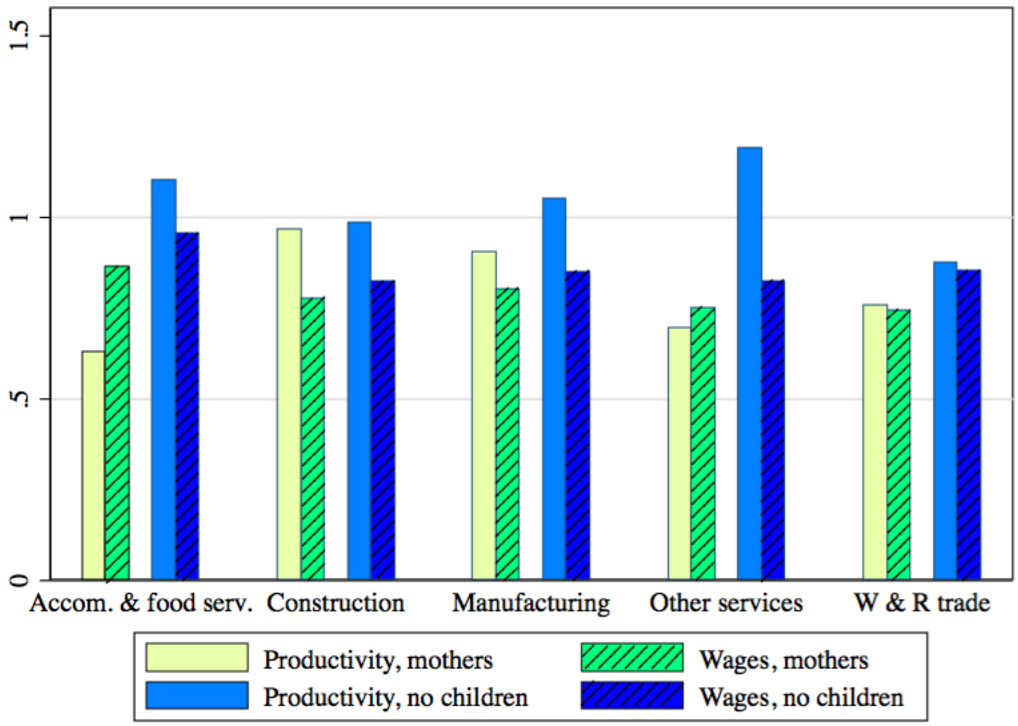

Figure 3: This figure gives the wage gap and productivity gap for women with children and without children compared to men. Hatched lines give the wage gap relative to men in the same industry. Solid lines give the productivity gap relative to men in the same industry. Mothers are on the left (in green) and women with no children are on the right (in blue).

One possible explanation for the result that non-mothers are relatively underpaid is that employers anticipate that these women may soon have children (which will cause a drop in their productivity). This uncertainty makes women without children relatively more "risky" to employ since legal protections prevent these women from being fired on the basis of having children and their relative productivity falls when they have children. If wages are somewhat sticky, employers will preemptively lower wages to compensate for the risk of having children in the future.

More specifically, the probability of having a child at age 30 is 13.8% in Denmark. Fertility rates are 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.

To test this expectations avenue, I split the male and female samples (female already split on presence of children) into two age groups: prime child-bearing age (25-25) and not. Women who are much older than 35 and have no children are very unlikely to have children. Similarly, women much younger than 25 are very unlikely to have children. In contrast, most women between the ages of 25 and 35 will have children at some point while they are in this age group. The interquartile range of age at birth is 28-34. Employers know that women without children between the ages of 25 and 35 will probably be having children in the near future. If they adjust wages for this group because of their expectation for future children, then the difference between the productivity gap and the wage gap (what I'm calling discrimination) should only show up for this group of women, not all women without children. In Figure 4 below, I report estimates of 5 βs, one for: non-prime age mothers, non-prime age women without children, prime-age men, prime-age mothers, and prime-age women without children. These measure productivity relative to non-prime age men (so $\beta = 1$ mechanically for non prime age men). Figure 4 below plots these estimates.

Prime child-bearing age women without children and men are more productive than other men. Prime child-bearing age mothers are slightly less productive than other mothers. In contrast, non-prime age women without children, though still more productive than mothers, are not more productive than their male counterparts and are much less productive than prime-age women without children[20].

Overall, the results suggest that women with children are less productive than men, as are non-prime-age women without children. Women between ages 25 and 35 who have no children are the most productive of any category. Nonetheless, in this age-group, women without children

[20]For some discussion of the differences between older women and men see for example Lusardi and Mitchell [2008]

are paid 18% less than prime-age men and mothers are paid 27% less than prime-age men. For the non-prime-age group, pay gaps coincide with productivity gaps. These results are consistent with some reduction in pay in anticipation of having children.

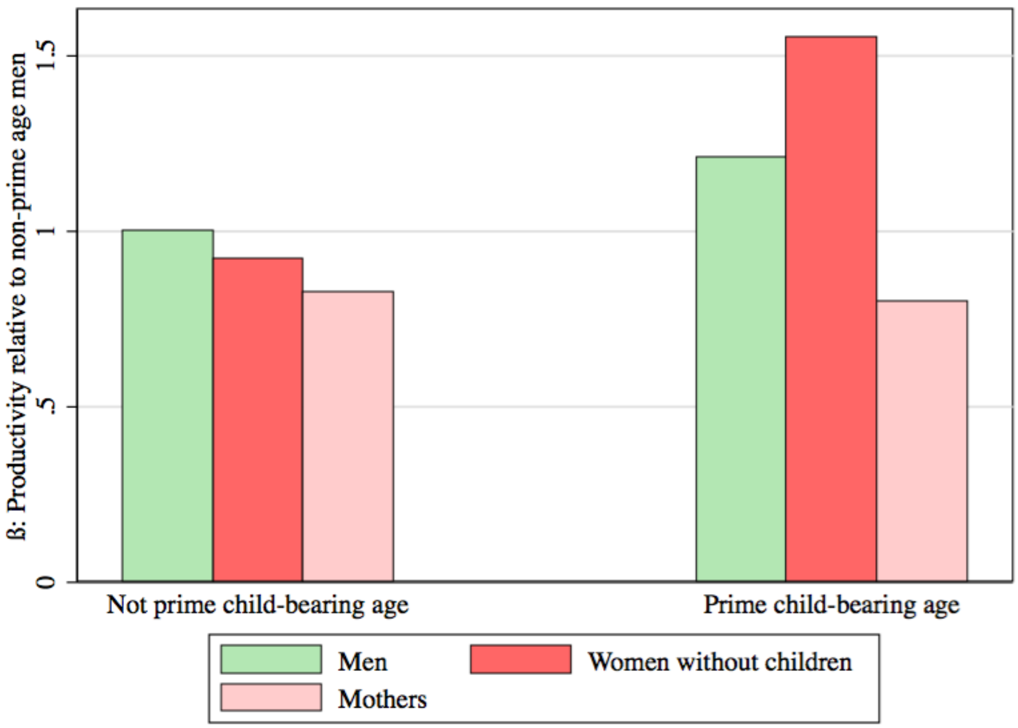

Figure 4: This figure gives the wage gap and productivity gap for prime child bearing age women without children, mothers, men and non-prime age women without children, and mothers compared to non-prime age men.

In the next section, I perform a variety of robustness checks. For simplicity, I will combine all women (regardless of age and number of children) into one category and compare them to all men, as in Table 6. Overall, I find that increasing the number of occupations does not affect estimation overall. I also find a great deal of heterogeneity by industry and occupation. Productivity gaps are quite small in Manufacturing and Construction and very large in accommodations and food services. In all industries, the productivity gap is smaller than the pay gap when controlling for selection. When allowing β to vary by occupation, I find that women who are white collar workers and managers are the most productive compared to men and are the only groups in which the productivity gap is smaller than the wage gap. I discuss these estimates in more detail below.

6.3 Heterogeneity and robustness

The overall estimate of the productivity gap has varied over time, most notably during the recession. Over the 15 year time period considered, β (my estimate of the productivity gap) fluctuates between just above 1 and just below 0.75. However, the large fluctuations are rare. In all but 3 years, β is estimated between 0.85 and 0.95. Figure 5 below plots the estimated β over time.

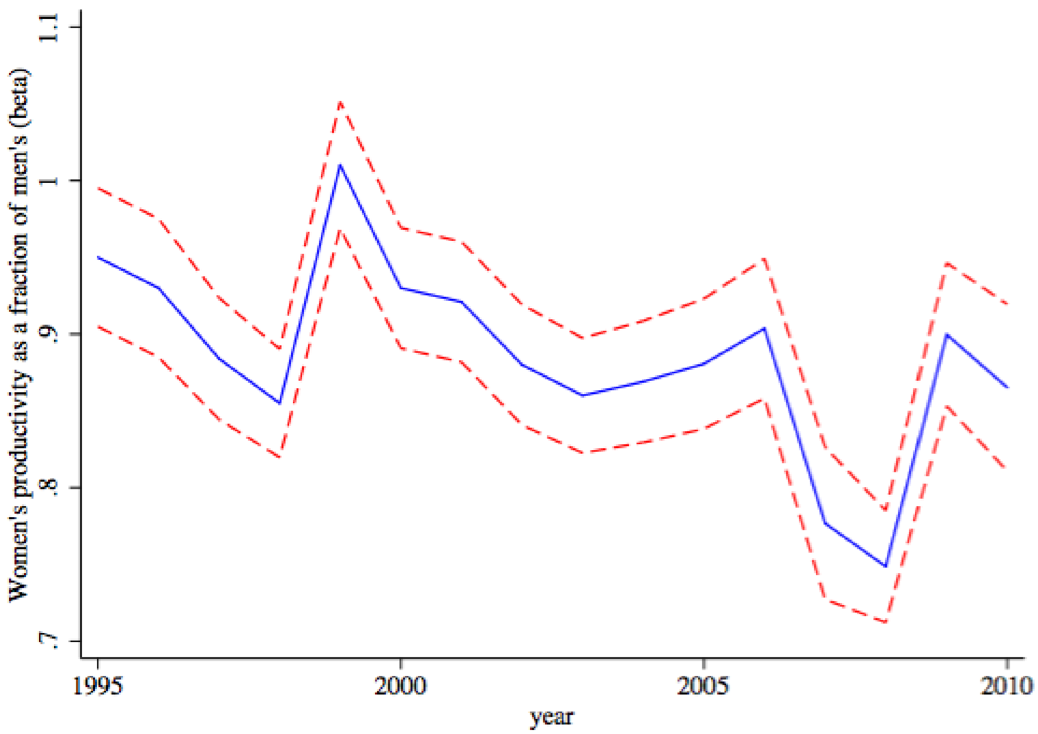

Figure 5: This figure display estimates β in(4), estimated year-to-year. The dashed lines are 95% confidence intervals with bootstrapped standard errors.

The largest changes come during the recession of 2007/2008. Denmark had a large recession (GDP growth of -5.1% in 2009, compared to -2.8% in the US). This recession followed a housing bubble and unemployment—particularly in construction and manufacturing—rose. The number of workers employed in the construction sector fell by 25% between 2007 and 2009 [Statistics-Denmark, b]. These changes asymmetrically affect men and women and result in a more selected sample of men compared to women in the industries studied here. The large productivity gap during the recession reflects compositional changes in the labor force resulting from layoffs of (primarily) the least productive men in the sample.

There is also a great deal of heterogeneity across occupations in the estimate of β. Table 10 below shows estimates of β for managers/executives, high skilled workers[21], white collar workers, and low skilled workers. The first column gives the wage gap by occupation (by occupation analogues of column 3 of Table 4). The second column provides estimates of β, by occupation, from the translog production function in capital and effective labor, where effective labor is now measured as

$$\mathcal{L} = \left[\left(\beta_m L_m^f + L_m^m \right)^{\frac{\sigma-1}{\sigma}} + \alpha_{hs}^i \left(\beta_{hs} L_{hs}^f + L_{hs}^m \right)^{\frac{\sigma-1}{\sigma}} + \alpha_{wc}^i \left(\beta_{wc} L_{wc}^f + L_{wc}^m \right)^{\frac{\sigma-1}{\sigma}} + \alpha_{ls}^i \left(\beta_{ls} L_{ls}^f + L_{ls}^m \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

The third column repeats this estimate on the subsample with non-negative investment data (the O-P sample). The final column estimates the production function (4) with $\mathcal{L}$ defined as above to obtain estimates of the productivity gap by occupation.

Cross-sectional estimates of β are similar to estimates which take into endogeneity for managers and white collar workers. For managers, the wage gap is very large (women are paid 76 cents on the dollar), and productivity gap is positive ten percent, implying a large amount of discrimination—in the form of asymmetrically compensated productivity—against female managers. Matsa and Miller [2013] and Bertrand et al. [2014] discuss a Norwegian reform which affected the number of women serving on boards and found that women in these high-level positions did have different management strategies than men, though they did not increase company profit.

In general my result is consistent with observational evidence from the US that female managers are more productive than male managers (Dezsö and Ross [2012], for example). Men make up a larger fraction of managers than of the other occupations. Facing some discrimination at

[21] workers with special technical skills and/or high education specific to their jobs

Table 10: Heterogenous β's by occupation

	w^f/w^m	β^o (Cross-sectional)	β^o (Cross-sectional, O-P sample)	β^o (O-P)
management	0.7649***	1.117**	1.129**	1.110**
[0.8742]	(0.0017)	(0.0277)	(0.0421)	(0.0518)
high skilled	0.8559***	0.9171**	0.8729**	0.7870***
[0.6938]	(0.0008)	(0.0211)	(0.0260)	(0.0314)
white collar	0.8256***	0.9352**	0.9317**	0.9678**
[0.6340]	(0.0003)	(0.0060)	(0.0070)	(0.0083)
low skilled	0.8609***	0.7132***	0.7071***	0.7849***
[0.6295]	(0.0006)	(0.0117)	(0.0090)	(0.0097)
N	18054253	527482	269445	269445

The fraction of workers in a given occupation is in hard brackets below that occupation name in column 1. The final three columns report the estimated productivity gap from 1. a cross sectional estimation of the translog production function with industry-specific occupation shares, 2. the same non-linear regression for the subsample of firms who are in the sample in 2 subsequent years with investment data, 3. a regression for the same subsample of firms with investment data which includes a control function in investment and capital (O-P regression). The number of observations for the wage regressions sum over all regressions. Standard errors are in parentheses.

the promotion to management stage of their careers, women who nonetheless become managers may be a very selected (and productive) group. Fryer [2007] gives a dynamic model of statistical discrimination in which discrimination benefits those members of minority groups who pass some threshold. The notion that female managers are more productive than male managers has been capitalized: the "Pax Ellevate Global Women's Index Fund" invests in companies which have women in high leadership positions such as Pepsico and Yahoo. A CEO is quoted on the fund's website: "Research suggests that where women are better represented, companies actually perform better." [Pax]

Wages do not reflect the higher productivity of female managers. Since negotiation plays a large role in determining the salary of managers, this result is consistent with Babcock and Laschever [2003] who find that MBA women are less likely to ask for raises than men. Bowles and McGinn [2008] discusses further the literature on gender and negotiation and argues women's traditional household responsibilities strongly affects their negotiating positions with employers.

White collar workers are workers who are doing non-manual labor tasks but do not have specialized education for their job. This category includes secretaries and most office workers.

Although white collar women are slightly less productive than white collar men, they are paid much less than white collar men. Accounting for endogeneity decreases the productivity gap, implying that women were sorted into less productive firms in this category.

In contrast to the case of managers and white collar workers, high skilled and low skilled women are paid more than their productivity gap. Over their lifetime, women change between these occupation categories. Figure 6 below plots the proportion of women in a given occupation by age in 2010 using only women working in the FIRE industries used in the regressions above.

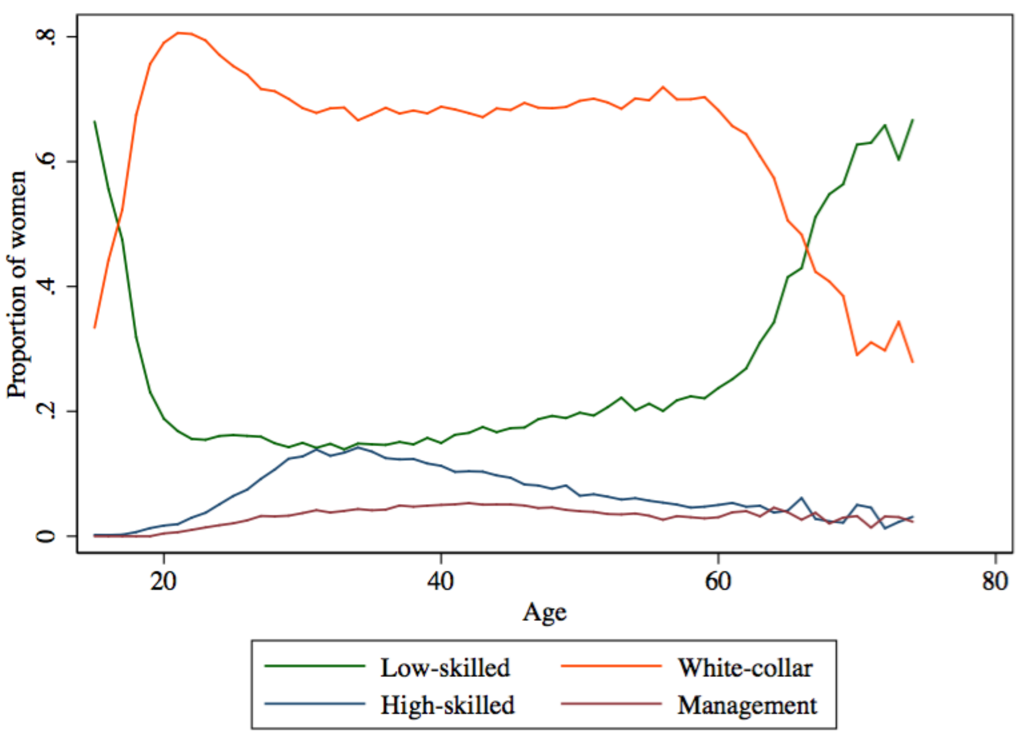

Figure 6: This figure graphs the proportion of women of a given age in a given occupation. Data is from 2010 and includes women between the age of 15 and 75.

When they are very young and very old, women tend to work low-skilled jobs. In-between, the majority of women are in white collar jobs. This suggests that most women do face discrimination in the form of being paid less relative to men than their marginal product relative to men. On the other hand, the 40% of women who do not work in white collar jobs (or management) are, if anything, overcompensated.

These estimates are robust to expanding the production function to include more occupations. If there is a great deal of error in measuring the occupation categories that workers belong in, and properly binning occupations matters for estimation of the production function (occupations are not perfect substitutes) then this will bias β towards 1. I consider two different methods for increasing the number of occupations. First, I re-write the CES production function to have 8 occupation inputs. By industry, I take the top 4 occupations (in terms of worker representation across firms) at the two digit level and then group workers not in these jobs into the 4 categories used previously. These categories can now be thought of as "other"-high-skilled, -white-collar, and -low-skilled (the management category stays the same). The data appendix lists the top 4 occupational categories, by industry. For example, in Accommodations and food-services, the top occupations are 1. Service and care work, 2. Cleaning and renovation work, budget and call service, telephone and doorstep selling, etc., 3. Retail sales and model work and 4. Internal office work. Table 11 below gives counts of the number of workers in each occupation.

Table 11: Occupation counts

Occupation	Worker count	Percent
Occupation 1	3,883,908	20.48
Occupation 2	1,912,575	10.08
Occupation 3	1,156,480	6.10
Occupation 4	958,049	5.05
Management	662,843	3.49
High-skilled	980,088	5.17
White-collar	5,625,392	29.66
Low-skilled	3,787,399	19.97

Occupation 1, 2, 3, 4 are the top occupations at the two digit level, by industry. See data appendix for the occupation names by industry.

Doubling the number of occupations in the CES production function increases the estimate of β by only half a percentage point, as in the second column of Table 12 below.

If we assume that occupations are perfect substitutes and within an occupation male and female workers have the same returns and are perfect substitutes themselves, then we can simply put occupation dummies into the regression estimating efficiency units from worker wages. Letting occupation enter the production function only by changing efficiency units expands the set of occupations greatly. In column 3 of Table 12 below, I use 100 distinct occupations (according to 2-digit DISCO codes) in the calculation of efficiency units and estimating a production function with $\mathcal{L} = \beta L_f + L_m$. This does not change the estimate of β.

Table 12: Estimates of β using a variety of occupation specifications

	Cross-sectional	O-P	O-P, 8 occupations	O-P all occupations
β	0.878***	0.882***	0.898***	0.897***
	(0.0038)	(0.0059)	(0.0060)	(0.0056)
N	527482	269445	269445	269445

This table provides OP estimates of β using expanded definitions of occupational categories in the production function. Column 1 gives the estimate using four occupation categories (management, high-skilled, white-collar, and low-skilled) in a CES production function for effective labor. Column 2 expands the number of occupation categories to 8, but adding the most frequent occupations by industry to the four categories above. Column 3 uses all information on occupation available (100 categories) by putting occupation dummies into the estimation of efficiency units.

In addition to occupational heterogenity in β, there are large differences in β across industries. This makes sense given the difference in the wage gap by industry. Also, there are differences in the proportion of a given industry that is female, and in how this varies by firm size across industries. In construction, manufacturing, and accommodation and food services, large firms have a high proportion of women relative to smaller firms. The opposite is true in wholesale and retail trade and other services. These differences suggest that selection may operate in different ways across industries.

Table 13 below provides by-industry estimates of β in the cross-section, adding occupations, and using an O-P correction for selection. Column 2 of Table 13 is analogous to the estimates of β reported in column 2 of Table 6, by industry (i.e.: β is fixed across occupations). Interestingly, the measured productivity gap is quite large in some industries, such as accommodations and food services, but it is negative in manufacturing (using cross sectional data).

Table 13: By-industry estimates

	w^f/w^m	β Cross-sectional	β Cross-sectional 8 occupations	β O-P	β O-P 8 occupations
All	0.8287***	0.878***	0.915***	0.884***	0.898***
	(0.0002)	(0.0039)	(0.0037)	(0.0056)	(0.006)
Accom. & food serv.	0.9241***	0.7884***	0.7339***	0.9351**	0.7961**
	(0.0013)	(0.0165)	(0.0153)	(0.0299)	(0.0264)
Construction	0.8193***	0.9616***	0.9313***	0.9778*	0.9132**
	(0.0010)	(0.0120)	(0.0019)	(0.0193)	(0.0165)
Manufacturing	0.8349***	1.0169*	0.9648**	0.9833*	0.9648*
	(0.0005)	(0.0101)	(0.0085)	(0.0111)	(0.0119)
Other services	0.7850***	0.9791*	0.9095**	0.9389*	0.8517**
	(0.0010)	(0.0154)	(0.0154)	(0.0414)	(0.0333)
Wholesale & retail trade	0.8226***	0.8585***	0.8946***	0.8295***	0.8804**
	(0.0006)	(0.0050)	(0.0051)	(0.0077)	(0.0079)

Number of observations is 269445 overall, 18519 Accommodations and food services, 60790 in Construction, 73443 in Manufacturing, 22391in Other services, and 94302 in Wholesale and retail trade. Standard errors in parentheses.

In most industries, selection does not dramatically change estimates. The exception is the accommodation and food services industry where β rises by 7-14 percentage points when using O-P controls for selection. Neumark [1996] conducts an small-scale audit experiment in the restaurant industry in which men and women were sent to apply to jobs and obtain pay quotes. He finds that women were more likely to be offered jobs in low-paying restaurants and men were more likely to be offered jobs in high-paying restaurants. This evidence on selection is consistent with women sorting into lower-productivity firms in the food services industry. In manufacturing, other services, and wholesale and retail trade, selection lowers estimates, suggesting overall positive sorting. In construction, selection doesn't seem to matter but the number of occupations included in estimation do affect β. In most industries, adding occupations causes β to fall, consistent with measurement error biasing β towards 1. The exception is wholesale and retail trade where β rises substantially when the number of occupations increases.

One potential reason for the productivity gap is customer discrimination: customers may dislike interacting with women rather than men. This sort of discrimination may be one reason that Azmat and Ferrer [2015] find female lawyers are less productive than male lawyers: if corporate clients only want male lawyers, billable hours and number of new clients (their measures of productivity) will reflect this discrimination. If this is the case, we might expect the productivity gap to correlate with proximity to customers: stronger in high-interaction industries like accommodation and food services, and low in low-interaction industries like manufacturing and construction. There is some support in my estimates for this customer-driven discrimination hypothesis, which from the firm perspective is a tangible difference in the relative benefit of employing a man vs. a woman. Estimates of the productivity gap are lowest in manufacturing and construction. Unfortunately, it is difficult to disentangle this avenue from selection—industries employing relatively few women have the smallest productivity gap.

7 Conclusion

This paper presented estimates of the relative productivity of men and women, accounting for age, education, experience, occupational choice, and hours worked. Overall, the results imply that the productivity of women is about 12 percent lower than men, controlling for age, education, experience, and hours worked. Since the residual wage gap is 16 percent, productivity differences explain three-quarters of the wage gap. This productivity difference may arise from differences in the effort, extra (undocumented) hours worked, or effectiveness of men relative to women. What remains of the residual wage gap—four percentage points—could be driven by discrimination by employers or differences in the bargaining ability of men relative to women. Another possibility is that wages are sticky and reflect the employer's beliefs about future productivity. This last possibility is consistent with evidence on the productivity gap for mothers compared to non-mothers: the gap in productivity is especially large for mothers. Women without children are as productive as men.

Even though mothers are less productive than non-mothers (who are approximately as productive as men), the earnings gap for mothers completely coincides with the productivity gap. This is also the case for fathers relative to men without children. Only women without children are paid less relative to men than their relative productivity would dictate. Delving more into the determinants of uncompensated productivity, I find that only women of prime child-bearing age who have no children are under-paid. If the employer takes into account the possibility that these women will have children in the future and cannot change the wages or employment of these workers based on whether they actually have children (but would like to) then he will optimally pay them less now in expectation of future productivity declines.

Ultimately, the structure of policy directed at reducing the residual wage gap depends on the source of the wage gap. I do not find on average that women are paid substantially less to do the same work as men. To the extent that there is discrimination in the labor market, it is more subtle than unequal pay for equal work. Knowing the magnitude of the productivity gap can help discipline policy aimed at reducing gender inequality. Knowing the sources of the gap between wages and productivity—women who may have children in the near future, women in management and doing white collar work—can help policy makers find effective avenues for reducing gender inequality.

8 Appendix 1: Data Appendix

I follow Baggar et al. [2014] exactly when calculating value added (revenue less cost of intermediate inputs) and capital. Their method of constructing value added and capital exactly follows the methods used by Statistics Denmark in calculating official national accounts from firm-level surveys. Table A1 below describes the definitions of variables used in the calculations [Statistics-Denmark, a].

Table A1: Variable names

Variables used in the calculation of value added	
Revenue	OMS
Work performed for own purposes and capitalized	AUER
Other operating income	ADR
End inventories minus starting inventories	DLG
Purchase of raw materials, finished goods, and packaging	KRH
Cost of energy	KENE
Purchase of contracting work, subcontracting	KLOE
Expenditure on rent	UDHL
Purchase of minor equipment	UASI
Other external expenses	OEEU
Secondary costs	SEUD
Claims in current assets (credits)	TGT
Expenses for temporary employment agencies	UDVB
Payments for long term rental and operational leasing	ULOL
External costs in general (except secondary items)	ANEU
Purchase of goods for resale (commodities)	KVV
Purchase of raw materials, finished goods, and packaging	KRHE
Variables used in the calculation of capital stock	
Operating equipment and other fixtures and inventory	AADI
Lands and buildings	GRBY
Plants and machinery	ATAM
Pre-paid material fixed assets and material fixed assets under construction	FMAA

Value added is revenue minus the cost of intermediate inputs. Between 1995 and 2011, there are three very minor changes in the definition of value added.

Table A2: Measures of value added

1995-1998	$Y = OMS + AUER + ADR + DLG$ $-KRH - KENE - KLOE - UDHL - UASI - OEEU - SEUD$
1999-2001	$Y = OMS + AUER + ADR + DLG + 0.0079TGT$ $-KRH - KENE - KLOE - UDHL - UASI - UDVB - ULOL - ANEU - SEUD$
2002-2003	$Y = OMS + AUER + ADR + DLG$ $-KRH - KENE - KLOE - UDHL - UASI - UDVB - ULOL - ANEU - SEUD$
2004-2011	$Y = OMS + AUER + ADR + DLG$ $-KVV - KRHE - KENE - KLOE - UDHL - UASI - UDVB - ULOL - ANEU - SEU$

Over the entire sample period, the stock of capital (K) is measured by the book value of material assets. This definition is unchanged over time:

$$K = AADI + GRBY + ATAM + FMAA$$

Efficiency wages are constructed using yearly pay to workers from tax records (JOBLON) and measures of education, experience, age, occupation, and hours worked. Hours worked are bracketed into interactions of

$$\left\{\begin{array}{l} \text{Full-time [37+hrs]} \\ \text{Part-time } [30-37\text{hrs}) \\ \text{Part-time } [29-30\text{hrs}) \\ \text{Part-time } [20-29\text{hrs}) \\ \text{Part-time } [10-19\text{hrs}) \\ \text{Part-time } (<10\text{hrs}) \end{array}\right\} \times \left\{\begin{array}{l} \text{Employed continuously} \\ \text{Employed serially} \end{array}\right\} \times \left\{\begin{array}{l} \text{Employed less than one year} \\ \text{Employed one year or more} \end{array}\right\}$$

The top four occupations, by industry are:

Accommodations and food services

1. Service and care work
2. Cleaning and renovation work, budget and call service, telephone and doorstep selling, etc.
3. Retail sales and model work
4. Internal office work

Construction

1. Working with mining and building crafts
2. Manual work in the construction sector, manufacturing, and transport
3. Metal and machine work
4. Internal office work

Manufacturing

1. Operation of industrial machinery
2. Metal and machine work
3. Manual work in the construction sector, manufacturing, and transport
4. Working with sales, finance, business administration, etc.

Other services

1. Working in the non-biological branches of science and computer science, statistics, architecture, and engeneering sciences
2. Cleaning and renovation work, budget and call service, telephone and doorstep selling, etc.
3. Working with sales, finance, business administration, etc.
4. Internal office work

Wholesale and retail trade

1. Retail sales and model work
2. Working with sales, finance, business administration, etc.
3. Internal office work
4. Manual work in the construction sector, manufacturing, and transport

9 Appendix 2: Appendix Tables

Table 14 below provides production function estimates

Table 14: Full production function estimates

	Cobb-Douglas	Translog	Translog, ind. specific shares	Translog, ind. specific shares O-P sample	O-P	O-P, ind. specific shares
β	0.930	0.909	0.878	0.859	0.891	0.884
	(.0051)	(0.0039)	(0.0038)	(0.0046)	(0.0059)	(0.0056)
σ	745.66	11.231	19.35	20.91	48.90	4066035
	(3285)	(0.4848)	(1.607)	(2.344)	(15.46)	(.)
ψ_1	0.684	1.198	1.175	1.197	0.746	0.741
	(0.0011)	(0.0041)	(0.0042)	(0.0051)	(0.0014)	(0.0014)
ψ_2	0.253	-0.349	-0.333	-0.3121	0.180	0.181
	(0.0007)	(0.0029)	(0.0029)	(0.0036)	(0.0008)	(0.0008)
ψ_3		0.068	0.0672	0.0633		
		(0.0002)	(0.0002)	(0.0002)		
ψ_4		-0.155	-0.156	-0.150		
		(0.0005)	(0.0005)	(0.0006)		
ψ_5		0.100	0.114	0.107		
		(0.0010)	(0.0012)	(0.0014)		
$\bar{a}$ Accom./food	4.597	5.867	5.670	5.700	5.001	4.726
	(0.0145)	(0.0182)	(0.0192)	(0.0244)	(0.0171)	(0.0866)
$\bar{a}$ Constr.	4.439	5.855	5.800	5.831	4.850	4.905
	(0.0139)	(0.0179)	(0.0186)	(0.0234)	(0.0162)	(0.0272)
$\bar{a}$ Manuf.	4.499	5.900	5.993	5.828	4.888	4.737
	(0.0141)	(0.0182)	(0.0181)	(0.0229)	(0.0164)	(0.0266)
$\bar{a}$ Other serv.	4.926	6.204	6.462	6.583	5.317	5.346
	(0.0143)	(0.0181)	(0.0187)	(0.0237)	(0.0167)	(0.0349)
$\bar{a}$ W/R trade	4.587	5.968	5.796	5.727	4.991	5.065
	(0.0140)	(0.0180)	(0.0182)	(0.0228)	(0.0613)	(0.0173)
N	527482	527482	527482	269445	269445	269445

Column 1 is the result of a non linear regression of value added on log effective labor units and log capital, with β, the coefficient measuring the substitutability between men and women in labor constrained to be the same across occupations. Column 2 and 3 estimate a higher order version of this regression. Column 3 in addition allows occupation-shares to vary by industry.

Table 15: Estimates of β when varying σ

	β Cross-sectional	β O-P
$\sigma = 3$	0.838	0.859
	(0.0048)	(0.0058)
$\sigma = 4$	0.846	0.868
	(0.0047)	(0.0057)
$\sigma = 5$	0.850	0.872
	(0.0046)	(0.0057)
$\sigma = 7$	0.854	0.876
	(0.0046)	(0.0056)
$\sigma = 10$	0.856	0.878
	(0.0046)	(0.0056)
$\sigma = 20$	0.859	0.880
	(00046)	(0.0056)
$\sigma = 50$	0.860	0.881
	(0.0046)	(0.0056)
$\sigma = 100$	0.860	0.882
	(0.0046)	(0.0056)
$\sigma = 1000$	0.860	0.882
	(0.0046)	(0.0056)
N	269445	269445

The estimates of β are analogous to column 5 and 7 of the appendix table above, but with σ fixed at the indicated values. β cross sectional is estimated from a translog production function with industry-specific shares in the CES labor aggregator and estimated on the O-P sample with positive investment data. β OP is also estimated with industry specific shares and a control function for investment.

The difference between male and female productivity seems driven by full-time workers, not part time workers. In the table below, I estimate the productivity of full-time women compared to full time men (β^f_{ft}), part-time women compared to full-time men (β^f_{pt}), and part-time men compared to full-time men (β^m_{pt}). I find that part-time women are as productive as full-time men and part-time men. In contrast, full-time women are less productive than full-time men. Since hours are accounted for in efficiency units, this suggests two things: 1. efficiency units do a good job of accounting for compensated productivity differences and 2. full time workers may have effort or hours heterogeneity which isn't captured by the data (and likely wouldn't be captured even with self-reported survey data). Full time men may be spending more time at work than full time women. This difference may explain a large portion of the productivity gap and the wage gap among full time, salaried workers.

Table 16: FT vs. PT

β^f_{ft}	0.845
	(0.0079)
β^f_{pt}	1.029
	(0.0090)
β^m_{pt}	0.961
	(0.0590)
N	269445

Standard errors in parentheses

Table 17: 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)
1996	0.014
	(0.0010)
1997	0.032
	(0.0010)
1998	0.104
	(0.0010)
1999	0.133
	(0.0009)
2000	0.151
	(0.0009)
2001	0.188
	(0.0009)
2002	0.200
	(0.0009)
2003	0.210
	(0.0009)
2004	0.223
	(0.0009)
2005	0.254
	(0.0009)
2006	0.296
	(0.0009)
2007	0.342
	(0.0009)
2008	0.383
	(0.0009)
2009	0.380
	(0.0009)
2010	0.382
	(0.0009)
2011	0.382
	(0.0010)

Standard errors in parentheses

This tables examines the importance of factors such as education and age (which are fairly well balanced across men and women) in estimation of the productivity and wage gap. The first column gives the estimated wage gap omitting this factor from wage regressions, while the second column gives the estimated productivity gap eliminating this factor from efficiency units. More precisely, the second column of each row gives the estimation of the production function in the third column of table 6 (so baseline β is 0.878), when eliminating one factor from the estimation of efficiency units (i.e.: "no age" means that efficiency units are calculated without including a quadratic in age). Experience and education affect the estimated productivity gap, while occupation—meaning having occupation specific returns—and age independent of experience do not.

Table 18: Robustness to factors in efficiency units

	w^f/w^m	β
No age	0.830	0.875
	(0.0003)	(0.0038)
No experience	0.805	0.858
	(0.0003)	(0.0037)
No education	0.815	0.858
	(0.0003)	(0.0038)
No age or experience	0.812	0.859
	(0.0003)	(0.0037)
No occupation	0.811	0.877
	(0.0003)	(0.0038)

Each row gives the estimation of the production function in the third column of table 6 when eliminating one factor from the estimation of efficiency units. In other words, "no age" means that efficiency units are calculated without including a quadratic in age. The first column gives the estimated wage gap omitting this factor from wage regressions, while the second column gives the estimated productivity gap eliminating this factor from efficiency units. Standard errors in parentheses.

10 Appendix 2: Appendix Figures

Over time, the relative productivity of mothers has fallen from just over 80 percent (compared to men) to a little under 80 percent. The estimates for non-mothers mover between being slightly more productive than men to slightly less. In all but three years, the relative productivity of non-mothers is statistically significantly higher than the relative productivity of mothers. Figure 4 plots the estimates of β for mothers and women without children year-by-year.

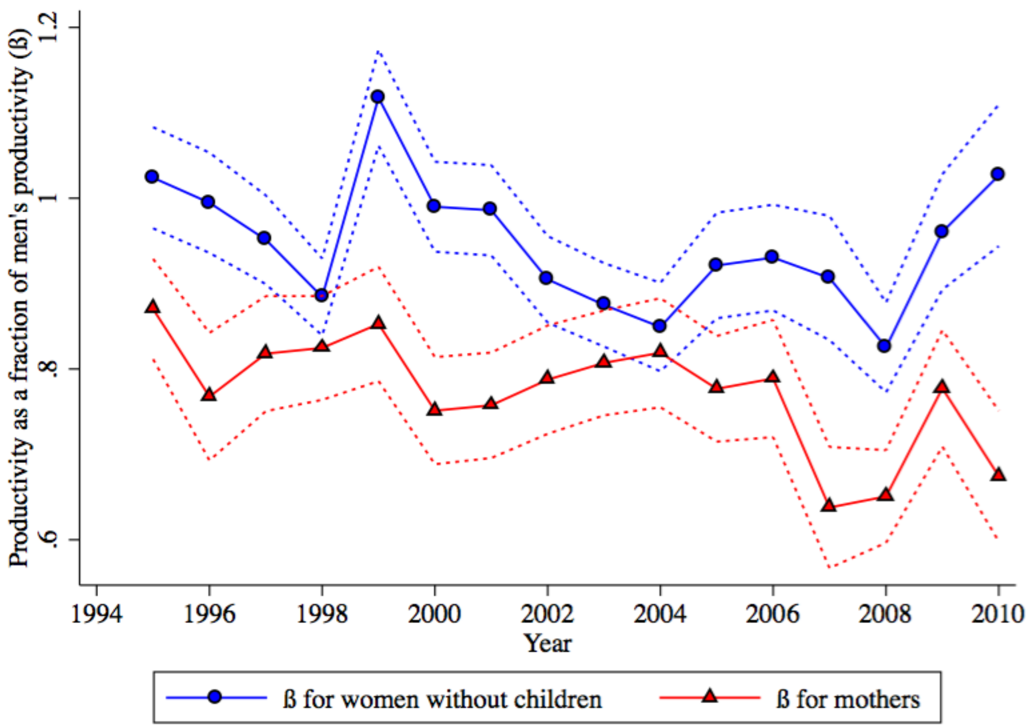

Figure 7: This figure gives the productivity for women with children compared to men and women without children compared to men over time.

Denmark has a labor market slightly more flexible than the US. The proportion of the workforce hired and separating within the year between 2000 and 2006 is about 27%. This is not true in other Nordic countries. In terms of employment flexibility, Sweden is comparable to France and Norway is comparable to Germany. The assumption that labor is a non-dynamic input is necessary for identification using investment as a control for unobservables known to the firm when they make hiring decisions. This assumption, to the extent that it is reasonable for any country, is reasonable for Denmark.

Figure 8: 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.

References

John M. Abowd, Francis Kramarz, and David N. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

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.

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.

Ghazala Azmat and Rosa Ferrer. Gender Gaps in Performance: Evidence from Young Lawyers. *Journal of Political Economy*, Forthcoming, 2015.

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

Jesper Baggar, Bent Jesper Christensen, and Dale T. Mortensen. Online Supplement to: Productivity and Wage Dispersion: Heterogeneity or Misallocation? Working Paper, 2014.

World Bank. World development indicators online (wdi) database. data.worldbank.org. Data retrieved May 15th, 2015.

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

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, 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, 2010. doi: 10.1257/app.2.3.228. URL `http://www.aeaweb.org/articles.php?doi=10.1257/app.2.3.228`.

Marianne Bertrand, Sandra E. Black, Sissel Jensen, and Adriana Lleras-Muney. Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labor Market Outcomes in Norway. NHH Dept. of Economics Discussion Paper No. 28, 2014.

Dan A. Black. Discrimination in an Equilibrium Search Model. *Journal of Labor Economics*, 13(2):309–333, 1995.

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.

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

Hannah Riley Bowles and Kathleen L. McGinn. Gender in Job Negotiations: A Two-Level Game. *Negotiation Journal*, 24(3):393–410, 2008.

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

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. *Quarterly Journal of Economics*, 2015.

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

Christian L. Dezsö and David Gaddis Ross. Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, 33, 2012.

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

Roland Fryer. Belief Flipping in a Dynamic Model of Statistical Discrimination. *Journal of Public Economics*, 2007.

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.

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.

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.

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.

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.

Annamaria Lusardi and Olivia S. Mitchell. Planning and Financial Literacy: How Do Women Fare? *American Economic Review*, 98(2):413–417, 2008.

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.

David A. Matsa and Amalia R. Miller. A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics*, 5(3):136–169, 2013.

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

David Neumark. Sex-discrimination in restaurant hiring: an audit study. *Quarterly Journal of Economics*, 111(3):915–942, 1996.

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.

Pax. Pax ellevate global women's index fund. www.paxellevate.com. retrieved September 22nd, 2015.

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

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

Statistics-Denmark. Accounting statistics for the private sector on the statistics denmark variable declarations page. www.dst.dk/extranet/staticsites/TIMES3, a. Data retrieved June 2nd, 2015.

Statistics-Denmark. Number of persons employed in the construction sector. www.dst.dk, b. Data retrieved Aug 5th, 2015.

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

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

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.