Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women*

David Card, UC Berkeley and NBER
Ana Rute Cardoso, IAE-CSIC and Barcelona GSE
Patrick Kline, UC Berkeley and NBER

June 2014

Abstract

There is growing evidence that firm-specific pay premiums are an important source of wage inequality. These premiums will contribute to the gender wage gap if women are less likely to work at high-paying firms or if women negotiate worse wage bargains with their employers than men. Using longitudinal data on the hourly wages of Portuguese workers matched with balance sheet information for firms, we show that the wages of both men and women contain firm-specific premiums that are strongly correlated with employer productivity. We then show how the impact of these firm-specific pay differentials on the gender wage gap can be decomposed into a combination of bargaining and sorting effects. Consistent with the bargaining literature, we find that women receive only 90% of the firm-specific pay premiums earned by men. Notably, we obtain very similar estimates of the relative bargaining power ratio from our analysis of between-firm wage premiums and from analyzing changes in firm-specific premiums over time. We also find that women are less likely to work at firms that pay higher premiums to either gender, with sorting effects being most important for lower-skilled workers. Taken together, the bargaining and sorting effects explain about one-fifth of the cross-sectional gender wage gap in Portugal. Our results suggest that regulatory policies aimed at ensuring equal pay are likely to have their greatest benefit for high skilled women, whereas policies ensuring that women are fairly represented in the hiring pool of firms will have effects throughout the skill distribution.

JEL Codes: J16, J31, J71

*We are grateful to Laura Giuliano, Michael Ransom, Jesse Rothstein, Andrea Weber, and seminar participants at California Polytechnic State University, Harvard University, RAND, University College Dublin, the University of Mannheim, the University of Potsdam and the University of Venice for many helpful comments and suggestions. We are also grateful to Alex Fahey for her expert assistance. We thank the Spanish Ministry of the Economy and Competitiveness (grant CO2012-38460) and the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075) as well as the Center for Equitable Growth and the Center for Labor Economics at UC Berkeley for generous funding support. An earlier version of this paper circulated under the title "Bargaining and the Gender Wage Gap: A Direct Assessment."

Labor market frictions give firms some control over wages. In a typical frictional equilibrium more profitable employers offer higher wages and workers search between jobs in pursuit of higher pay.[1] While a growing body of evidence confirms that inter-firm wage differentials are a significant source of earnings inequality, there has been little systematic analysis of how firm wage premiums contribute to pay gaps between demographic groups.

One of the most heavily studied and debated of these gaps is the wage differential found between men and women in virtually all developed economies. Despite rapid advances in the educational attainment and job experience of women, a large literature argues that the gender wage gap is still primarily driven by differences in the effective productivity of working men and women (e.g., Mulligan and Rubinstein, 2008). A more expansive view, suggested by models of frictional labor markets, is that equally productive men and women also often face different job prospects and strike different wage bargains with the same employers. Concern for such possibilities permeates the U.S. legal system, which emphasizes equal access to jobs at different firms and equal treatment *within a firm* by gender or race.

Two long-established strands of economic research suggest that firm premiums may in fact be important for understanding male/female wage differences. One emphasizes the role of bargaining and the possibility that women negotiate less aggressively than men (e.g., Bowles et al., 2005, 2007; Babcock et al., 2006).[2] The other focuses on the relative rate that women move to higher-paying jobs (e.g., Loprest, 1992; Hospido, 2009; Del Bono and Vuri, 2011).[3] These studies point to two complementary channels for generating gender disparities: a *bargaining* channel that arises if women obtain a smaller share of the surplus associated with their job, and a *sorting* channel that arises if women are less likely to be employed at higher-wage firms.

In this paper we provide the first comprehensive analysis of the impact of firm-specific pay premiums on the gender wage gap, using matched worker-firm data from Portugal that combines detailed information on the earnings of private sector employees with balance sheet data for employers.[4] Building on a simple rent-sharing model, we develop an approach to measuring the bargaining and sorting channels via a Oaxaca-style decomposition (Oaxaca, 1973; Fortin, Lemieux and Firpo, 2011) of gender-specific firm wage effects. Like Abowd, Kramarz and Margolis (1999) – hereafter, AKM – our model includes fixed effects for individual workers and fixed effects for employers that measure the wage premium paid by each firm relative to some reference firm or group of firms. A key issue for assessing the contribution of the bargaining channel is the need to define the relevant reference group for each

[1]See Burdett and Mortensen (1998), Coles and Mortensen (2012) and Moscarini and Postel Vinay (2013). Manning (2003), Postel-Vinay and Robin (2006), and Lentz and Mortensen (2010) provide overviews of microeconomic models of frictional labor markets.

[2]See Bertrand (2011) for a review. Related bodies of work suggest that women are less competitive and less confident than men, and as a result tend to choose occupations and jobs that are associated with lower wages, e.g., Gneezy, Niederle and Rustichini (2003), Niederle and Vesterlund (2007), and Buser, Niederle and Oosterbeek (2013).

[3]A related literature looks at between-firm distribution of employment by gender: e.g., Groshen (1991) and Petersen and Morgan (1995). Another set of studies focus on mobility rates of men and women without focusing on wage outcomes of the moves, e.g., Royalty (1998).

[4]An earlier study by Nekby (2003) relates male and female wages to measured profitability in a cross section of Swedish firms, but does not address the potential selectivity issues caused by non-random sorting of men and women with different unobserved skill characteristics to more profitable firms. Barth and Dale-Olsen (2009) examine firm-specific gender wage differences in a monopsony framework.

gender, a problem highlighted by Oaxaca and Ransom (1999). To solve this problem we exploit data on firm value-added to develop a data driven normalization that generates a lower-bound estimate of the differential bargaining power of women.

Since our decomposition approach builds directly on the two-way fixed effects model of AKM, we begin our empirical analysis by providing some non-parametric evidence on the plausibility of the additive separability and exogenous mobility assumptions underlying the AKM approach. Corroborating earlier exercises by Card, Heining and Kline (2013) with German data, and by Macis and Schivardi (2013) with Italian data, we find little evidence against the working hypothesis that these assumptions are approximately satisfied for both men and women in Portugal. Comparing the average wage changes for men and women who move between matched sets of firms we also provide non-parametric evidence that women benefit less from firm-to-firm mobility than men.

We then estimate separate AKM models for male and female workers in Portugal. We find that firm-specific pay premiums explain about 20% of observed wage variation among both men and women, while positive assortative matching (i.e., the positive covariation between worker and firm effects) explains another 10%. The firm-specific pay premiums for men and women are also highly (but imperfectly) correlated across firms. Decomposing the estimated firm premiums for firms that employ both male and female workers, we find that firm wage premiums tend to widen the gender wage gap. Part of the effect – accounting for up to 5% of the overall gender wage gap in Portugal – is attributable to the fact that women gain less than men from higher-wage firms. A larger share – 15% or more of the overall gender gap – is due to the under-representation of women at high-wage firms. In total, firm-specific pay premiums explain just over one-fifth of the average gender wage gap. In light of studies emphasizing the role of occupational choice in generating gender gaps (e.g., Goldin, 2014), we also examine the contributions of firm wage premiums to the gender gaps between workers in mainly female and mainly male occupations. We find that gender differences persist even within occupation groups. We also find that sorting effects are more important among less skilled workers, while bargaining effects are larger for the highly skilled.

We then narrow our focus to the component of the firm-specific wage premiums paid to men and women that is directly related to measured productivity. We find that women's wages are only 90% as responsive to average value-added per worker as men's, and we can easily reject the hypothesis of equal responsiveness, thus confirming that women gain a smaller share of firm-wide rents than their male co-workers. We also confirm that women are under-represented at firms with higher measured profitability. Bargaining and sorting based on measured productivity account for about 80% of the overall impact of firm-specific pay premiums on the gender wage gap.

As a final step in our analysis we examine the effects of changes in firm-specific profitability on the wage changes of men and women who remain with the firm over a multi-year period. This approach, which mirrors the design employed in the modern rent-sharing literature (e.g., Guiso, Pistaferri, and Schivardi, 2005; Carlsson, Messina and Skans, 2011; Card, Devicienti and Maida, 2014), uses an entirely different component of wage variation than our analysis of firm-specific pay premiums. Reassuringly, we obtain a nearly identical 90% estimate of women's relative bargaining power. Importantly, we suspect that our 90% estimate of the relative bargaining power of women in Portugal actually overstates the relative bargaining power of women in other labor markets where there are fewer institutional constraints on wages.

We conclude that, in the aggregate, firm-specific pay premiums explain a fifth of the gender wage gap in Portugal, with about three quarters of this effect arising through a between-firm sorting channel, and one quarter arising through a relative bargaining power channel. While modest in size, the relative bargaining power effect provides important confirmation of the hypothesis that women gain a smaller share of the rents than their male colleagues. Consistent with evidence from other studies, we find that bargaining effects are more important for high skilled women, while sorting effects are consequential throughout the skill distribution. These findings suggest that regulations aimed at ensuring equal pay are most likely to benefit high skilled workers, while regulations promoting inclusive hiring practices may be beneficial to all skill groups.

1 Firm-specific Determinants of the Gender Wage Gap

In traditional competitive labor market models, wages are determined by market-level supply and demand factors rather than the wage premiums of particular firms or the tenacity of particular workers (or, in the parlance of Sheryl Sandberg (2013)'s recent bestselling monograph on the subject, a worker's willingness to "lean-in").[5] A market-based perspective is central to Becker's (1957) model of employer-based discrimination. In Becker's model a *market-wide* discriminatory wage premium is determined by the preferences of the marginal employer of women. The mean wage gap between men and women then depends on their relative skills and a constant that reflects the supply of less-discriminatory firms.[6] Building on this framework, most studies of the gender wage gap focus on measured skill differences between men and women and treat any unexplained component as a combination of discriminatory factors and unobserved skill gaps (see Altonji and Blank, 1999 and Blau and Kahn, 2000, for reviews).

Despite the market-level focus of most economic studies, legislation aimed at eliminating gender discrimination is primarily directed at firms. In the U.S., for example, the Equal Pay Act requires that employers offer equal pay to men and women for "substantially equal" work, while Title VII of the Civil Rights Act prohibits firms from discriminating against women (and other protected groups) in decisions over hiring, layoffs, and promotions. Similarly, audit-based studies of potential discrimination (e.g., Heckman and Siegelman, 1993; Neumark, Bank and Van Nort, 1996; Bertrand and Mullainathan, 2004) examine the hiring practices of individual employers. Anti-discriminatory policies such as the use of blind auditions by orchestras (Goldin and Rouse, 2000) also focus on firm-specific behaviors.[7]

[5] Firms can still matter for observed wage outcomes if there are (market-based) compensating differentials for firm-wide amenities or disamenities, or if there is firm-specific human capital accumulation that is rewarded in pay. Robinson's (1933) model of monopsonistic wage setting was motivated in part by trying to explain why a firm might pay lower wages to women than men. As pointed out by Barth and Dale-Olsen (2009), this framework has been largely ignored in the gender wage literature. Lang and Lehmann (2012) discuss models of employer wage setting in a racial discrimination context.

[6] See Charles and Guryan (2008, 2011) for a recent application to the black-white wage gap and reviews of related work.

[7] Giuliano, Levine and Leonard (2009) show that changes in the race of a manager responsible for hiring can affect the fraction of non-whites who are hired. Such evidence suggests that a firm's policies and practices may be endogenous to the characteristics of its workforce.

The emerging literature on frictional labor markets provides a framework for understanding how firm-level behavior can matter for the gender wage gap.[8] To the extent that firms have some control over the wages offered to a given worker, the average wages of women relative to men will be affected by two factors. The first is whether firms pay different average wage premiums for men and women relative to the "market" (or a reference employer). The second is whether firms that pay higher wages, on average, are more or less likely to hire women.

One reason to suspect that the wage premiums paid to women and men may differ in a frictional labor market equilibrium is the finding in the social psychology literature that women are less likely than men to initiate negotiations with their employer (Babcock and Laschever, 2003; Bowles, Babcock and Lai, 2007), and in (lab-based) negotiation games are typically less successful negotiators.[9] As noted by Hall and Krueger (2012) about one-third of U.S. workers report having bargained with their employer over their starting wage, with bargaining being more common among men and higher skilled workers. Gender differences in bargaining behavior might be expected to translate into lower wages for women in these types of jobs. In fact, a study of Swedish college graduates by Save-Soderbergh (2007) found that women who obtained jobs where they were asked to submit a salary demand at the start of the job tended to ask for lower salaries (and ended up receiving lower salaries) than men.

The literature on inter-firm mobility and career progression provides additional insight into the potential role of firms in mediating the gender wage gap. Loprest (1992) documents that young women in the U.S. are about as likely to move between firms as young men, but experience smaller wage gains than men from each move. Similar patterns are reported for Italy by Del Bono and Vuri (2011) and for Spain by Hospido (2009). The smaller wage gains for women may be due to less aggressive negotiating by women at the start of their jobs. They could also arise because women place greater weight than men on non-wage factors like distance to home and flexibility of hours in choosing between the job offered by different firms, and are therefore less likely to move simply to raise their wages.

Both sources of gender differences are targeted by labor market regulations. Equal pay regulations aim to limit the ability of employers to bargain differently with men and women. Fair hiring rules ensure that hiring pools contain a proportionate share of women. Empirically assessing the relative importance of the bargaining and sorting channels, however, requires a comprehensive data set that includes men and women making transitions to and from the same groups of firms.

2 Institutional Setting and Data Overview

Our analysis relies on an unusually rich matched employer-employee data set from Portugal that includes *hourly* wage information – enabling us to control for differences in hours of work between men and women – as well as firm identifiers that allow us to link workers to financial information on

[8] See Manning (2003) for a discussion of gender-related pay differences in a wage posting model.

[9] Stuhlmacher and Walters (1999) present a meta-analysis of studies of gender in the bargaining literature from the 1974-1996 period. Nearly all the studies in their sample are from lab experiments. They conclude that on average women obtain a smaller surplus than men - their estimate of the differential is 10% of the standard deviation of the surplus share among men and women combined. Rigdon (2012) presents a lab experiment based on the Demand-Ultimatum game and finds relatively large differences between male and female players, both in terms of cheap talk during the game and game outcomes.

the owners of their workplace. Although our focus on Portugal is driven by the quality of these data, three features of the Portuguese labor market suggest that our findings may be broadly generalizable to other settings. First, women in Portugal have relatively high labor force participation rates. Fifty eight percent of adult women in the country were in the labor force in 2010 (ILO, 2012), comparable to the rates in the U.S. and Northern Europe. The participation rate of women between the ages of 25 and 45 is particularly high (over 85%) reflecting the strong commitment to work among younger cohorts of women (INE, 2012). Second, the vast majority of women in Portugal (over 90% of those in private sector jobs) work full time. As a result, wage comparisons between men and women are unlikely to be driven by differences between full-time and part-time jobs. Third, the gender wage gap in Portugal is within a few percentage points of the gaps in the U.S. and U.K., and is very close to the OECD-wide average.[10]

Most jobs in Portugal are covered by sector-wide pay agreements negotiated by employer associations and trade unions. Since these contracts are gender-neutral (i.e., they set wages for jobs regardless of gender), they presumably exert some equalizing effect on the relative pay of women.[11] Nevertheless, employers have wide latitude in assigning newly hired employees to job categories and in deciding who to promote to higher categories. In addition, most male and female workers earn substantial wage premiums over the base pay rates for their job category (Cardoso and Portugal, 2005). We suspect that these features lessen the impact of sectoral bargaining on the gender pay gap and leave substantial room for firm-specific factors to affect the relative wages of women.

2.1 Data Sources

Our wage data come from the Quadros de Pessoal (QP), an annual census of private sector employees conducted by the Portuguese Ministry of Employment. Firms with at least one paid employee are required to submit information on their full workforce as of the survey reference week (in October). Government employees and individuals working as independent contractors are excluded from coverage.[12] Over our 8-year sample period from 2002 to 2009 we have useable information on 4 million workers at 500,000 firms. Most Portuguese firms (over 96%) have only one workplace. For multi-plant firms we aggregate employees at different establishments to the firm level to reflect the fact that our financial data are firm-wide.

The QP asks firms to report each employee's gender, education, occupation, date of hire, regular monthly salary, wage supplements, and hours of work. Information is also collected on the industry, location, founding date of the firm, and gross sales in the preceding calendar year. We construct hourly wages by dividing each worker's normal salary and regular earnings supplements by his or her normal hours of work. The availability of hours information is a unique strength of the QP and reduces the

[10]The OECD "Family database" shows the gender gap in median full time earnings was 16% in Portugal, 19% in the U.S., and 16% on average across 26 OECD countries (OECD, 2012).

[11]Freeman and Leonard (1987) showed that trade unions in the U.S. have a (small) narrowing effect on the gender wage gap.

[12]Firm owners are included in the data set but do not report wages, and so are excluded from our analysis. Individuals who are on temporary leave (e.g., sick leave or maternity) are included in the data set, but those with no employer in the reference week are excluded.

potential impact of differential hours of work between men and women on the gender wage gap.[13]

We augment the information in the QP with balance sheet information from the "SABI" (Sistema de Analisis de Balances Ibericos) database. Businesses in Portugal are required to file annual balance sheets as well as profit and loss statements.[14] These reports are publicly accessible and are collected by financial service firms and assembled into the SABI database by Bureau van Dijk. Information in SABI includes the firm's name, address, industry, and founding date, as well as balance sheet data and total employment. SABI data are available from 2000 onward, but coverage of the database was relatively incomplete before 2006.

Since the QP does not include firm names or tax identification codes we use a combination of other variables that are reported in both QP and SABI to perform a fuzzy match. Specifically, we use location, industry code, firm creation date, annual sales, and end-of-year shareholder capital as matching variables. As described in the Data Appendix, we successfully match about 53% of firms that appear in the QP between 2002 and 2009 to a firm with at least one year's information in SABI. More information on our matching procedure and the match rates for various subgroups is presented in the Data Appendix. Overall we have current-year employer financial data for about 66% of the person-year observations in our QP sample from 2006-2009.

2.2 Descriptive Overview

As background for our analysis we begin by presenting a brief overview of the differences between male and female employees in Portugal. We focus on individuals who are between 19 and 65 years of age, have more than one year of potential labor market experience, and worked as a paid employee in the QP reference week. Our primary analysis sample – described in columns 1 and 2 of Table 1 – contains annual wage observations from 2002 to 2009 for 2.1 million men and 1.7 million women.[15]

Comparisons between the two columns show that female workers in Portugal are slightly younger than their male counterparts but are more likely to have completed secondary or tertiary education. Despite this education advantage women earn about 18% less per hour than men – very similar to the gender gap in median hourly wages in the U.S. in 2007 (EPI, 2010). Women also work slightly fewer hours per month, though the difference (3%) is small by international standards.[16] The dispersion in monthly hours is larger for women than men while the dispersion in hourly wages is smaller for women. Thirty-five percent of both male and female employees work in the Lisbon area, another 13 percent work in the Porto area, and the remainder work in smaller cities and rural areas.

[13]Differences in hours of work between men and women play a major role in explaining earnings differences in the U.S., particularly among the highly skilled. Wood, Corcoran and Courant (1993) and Bertrand, Goldin and Katz (2010) find important hours gaps between male and female lawyers and MBA graduates, respectively.

[14]These are filed with the Conservatoria do Registo Comercial. The same agency also keeps track of changes in ownership and organizational structure of firms. Based on informal discussions with firm owners we believe that the penalties for non-filing are small, possibly accounting for missing data for many firms.

[15]See the Data Appendix for details on the derivation of this sample, and comparisons with the overall population of 16-65 year old workers in the QP. In the small number of cases where an individual is employed at two or more firms in the reference week, we assign them to their highest earnings job.

[16]Data reported by the OECD (2012) for Portugal (based on labor force survey data that includes the government sector and independent contract workers excluded from QP) shows part-time employment rates for men and women of 8% and 14%, respectively. The same source shows part-time employment rates for men and women in the US of 8% and 17%.

As shown in Figure 1, the gender wage gap in Portugal was gradually narrowing over our sample period, reflecting stagnant real wages for men and modest wage growth for women. In the final year of our sample period nominal wages of men and women rose by about 3% (the same pace as in the previous 6 years) but the rate of inflation dropped sharply, causing a jump in real wages for both genders and leaving the gender wage gap at 16 percent.

Looking at workplace characteristics, women tend to work at larger firms than men (858 employees vs. 730), a characteristic that is also true in the U.S. and the U.K.[17] More striking is the difference in the share of female employees at women's and men's workplaces – 70% vs. 24%. This gap means that there is significant gender segregation across firms.[18] Indeed, about 21% of men work at all-male firms, while 19% of women work at all-female firms.[19] The presence of single-gender firms poses a problem for assessing the role of firms in the gender wage gap, since we cannot observe the wages that would be offered to women at all-male firms, or to men at all-female firms. For most of our analysis below we therefore limit attention to firms that hire at least one worker of each gender at some point in our sample period. Wages at single-gender firms tend to be relatively low, particularly for men: the mean log wage for men at all-male firms is 1.28 (31% below the average for all men) while the mean wage for women at all-female firms is 1.19 (22% below the average for all women). Paradoxically, the presence of single-gender firms therefore contributes to a narrowing of the gender wage gap relative to the gap at integrated firms.

An important issue for an analysis of between-firm wage differentials is the rate of job mobility, since these differentials are identified by the wage changes of job movers. In Appendix Table A3 we show the distributions of the number of jobs held by men and women in our overall QP sample (distributions for the subsamples shown in Table 1 are broadly similar). Approximately 73% of men and 74% of women hold only 1 job during our 2002-2009 sample period; 19% of both gender groups have 2 jobs; and 6% have 3 jobs. The remaining approximately 2% of men and 1% of women hold 4-8 jobs. The shares of the overall person-year observations in our samples contributed by these job-groups are similar. We also examined survival rates of new jobs that are observed starting during our sample period. As shown in Appendix Figure A1, about 40% of new jobs for both men and women last less than 1 year, and only about 20% of new jobs of both gender groups survive 5 years or longer. These comparisons suggest that between-firm mobility rates of women in Portugal are only slightly lower than the rates of men. Interestingly, similar findings are reported for full-time workers in the U.S. by Loprest (1992), and for broader samples of male and female workers in Spain by Hospido (2009), whereas in Italy, Del Bono and Vuri (2011) find a lower rate of firm-to-firm mobility for women than men.

[17] Papps (2012) and Mumford and Smith (2008) report roughly 10% larger workplace sizes for women than men in the U.S. and U.K., respectively.

[18] Mumford and Smith (2008, online Appendix Table A2) report that in the U.K. in 2004 the average fraction of female employees at the workplace was 70% for women and 34% for men, which implies somewhat less segregation by gender than in Portugal. We are unaware of any broad estimates of the degree of gender segregation in the US.

[19] About 20% of men and women at single-gender firms are the only employee at their workplace. Construction and trade account for 43% and 20%, respectively, of the person-year observations at all-male jobs. All-female workplaces are prevalent in trade (23% of person-years at all female firms), health services (17%), hotels (14%), and textiles (13%).

3 Modeling Framework

We now develop an econometric model that allows us to evaluate the effect of firm-specific pay premiums on the observed wages of women and men. Assume that we observe the point-in-time wages of a group of workers (indexed by $i \in \{1, ..., N\}$) in multiple periods (indexed by $t \in \{1, ..., T\}$). We denote worker i's gender by $G(i)$ which takes on values in the set $\{F, M\}$, and the identity of his or her employer in a given year by $J(i,t)$ which takes on values in the set $\{1, ..., J\}$. We refer to a particular gender as g and a particular firm as j.

We posit a rent-sharing model where the logarithm of the real wage earned by individual i in period t (w_{it}) is determined by:

$$w_{it} = a_{it} + \gamma^{G(i)} S_{iJ(i,t)t}. \tag{1}$$

Here, a_{it} represents the alternative wage available to worker i in period t, $S_{iJ(i,t)t} \geq 0$ is the match surplus between worker i and firm $J(i,t)$ in period t, and $\gamma^g \in [0,1]$ is a gender-specific share of the surplus captured by a worker of gender $g \in \{F, M\}$. In light of studies such as Bowles et al. (2005, 2007) we are specifically interested in the question of whether women get a smaller share of the surplus associated with their job (i.e. $\gamma^F < \gamma^M$).

We assume that $S_{iJ(i,t)t}$ can be decomposed into three components:

$$S_{iJ(i,t)t} = \bar{S}_{J(i,t)} + \phi_{J(i,t)t} + m_{iJ(i,t)}. \tag{2}$$

The first term, $\bar{S}_{J(i,t)}$, captures time-invariant factors like market power or brand recognition that raise the overall profitability of the firm. The second component, $\phi_{J(i,t)t}$, measures random time-varying factors that potentially raise or lower the overall rents available at the firm. The third component, $m_{iJ(i,t)}$, captures any person-specific surplus associated with the match between the worker and the firm, due for example to idiosyncratic skills or job requirements.

We also assume that the alternative wage a_{it} can be decomposed into a permanent component α_i (due, for example, to ability or general skills), a time-varying component associated with an observed set of characteristics X_{it} (e.g., labor market experience and changing returns to education), and a transitory component ε_{it}:

$$a_{it} = \alpha_i + X'_{it}\beta^{G(i)} + \varepsilon_{it}, \tag{3}$$

where β^g is a gender specific vector of coefficients.

Equations (1) through (3) imply the wage of worker i in period t takes the form:

$$w_{it} = \alpha_i + \psi^{G(i)}_{J(i,t)} + X'_{it}\beta^{G(i)} + r_{it}, \tag{4}$$

where $\psi^{G(i)}_{J(i,t)} \equiv \gamma^{G(i)} \bar{S}_{J(i,t)}$ and $r_{it} \equiv \gamma^{G(i)} \left(\phi_{J(i,t)t} + m_{iJ(i,t)}\right) + \varepsilon_{it}$ is a composite error. Equation (4) is consistent with an additive "two-way" worker-firm effects model of the type considered by Abowd, Kramarz and Margolis (1999) and many subsequent authors, with person effects, *gender-specific* firm effects, and gender-specific returns to the covariates X_{it}.

3.1 Exogeneity

We estimate models based on equation (4) by OLS, yielding gender-specific sets of firm effects. For these estimates to be unbiased we require the following orthogonality conditions to hold:

$$E\left[\left(r_{it}-\bar{r}_{i}\right)\left(D_{it}^{j}-\bar{D}_{i}^{j}\right)|G\left(i\right)\right]=0\ \forall j\in\left\{1,...,J\right\}, \tag{5}$$

where $D_{it}^{j}\equiv 1\left[J\left(i,t\right)=j\right]$ is an indicator for employment at firm j in period t and bars over variables represent time averages. To develop insight into the substantive restrictions involved in (5), it is useful to consider the special case where $T=2$. With two time periods, fixed effects estimation of firm effects is equivalent to first differences estimation and (5) can be restated as:

$$E\left[\left(r_{i2}-r_{i1}\right)\left(D_{i2}^{j}-D_{i1}^{j}\right)|G(i)\right]=0\ \forall j\in\{1,...,J\}. \tag{6}$$

Using the fact that $\left(D_{i2}^{j}-D_{i1}^{j}\right)$ takes on values of $+1$ for workers who move to firm j in period 2, -1 for those who leave firm j in period 1, and 0 for all others, we can write:

$$\begin{aligned}E\left[\left(r_{i2}-r_{i1}\right)\left(D_{i2}^{j}-D_{i1}^{j}\right)|G(i)\right] \;=\;& E\left[r_{i2}-r_{i1}|D_{i2}^{j}=1,D_{i1}^{j}=0,G(i)\right]\\ &\times P\left(D_{i2}^{j}=1,D_{i1}^{j}=0|G\left(i\right)\right)\\ &-E\left[r_{i2}-r_{i1}|D_{i2}^{j}=0,D_{i1}^{j}=1,G(i)\right]\\ &\times P\left(D_{i2}^{j}=0,D_{i1}^{j}=1|G\left(i\right)\right).\end{aligned}$$

The term $E\left[r_{i2}-r_{i1}|D_{i2}^{j}=1,D_{i1}^{j}=0,G(i)\right]$ gives the mean change in the unobserved wage determinants for the *joiners* of firm j, while the term $E\left[r_{i2}-r_{i1}|D_{i2}^{j}=0,D_{i1}^{j}=1,G(i)\right]$ gives the corresponding change for the *leavers* of this firm. Clearly, if worker mobility is independent of all three components in the composite error term, these expectations will be zero and OLS will be unbiased. Even if the expected changes in the residuals of joiners and leavers are non-zero, these biases will cancel out if the biases for joiners and leavers are the same and the probabilities of joining and leaving firm j are equal. The latter condition will true if the firm's employment is in steady state.

It is illustrative to consider potential reasons why workers joining (or leaving) a given firm might have unusual values for the change in their residual wage terms:

$$r_{i2}-r_{i1}=\gamma^{G(i)}\left[\phi_{J(i,2)2}-\phi_{J(i,1)1}+m_{iJ(i,2)}-m_{iJ(i,1)}\right]+\varepsilon_{i2}-\varepsilon_{i1}.$$

One reason is that mobility is related to the firm-wide shocks ϕ_{jt}. For example, workers may be more likely to leave firms that are experiencing negative shocks and join firms that are experiencing positive shocks. If this is true, however, then we would expect to see systematic dips in the wages of leavers just prior to their exit, and unusual wage growth for recent joiners. We look for such patterns below and find no evidence that they are present in the data.

A second source of correlation arises if workers tend to move to firms where they have a larger positive match effect (m_{ij}) and leave those where they have a smaller match effect. Indeed, mobility based on "comparative advantage" in wages is *assumed* in many formal models of worker-firm matching. A direct implication of such mobility is that workers who move from one firm to another will tend to experience wage changes of a different magnitude than people who move in the opposite direction. For example, suppose that firm A offers a 10% larger average premium than firm B. If mobility is independent of the match effects obtained by workers at the two firms, movers from firm B to firm A will experience a 10% average wage *gain*, while movers from firm A to firm B will experience a 10% average wage *loss.* If instead mobility is based in part on comparative advantage then the expected wage losses associated with moving from A to B will tend to be offset by an improvement in match effects. In the limit, if all firm transitions are voluntary and selection is based solely on the match components, *all* moves will lead to wage gains, as in the dynamic matching model of Eeckhout and Kirchner (2011). In our analysis below we examine workers moving in opposite directions between groups of high and low wage firms, and find that their wage changes exhibit the approximate symmetry (i.e., equal magnitude and opposite sign) predicted by an additive model with exogenous mobility. This symmetry is inconsistent with selection based on the match component of wages.

A third reason why the composite wage errors may be correlated with mobility is that the direction of firm-to-firm mobility is correlated with the transitory wage shock ε_{it}. For example, a worker who is performing well and receiving promotions may be more likely to move to a higher wage firm, while workers who are stalled in their job may be more likely to move down the job ladder to a lower-paying firm. Systematic mobility of this form implies that people moving to higher-wage firms will have different trends prior to moving than those who move to lower-wage firms. Again, in our analysis below we find no evidence for any of these predictions.

A final question is: What drives firm-to-firm mobility if it is *not* related to the elements in r_{it}? The most straightforward explanation is that worker-firm matching is based on a combination of the permanent component of worker ability (the α_i component in equations 3 and 4) and the average wage premiums offered by firms. Skilled workers, for example, are more likely to engage in on-the-job search (Pissarides and Wadsworth, 1994; Hall and Krueger, 2012) suggesting that they may be more likely to find jobs at high-wage firms. Skilled workers also may have networks of friends and family members that are more likely to work at high wage firms, leading to network-based sorting (as in Kramarz and Skans, 2013). Firms may be less likely to lay off high-skilled workers when demand is slack, leading to lower job destruction rates and a higher average probability of working at high-wage firms. These forms of sorting create no bias for our estimation strategy because we condition on time-invariant worker and firm characteristics. Finally, sorting based on non-wage dimensions such as the geographic location of the firm, recruiting effort, or firm size, creates no bias provided that these factors are uncorrelated with the time varying error component in (4).

3.2 Normalization

Exogeneity is necessary but not sufficient for identification of firm effects. As explained by Abowd, Creecy and Kramarz (2002) firm effects are only identified up to a normalizing constant within each "connected set" of firms linked by worker mobility – that is, the set of firms that have movers in common. In our analysis below we limit attention to workers and firms in the largest connected set for each gender. Even within a single connected set of firms, we still require a linear restriction on each gender's firm effects for identification. Intuitively, if a female worker moves from firm j to firm k, we can only use her wage history to infer the difference in female wage premiums between these firms. The levels of the wage premiums are not identified.

According to our simple model, the true firm effects for each gender should be non-negative, and zero at firms that offer no surplus above a worker's alternative wage. Using measured value-added per worker as a proxy for the average surplus at a given firm, we impose the normalization that firms with an observed level of value-added per worker below some threshold level τ have a (person-year weighted) mean firm effect of zero for both genders. More precisely, letting $\overline{VA}_j$ denote the average log value-added per worker at firm j over the sample period, we assume that:

$$E\left[\psi^g_{J(i,t)}|\overline{VA}_{J(i,t)} \leq \tau\right] = 0 \ \forall g \in \{F, M\} \tag{7}$$

If (7) is correct then imposing this condition will yield a set of normalized firm effects that coincide with the true firm effects. Otherwise, the normalized effects for each gender will be equal to the true firm effects, minus the average value of the firm effects for that gender group at firms with low value-added. Assuming that the wage premiums received by male workers at low value-added firms are at least as big as the premiums received by female workers, our normalization procedure will understate the relative size of the wage premiums received by male workers at higher productivity firms.

While the normalized effects could, in principle, be estimated in a single step by constrained least squares, we opt instead for a computationally convenient two-step approach. We first estimate the gender specific firm effects via unrestricted OLS, normalizing relative to a particular large firm by setting its firm effects to 0 for both genders. We then renormalize the effects by subtracting off the average value of the gender-specific firm effects at low value-added firms. We explain how we estimate the threshold τ in Section 5.3, below.

3.3 Decomposing the Effect of Firm-Level Pay Premiums

Equation (4) provides a simple framework for evaluating the impact of firm-level pay premiums on the gender wage gap. The average pay premium received by men is $E[\psi^M_{J(i,t)}|G(i) = M]$, while the average pay premium received by women is $E[\psi^F_{J(i,t)}|G(i) = F]$. As in the traditional Oaxaca wage decomposition (see e.g., Oaxaca, 1973; Fortin, Lemieux and Firpo, 2011), we can decompose the difference in pay premiums into the sum of a bargaining power effect and a sorting effect in either of two ways:

$$\begin{aligned} E[\psi^M_{J(i,t)}|G(i) = M] - E[\psi^F_{J(i,t)}|G(i) = F] &= E[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|G(i) = M] && (8)\\ &\quad +E[\psi^F_{J(i,t)}|G(i) = M] - E[\psi^F_{J(i,t)}|G(i) = F] \\ &= E[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|G(i) = F] && (9)\\ &\quad +E[\psi^M_{J(i,t)}|G(i) = M] - E[\psi^M_{J(i,t)}|G(i) = F]. \end{aligned}$$

The first term in equation (8) is the average bargaining power effect, calculated by comparing ψ_j^M and ψ_j^F across the distribution of jobs held by men. The second term in (8) is the average sorting effect, calculated by comparing the average value of ψ_j^F across the jobs held by men versus women. In the alternative decomposition (equation 9) the bargaining power effect is calculated using the distribution of jobs held by women, and the sorting effect is calculated by comparing the average value of the *male* pay premiums across jobs held by men versus women.

As noted by Oaxaca and Ransom (1999), the estimated sorting effect from either of these expressions is invariant to the normalization chosen for the firm effects. The estimated bargaining effect, however, is not invariant to the choice of normalization. Our procedure of normalizing the estimated effects to zero for a common set of low-value-added firms is equivalent to subtracting different constants from the two sets of fixed effects. Subtracting off different constants will obviously lead to a different value for the first line of either equation (8) or (9). Provided that the rents received by female workers at low value-added firms are no larger than the rents received by male workers at these firms, however, our choice of normalization will yield a lower bound estimate of the bargaining effect.

Note that under our assumed model of wage determination, the pay premiums for female and male workers at a given firm j are related by:

$$\psi_j^F = (\gamma^F/\gamma^M)\psi_j^M. \tag{10}$$

This suggests that we can estimate the relative bargaining power ratio γ^F/γ^M by examining the relationship between the estimated pay premiums for men and women across firms. Since these estimated premiums contain sampling errors, we average the pay premiums across groups of firms with similar characteristics and regress the average premium for female workers in each group on the corresponding average premium for male workers, providing an estimate of γ^F/γ^M.[20]

We can also use this estimate of the relative bargaining power of women to conduct a consistency check on the magnitude of the bargaining power effects estimated from (8) or (9). Re-arranging equation (10), notice that

$$\psi_j^M - \psi_j^F = (1 - \gamma^F/\gamma^M)\psi_j^M \tag{11}$$

Thus, the model predicts that $E[\psi^M_{J(i,t)} - \psi^F_{J(i,t)}|G\,(i)] = (1 - \gamma^F/\gamma^M)E[\psi^M_{J(i,t)}|G\,(i)]$. If, for example, the estimate of γ^F/γ^M is 0.9, then we expect that the estimated bargaining power effects from (8) and (9) should be approximately 10% of the estimated average pay premiums for men and women, respectively.

3.4 Relating the Estimated Firm Effects to Measures of Productivity

An alternative estimate of the relative bargaining power of male and female employees can be obtained by relating the estimated firm effects from equation (4) to value-added per worker. We assume:

$$E\left[\bar{S}_{J(i,t)}|\overline{VA}_{J(i,t)}, G\,(i)\right] = \kappa \max\left\{0, \overline{VA}_{J(i,t)} - \tau\right\}. \tag{12}$$

[20] Assuming that the groups are equal-sized, this approach corresponds to an instrumental variables estimate of the relationship given by (10), using as instruments the grouping dummy variables.

Note that this formulation imposes the normalizing assumption from equation (7) that firms with average log-value added per worker below τ have no surplus. We refer to the quantity $\max\left\{0, \overline{VA}_{J(i,t)} - \tau\right\}$ as "excess average value-added," $\overline{EVA}_{J(i,t)}$. Given the value of τ (which we estimate in a prior step, as explained in Section 5.3) we can write:

$$\psi^g_{J(i,t)} = \pi^g \overline{EVA}_{J(i,t)} + \nu^g_{J(i,t)} \tag{13}$$

where $\pi^g \equiv \gamma^g \kappa$ and $E\left[\nu^g_{J(i,t)} | \overline{EVA}_{J(i,t)}, G(i)\right] = 0$. Notice that $\pi^F/\pi^M = \gamma^F/\gamma^M$. By taking the ratio of the estimated gender specific slopes after estimating equation (13) for male and female workers we obtain an additional estimate of the bargaining power ratio γ^F/γ^M.

We can also use equation (13) to decompose the part of the firm-specific pay premiums received by men and women that is directly linked to our measure of firm-specific profitability. Specifically, the gender gap in these components is:

$$\begin{aligned}
& E[\pi^M \overline{EVA}_{J(i,t)}|G(i) = M] - E[\pi^F \overline{EVA}_{J(i,t)}|G(i) = F] \\
&\quad = (\pi^M - \pi^F)E[\overline{EVA}_{J(i,t)}|G(i) = M] + \pi^F \begin{pmatrix} E[\overline{EVA}_{J(i,t)}|G(i) = M] \\ -E[\overline{EVA}_{J(i,t)}|G(i) = F] \end{pmatrix} && (14) \\
&\quad = (\pi^M - \pi^F)E[\overline{EVA}_{J(i,t)}|G(i) = F] + \pi^M \begin{pmatrix} E[\overline{EVA}_{J(i,t)}|G(i) = M] \\ -E[\overline{EVA}_{J(i,t)}|G(i) = F] \end{pmatrix}. && (15)
\end{aligned}$$

Focusing only on the part of the firm surplus that is correlated with our measure of profitability, the contribution of the bargaining channel to the male-female wage gap is simply the difference in coefficients $(\pi^M - \pi^F)$, weighted by excess average value-added at men's jobs (equation 14) or at women's jobs (equation 15). The corresponding contribution of the sorting channel is the difference in mean excess value-added at men's jobs and women's jobs, weighted by either π^F (equation 14) or π^M (equation 15).

3.5 Within-Firm Changes in Productivity and Wages

While our main focus is on gender differences in *between-firm* wage differentials, our model also implies that the wages of male and female employees who are observed working at the same firm over time will respond differently to *changes* in firm surplus. Define VA_{jt} as the log value-added per worker at firm j in year t and $S_{jt} \equiv \bar{S}_j + \phi_{jt}$ as the firm-wide surplus in period t. Building on our normalization assumption, we assume that the surplus available for rent-sharing in period t is a function of the excess value-added per worker at the firm *in that period* (EVA_{jt}):

$$\begin{aligned}
S_{J(i,t)t} &= \lambda \max\left\{0, VA_{J(i,t)t} - \tau\right\} + \varsigma_{J(i,t)t} && (16) \\
&\equiv \lambda EVA_{J(i,t)t} + \varsigma_{J(i,t)t}.
\end{aligned}$$

We assume that the error $\varsigma_{J(i,t)t}$ has mean zero when we condition on the firm's value-added per worker in all periods and the characteristics of workers observed working at the firm continuously between an initial period $t = 1$ and a later period $t = T$ (i.e., a "stayer"). Using equation (4) we can therefore write:

$$E\left[w_{iT} - w_{i1} | VA_{J(i,1)1}, VA_{J(i,1)T}, X_{i1}, X_{iT}, G(i), Stayer\right] \\ = (X_{iT} - X_{i1})' \beta^{G(i)} + \theta^{G(i)}[EVA_{J(i,1)T} - EVA_{J(i,1)1}], \tag{17}$$

where $\theta^g = \gamma^g \lambda$ and $Stayer$ is shorthand for the conditioning event that worker i is continuously employed at the same firm throughout the sample period. Estimating this equation by OLS separately by gender yields estimates of the slope parameters θ^M and θ^F which can be used to form a third estimate of the relative bargaining power ratio γ^F/γ^M . Note that this estimate is based on within-firm variation in wages and profitability, while the previous estimates are all based on between-firm comparisons. Similarity of the various estimates of the relative bargaining power ratio therefore provides support for the simple rent-sharing model specified by equations (1)-(3).

A problem with OLS estimation of (17) is that annual value-added data is extremely noisy. To deal with this problem we also implement an IV strategy that uses the change in excess value-added between years 2 and $T - 1$ as an instrument for the change between years 1 and T. This is an imperfect solution both because we are considering a nonlinear function of value-added and because measurement errors might be serially correlated.

An alternate solution relies on the insight that according to our model:

$$E\left[w_{iT} - w_{i1} | J(i,1) = j, X_{iT} - X_{i1}, G(i) = F, Stayer\right] \\ = \frac{\gamma^F}{\gamma^M} E\left[w_{iT} - w_{i1} | J(i,1) = j, X_{iT} - X_{i1}, G(i) = M, Stayer\right].$$

That is, the covariate-adjusted average wage changes of male and female stayers at the same firm should be deterministically related by the gender bargaining power ratio. While the small number of stayers at each firm makes OLS estimation based upon this relationship between firm means infeasible, a straightforward instrumental variables solution is available. For each gender, we regress the change in wages on covariates and firm dummies to obtain adjusted average firm wage changes by gender. We then regress the adjusted average change in male wages at each firm on the corresponding average female change using the change in excess value-added between period 1 and T as an instrument and weighting by the total number of stayers at each firm. This yields a final (and arguably most robust) estimate of the bargaining power ratio γ^F/γ^M .

4 Descriptive Evidence on Firm-Specific Pay Premiums

Although the two-way effects model specified in equation (4) has been widely used over the past decade, the simple additive structure of the model and the restrictive assumptions needed for OLS estimation have been strongly criticized by some authors (e.g., Lopes de Melo, 2009; Eeckhout and Kirchner, 2011). It is therefore useful to scrutinize the model empirically and verify that the assumptions needed for OLS estimation are at least approximately satisfied.

Following Card, Heining and Kline (2013), we present simple non-parametric evidence on the wage changes of people who move between jobs with higher- and lower-paid co-workers. We document five basic facts that are all consistent with equation (4) and the related exogenous mobility condition (5). First, men and women who move between jobs with higher- and lower-paid co-workers experience systematic wage gains and losses, confirming that there are significant firm-specific pay premiums for both genders. Second, there is no indication that movers to firms with higher- or lower-paid co-workers experience differential wage trends prior to their move. Third, wage changes for people who move between firms with similarly-paid co-workers experience little or no excess wage growth relative to job stayers. Fourth, the gains and losses from moving between jobs with higher-paid and lower-paid co-workers are approximately symmetric, suggesting that the firm-specific pay premiums are additively separable (in logarithms) from other pay components and that mobility patterns are not driven by comparative advantage in wages. Fifth, women gain less than men from moving to jobs with more highly paid co-workers, as predicted by a rent-sharing model in which women get a smaller share of the rents than men.

We begin by selecting men and women from the overall analysis sample described in columns 1 and 2 of Table 1 who are employed at firms with at least one worker of each gender at some point in our sample period.[21] We construct mean log co-worker wages for each person in each year and divide all jobs for both men and women into four quartiles of co-worker wages, excluding the small number of workers at firms with only one employee. Then we identify job changers who are observed for at least two years at their origin firm and two years at their destination firm.[22] Finally, we construct average wages in the years before and after a move for male and female job-changers, classified into 16 cells based on the co-worker wage quartiles of their origin and destination jobs.

Figures 2a and 2b plot the wage profiles before and after the job change for men and women who moved from jobs in the lowest ($1st$) quartile of co-worker wages, and for those who moved from jobs in the highest ($4th$) quartile. The figures show that men and women who move from jobs with highly paid co-workers to jobs with poorly-paid co-workers experience large average wage losses, while those who move in the opposite direction experience large wage gains. Moving within a quartile group, by comparison, is associated with relatively small wage changes. Moreover, although the *levels* of wages on the old job differ between people from the same origin quartile who move to different destination quartiles, the *trends* prior to moving are very similar across groups. Likewise, wage trends on the new job are very similar across groups. These observations imply that inter-firm mobility is correlated with the permanent component of individual wages (i.e., the α_i component of equation 4) but not systematically correlated with the transitory error components (i.e., ϕ_{jt} or ε_{it}).

Table 2 summarizes all 16 groups of men and women, including information on the numbers of

[21] This sample contains 7.78 million person-year observations for men (86% of the male sample) and 6.30 million person year observations for women (87% of the female sample). In this sample the mean log wages of men and women are 1.646 and 1.440, respectively, and the gender wage gap is 0.21.

[22] This restriction means that our sample includes only new jobs that are observed starting between 2004 and 2008.

observations in each origin/destination group, the fractions of each origin group that move to each of the four possible destination groups, and the average wage change experienced by each group from two years before to two years after the move. Since the job changers in different origin and destination groups are heterogeneous in terms of age and education, we also construct an adjusted wage change, using the coefficients from a model of wage changes fit to the sample of job *stayers* who remain on the same job over a given four-year interval. We fit separate models for men and women, including year effects for the beginning year of the interval, dummies for education, and interactions of education dummies with age and age-squared. We then use the coefficients from these regressions to calculate predicted wage changes for movers, and calculate the deviation of each mover's actual wage change from his or her predicted change. Finally, we average the deviations for all workers in each origin-destination group.

The average adjusted wage changes for job changers who stay in the same co-worker wage quartile are all relatively small – e.g., 0.1% for male movers from quartile 1 jobs to other quartile 1 jobs, and -1.7% for female movers from quartile 2 jobs to other quartile 2 jobs – suggesting that mobility *per se* has little effect on wage growth. The only exception is for movers who stay in quartile 4. Male movers in this group experience a modest 5.3 percent wage gain (or 1.8% per year) faster wage growth than stayers, while women in this group experience a 6.1% wage gain (2% per year).

Mobility between co-worker pay quartiles, on the other hand, has relatively large effects on individual wages, even controlling for experience. Moreover, while not precisely symmetric, the mean wage changes for people who move in opposite directions between quartile groups (e.g, from quartile 1 to quartile 2, versus from quartile 2 to quartile 1) are of similar magnitude and uniformly of opposite sign. This is illustrated in Appendix Figures A2 and A3, where we graph the mean adjusted wage changes for all 16 origin-destination groups. Examination of these figures suggests that the symmetry restriction implied by the "no comparative advantage in firm-specific wages" assumptions is approximately satisfied for both men and women.[23]

Comparisons between the wage changes for men and women in Table 2 point to another important fact, which is that the changes for female movers tend to be smaller in absolute value than the corresponding changes for men. This is illustrated in Figure 3, where we plot the adjusted wage changes for each of the 16 origin-destination quartiles for women against the corresponding adjusted changes for men. The points lie tightly clustered around a line with a slope of approximately 0.8, confirming that women gain less from moving to jobs with more highly paid co-workers, and lose less from moving in the opposite direction. According to equation (4), the expected wage change for men who move from firm j to firm k is $\psi_k^M - \psi_j^M$ (ignoring any impact of the $X's$), while the expected wage change for women is $\psi_k^F - \psi_j^F = (\gamma^F/\gamma^M)(\psi_k^M - \psi_j^M)$ (making use of equation 10). Thus, the scatter in Figure 3 points to a relative bargaining power ratio for women of just under 80%.

A potential concern with these simple comparisons is that even within co-worker quartile groups,

[23]Since we only observe job status at a single point in time each year we cannot tell when the previous job ended, or why. We suspect that many of the transitions to higher-quartile firms are worker-initiated voluntary moves, while many of the transitions to lower-quartile firms arise from layoff and firing events. In any case, the approximate symmetry of the average wage changes for upward and downward transitions implies that the average match component in firm-specific wages is not systematically different for voluntary joiners and involuntary leavers, which is consistent with the assumptions needed for OLS estimation of (4).

males and females are not equally distributed across origin and destination firms. We therefore matched all female movers in the sample underlying Table 2 to male movers who made *exactly the same* firm-to-firm transition. Out of the 98,000 female job movers in Table 2 we were able to match about 8,200 (8.3%) to at least one male making the same transition. We then regressed the mean wage change for these female movers on the mean wage change for men making the same transition, instrumenting the male wage change with the change in their male co-workers' wages between the origin and destination firms. (Using the wage change for *male* co-workers as an instrumental variable ensures that the wages of the female movers are not included in the co-worker wages of the male movers).[24] The resulting coefficient estimate is 0.76, with a standard error of 0.03, providing a very similar estimate of the relative bargaining power of women. Given the highly selective nature of the matched female/male mover subsample, however, this estimate must be interpreted cautiously.

To summarize, our analysis thus far provides suggestive nonparametric evidence that firm wage effects are present and that firm mobility is (at least over the horizon we study) related to time invariant person components of wages but not to time varying or match components of wages. Furthermore, mobility *between the same sets of firms* influences the wages of men proportionally more than the wages of women, which is a qualitative sign that men possess greater average bargaining power than women.

5 Estimation of Worker-Firm Models

5.1 Estimation Sample

We turn now to a more systematic analysis of gender differences in pay premiums across firms. Building directly on equation (4), we fit models that include person effects, gender-specific firm effects, and a set of time-varying observable covariates with gender-specific coefficients. As noted earlier, the worker and firm effects for each gender group are only separately identified within a connected set of firms. For simplicity, we restrict our analysis to the largest connected set of firms for each gender. This allows us to identify the worker and firm effects for each gender, subject to a single normalization for each group. For estimation purposes we normalize the firm effects relative to the largest firm in the sample. We then re-normalize the estimated effects as described below.

The estimation samples are described in columns 3 and 4 of Table 1. Overall, 91% of all person-year observations for male workers and 88% of all person-year observations for female workers are included in the largest connected sets. The included workers have demographic characteristics and labor market outcomes very similar to those in our overall analysis sample. In particular, the mean gender wage gap is only 1 point wider (0.19 versus 0.18) for men and women in the largest connected sets than in our overall analysis sample. In much of our analysis below we further limit attention to workers who are employed at firms that are in the connected sets for both men and women. This dual-connected sample – described in columns 5 and 6 of Table 1 – includes 66% of the person-years of male workers in our analysis sample, and 69% of the person-years of female workers. Individuals in the

[24] The IV model is estimated on a subsample of 6,912 observations in which there are observed wages for male co-workers at the origin and destination firm.

dual-connected set have higher education than in the workforce as a whole, and also have somewhat higher average wages. The gender wage gap is also larger in this sample than in our overall sample (23% versus 18%), reflecting the omission of the single-gender firms, which as noted earlier have a relatively small gender gap.

5.2 Estimation Results

Columns 1 and 2 of Table 3 summarize the parameter estimates and fit of our models for men and women in the samples described in columns 3 and 4 of Table 1 (i.e., the largest connected sets of workers of each gender). The models include fixed effects for workers and firms (a total of 2.1 million dummies for men and 1.7 million dummies for women) as well as year dummies, fully interacted with education dummies (for the 4 education groups shown in Table 2), and quadratic and cubic terms in age interacted with education dummies.[25]

We show the standard deviations of the estimated person and firm effects and the covariate indexes $\left(X_{it}'\widehat{\beta}^g\right)$ for each observation, as well as the correlation of the person and firm effects, the residual standard deviation of the model, and the adjusted R-squared statistics. For both males and females, the standard deviations of the person effects are nearly twice as big as the standard deviations of the firm effects, implying that a relatively large share of wage inequality for both genders is attributed to worker characteristics that are equally rewarded at all firms. The correlations between the estimated person and firm effects are both positive, implying that more highly-skilled men and women are disproportionately employed at firms that pay all their workers a bigger wage premium. Such positive assortative matching has been found in many recent studies of wage determination.[26]

The middle panel of Table 3 shows fit statistics for a generalized model of wage determination that includes unrestricted dummies for each job match. This model provides a slightly better fit to the wage data for both men and women, with adjusted R-squared statistics that are about 1 percentage point higher (e.g., 0.951 versus 0.940 for women). Comparing the residual standard error of the job match model to the corresponding standard error for the model with worker and firm effects we can construct an estimate of the standard deviation of the permanent job match effects (the $m_{iJ(i,t)}$) that are absorbed in the job match model but included in the residual of (4). The estimates are 0.062 for men and 0.054 for women – only about one-quarter as big as the standard deviations of the firm effects for the two genders. We conclude that the firm-wide component of job match surplus is considerably larger than the purely idiosyncratic component.

We have also conducted a series of additional specification checks of the fit of our basic models. In one check, we examine the mean residuals from equation (4) for subgroups of observations classified by the decile of the estimated person effect and the decile of the estimated firm effect. As shown in Appendix Figures A4 and A5, we find that the mean residuals are very small in all 100 cells for

[25]Estimates were computed using a preconditioned conjugate gradient algorithm as in Card, Heining and Kline (2013).

[26]See e.g., Card, Heining and Kline (2013) for West Germany, Mare and Hyslop (2006) for New Zealand, Skans, Edin and Holmlund (2008) for Sweden, and Bagger, Sorensen and Vejelin (2012) for Denmark. The sampling errors in the estimated person and firm effects from a model such as (4) are in general negatively correlated (see e.g., Mare and Hyslop, 2006; Andrews, Shank and Upward 2008), implying that the correlations between the estimated effects are downward biased estimates of the degree of assortative matching.

both genders, suggesting that the additive structure of (4) provides a good approximation to the wage-setting process in Portugal. In a second check, we examined the mean residuals for workers who transition between groups of firms, classified by the quartile of the (gender-specific) estimated firm effects. We find that the mean residuals are small in magnitude for groups of men and women who move up or down the firm quality distribution.

The bottom rows of Table 3 present the main components of a simple decomposition of the variance of wages across workers implied by the fitted version of equation (4):

$$\begin{aligned} Var\left(w_{it}\right) &= Var\left(\hat{\alpha}_i\right) + Var\left(\hat{\psi}^{G(i)}_{J(i,t)}\right) + 2Cov\left(\hat{\alpha}_i, \hat{\psi}^{G(i)}_{J(i,t)}\right) \\ &\quad + Var\left(X'_{it}\hat{\beta}^{G(i)}\right) + Var\left(\hat{r}_{it}\right) + 2Cov\left(\hat{\alpha}_i + \hat{\psi}^{G(i)}_{J(i,t)}, X'_{it}\hat{\beta}^{G(i)}\right). \end{aligned} \tag{18}$$

Among both male and female workers, person effects accounts for about 60% of overall wage variation, firm effects account for about 20%, and the covariation in worker and firm effects accounts for an additional 10%. In both cases the contribution of the measured covariates (including the main effects and the covariances with the person and firm effects) is relatively small, and the residual component is also small, reflecting the relatively high R-squared coefficients for the underlying models.

5.3 Normalizing the Estimated Firm Effects

As noted in Section 3.2, we normalize the firm effects relative to a set of firms that arguably offer no surplus to workers of either gender. Figure 4 shows the relationship between our estimated male and female firm effects (normalized for purposes of estimation by setting the effects to zero for the largest firm in the sample) and average log value-added per worker. We group firms into percentile bins of value-added and plot the mean estimated firm effects for men and women at the firms in each bin against average log value-added per worker for firms in the bin.[27]

A striking feature of this figure is the piecewise linear nature of the relationship with value-added. Firms in the bottom percentile groups pay very similar average wages, while at higher percentiles the firm-specific wage premiums for both men and women are linearly increasing in log value-added per worker, suggesting a constant elasticity relationship between wages and value-added above a value-added "kinkpoint." To identify the value-added threshold more formally, we fit a series of bivariate regression models of the form:

$$\begin{aligned} \hat{\psi}^M_{J(i,t)} &= \pi^M_0 + \pi^M \max\left\{0, \overline{VA}_{J(i,t)} - \tau\right\} + \nu^M_{J(i,t)} \\ \hat{\psi}^F_{J(i,t)} &= \pi^F_0 + \pi^F \max\left\{0, \overline{VA}_{J(i,t)} - \tau\right\} + \nu^F_{J(i,t)}. \end{aligned} \tag{19}$$

where (as above) $\overline{VA}_j$ represents the average of log value-added per worker at firm j, and τ is a threshold beyond which the firm begins to share rents. We estimated these equations for a range of values of τ using data on all firms in the dual connected sample that can be matched to the financial data set.[28] We then selected the value of τ that minimized the mean squared error of the system of

[27] We define the groups so they have equal numbers of person-year observations, using the sample of men and women at dual-connected firms with merged financial data.

[28] We fit the model to firm-level data using the 47,477 dual connected firms with matched financial data. These firms

two equations. Consistent with the pattern of the points in Figure 4, this procedure selects a value of $\hat{\tau} = 2.45$: the estimated values of the coefficients π^M and π^F are 0.156 and 0.137, respectively.[29] We show the fitted relationship with blue and pink lines in Figure 4.

The implied set of "no rent" firms (i.e., those with $\overline{VA}_j < \hat{\tau}$) account for 9.4% of all person-years at dual-connected firms in our sample with financial information (6.6% of male person-years and 12.9% of female person-years). Workers at these firms have relatively low education and receive relatively low wages (42% below average among men, and 28% below average among women). Given the estimate $\hat{\tau}$ we then determine the values of normalizing constants such that the re-normalized effects for both genders have employment-weighted averages of zero across all firms with $\overline{VA}_j < \hat{\tau}$.[30] To check the sensitivity of our normalization procedures, we used a nonparametric bootstrap procedure to estimate the standard error of the estimate of τ. The resulting estimated standard error is 0.09, implying a 95% confidence interval of $[2.27, 2.63]$. We then re-calculated the normalizing constants using the upper and lower values of this confidence interval. We obtained constants that are quite close to the baseline constants for $\hat{\tau} = 2.45$, suggesting that our procedure is relatively insensitive to sampling errors in $\hat{\tau}$.

Figure 5 graphs the normalized firm effects for women against the corresponding effects for men, using the same 100 groups as in Figure 4. There is clearly a strong relationship between the average premiums paid to male workers in each group and the average premiums for female workers. In fact, even at the individual firm level the employment-weighted correlation of $\hat{\psi}_j^F$ and $\hat{\psi}_j^M$ is 0.59 across firms in the dual-connected set, and the corresponding regression of $\hat{\psi}_j^F$ on $\hat{\psi}_j^M$ has a slope of 0.56. Given the presence of sampling errors in the estimated firm effects, however, this is a downward-biased estimate of the rent-sharing ratio γ^F/γ^M. Grouping firms by values of $\overline{VA}_j$ averages out the sampling errors and yields an estimated slope coefficient of 0.89, implying that women have about 10% less bargaining strength than men.

6 Firm-specific Pay Premiums and the Gender Wage Gap

6.1 Decompositions Based on Estimated Firm Effects

Given the normalized firm effects for men and women we can quantify the impact of firm-specific pay premiums on the gender wage gap, and use equations (8) and (9) to decompose the total effect into components due to sorting and relative bargaining power. We begin in Table 4a by showing the various terms in these equations. Row 1 shows the mean log wages of men and women in the dual connected sample and the mean gender wage gap, which is 23.4%. Row 2 shows the means of the estimated firm effects for men across male person-year observations (column 1) and across female person-year observations (column 2), and the difference between these means (column 3). Similarly, row 3 shows

account for 6.95 million person-year observations in our data set, or 63% of the person-year observations at dual-connected firms.

[29] Appendix Figure A6 shows the adjusted R-squared from the bivariate system for a range of values of τ and the associated estimates of the coefficients (π^M, π^F).

[30] This is approximately the same as subtracting the estimated values of the constants $\hat{\pi}_0^M$ and $\hat{\pi}_0^F$ from $\hat{\psi}_j^M$ and $\hat{\psi}_j^F$, respectively. The difference arises because we choose the normalizing constants based on the employment-weighted average of the firm effects for all firms with $\overline{VA}_j < \hat{\tau}$.

the means of the estimated female firm effects among male and female person-year observations, and the difference. Our estimate of $E[\psi_j^M|G(i) = M]$ – the average rents received by male workers – is simply the mean of the estimated male firm effects among males, which evaluates to 14.8%. Our estimate of $E[\psi_{J(i,t)}^F|G(i) = F]$ – the average rents received by female workers – is the mean of the estimated female firm effects among females, which evaluates to 9.9%. Thus, our estimate of the total contribution of firm-specific pay premiums to the gender wage gap is 14.8% - 9.9% = 4.9%, which is displayed in row 5, column 3.

The part of this total that is attributable to the bargaining channel can be calculated in either of two ways. One is to compute the average difference between the male and female firm effects for each firm, weighted by the shares of men at each firm (i.e., $E[\psi_{J(i,t)}^M - \psi_{J(i,t)}^F|G(i) = M]$) – this is the difference shown in row 4, column 1, and is equal to 0.3%. The alternative is to compute the average difference between the male and female firm effects for each firm, weighted by the shares of women at each firm (i.e., $E[\psi_{J(i,t)}^M - \psi_{J(i,t)}^F|G(i) = F]$). This difference, shown in row 4 of column 2, is equal to 1.5%. We therefore estimate that the bargaining channel explains a 0.3 to 1.5 percent gap in male-female wages, 1.2 to 6.3 percent of the overall gender wage gap.

Similarly, the impact of the sorting channel can be calculated in two ways. The first is to compute $E[\psi_{J(i,t)}^M|G(i) = M] - E[\psi_{J(i,t)}^M|G(i) = F]$. As shown in row 2, column 3 this implies a 3.5% effect. The second is to compute $E[\psi_{J(i,t)}^F|G(i) = M] - E[\psi_{J(i,t)}^F|G(i) = F]$. As shown in row 3, column 3 this is 4.7%. Thus, the sorting channel explains 14.9 to 19.9 percent of the overall male-female wage gap.

The results in Table 4a point to several interesting conclusions. Relative to the least productive firms in the labor market, men earn about a 15% average wage premium while women earn a 10% premium. The 5 percentage point gap accounts for a little over one-fifth of the gender wage gap at mixed-gender firms. Three quarters or more of this overall effect is explained by the relative concentration of women to firms that offer lower wages to both men and women. This finding agrees with the results of Cardoso, Guimarães and Portugal (2012), who conclude that roughly a fifth of the gender wage gap in Portugal can be attributed to the differential sorting of women across firms.[31] The estimated impact of differential bargaining power is smaller, and explains at most 1.5 points of the 23.4 point gender gap. While modest, this 1.5 point effect is consistent with the prediction following from equation (11), given an estimate of the rent-sharing ratio γ^F/γ^M equal to 0.9 and an estimate of the mean rents of male workers of roughly 15%.

Table 4b present a similar analysis for different age and education subgroups. Column 1 of the table shows the mean log wage gap between men and women for the subsample indicated by the row heading (e.g., 23.4% for the overall sample in row 1). Column 2 shows the mean value of the (normalized) male firm effects for men in the dual connected sample, which is our estimate of the rents earned by men – while column 3 shows the mean value of the (normalized) female firm effects for women, which is our estimate of the rents earned by women. The gap between these two (column 4) is our estimate of the total contribution of firm-specific pay premiums to the gender wage gap (e.g., 4.9%

[31] Cardoso, Guimarães and Portugal (2012) estimate a worker-firm model imposing that the firm effects are identical across the genders. This ignores the relative bargaining power effect.

for the overall sample). Columns 5-8 present the terms in the alternative decompositions described by equations (8) and (9). The entry in column 5 is the difference between the average value of the *male* firm effects at jobs held by men versus women, and is our estimate of the sorting effect using male firm effects. The alternative estimate of the sorting effect in column 6 uses female firm effects. Finally, the entries in columns 7 and 8 show the estimated contributions of relative bargaining channel, using the distribution of men across firms (column 7) or the distribution of women across firms (column 8).

Comparing across age groups (rows 2-4) the entries in column 1 show that the male-female wage gap in Portugal widens dramatically with age, from around 10% for workers under 30 to 33% for workers over the age of 40. Firm-specific pay premiums contribute to this age pattern, rising from 2.8% for workers under the age of 30 to 6.9 percentage points for workers over 40. The rise is mainly due to the sorting effect. The relative bargaining power advantage of men also rises with age but only by a percentage point or so.

Looking across education groups in the bottom rows of Table 4b, we see that the wage gap is roughly constant across education groups, but the average rents received by both men and women are increasing with years of schooling. This pattern implies that some of the "return to education" in Portugal is attributable to the impact of firm-specific pay premiums. Among men, for example, the average rent component rises from 11.5% to 25.9% between workers with less than high school education, and those with university education. This 14 point rise accounts for about one-sixth of the 95 log point wage gap between university educated men and those with less than high school education in the dual connected sample. Among women the rent component rises from 5.5% to 21.3%, again enough to account for about one-sixth of the overall wage gap between university-educated women and those with less than high school education.

As shown in column 4, the net effect of firm-specific pay premiums on the gender wage gap is about the same for workers with less than high school or high school education, but is somewhat smaller for university-educated workers. The smaller effect for the latter group is a result of a substantially smaller sorting effect for university educated women, coupled with a relatively large bargaining power effect. Indeed, the biggest impact of differential bargaining power between women and men across all the groups in Table 4b is for the university-educated group.

To summarize, we find evidence of substantial gender differences in firm wage premia, with most of the differences attributable to workers sorting, but a non-negligible component attributable to bargaining, particularly among more educated workers. As noted earlier, our normalization procedure is likely to understate (if mildly) the importance of bargaining effects by understating the magnitude of rents in the labor market. Therefore, our estimates are in keeping with recent popular discussion (e.g. Sandberg, 2013) in suggesting the bargaining factors may be quite important for understanding the success of high skilled women.

6.2 Gender or Occupation?

A well-known feature of labor markets is that women and men work in different occupations (see e.g. Petersen and Morgan, 1995, Manning and Swaffield, 2008, and Goldin, 2014 for recent analyses, and Cardoso, Guimarães and Portugal, 2012 for a discussion in the Portuguese context). Even within a given firm, the jobs held by men and women are often quite different. This raises the question of whether some of the differential bargaining power illustrated in Figure 5 is actually due to differences in the extent to which workers in different occupations receive different shares of any firm-wide rents.[32] To address this question, we fit separate fixed effect models for male and female workers who work in mainly male or mainly female occupations, allowing unrestricted firm effects for each group. We then investigated whether there is a systematic difference in bargaining power between men and women who work in similar occupations, or if the differences we see in Figure 5 are really due to differences in bargaining power between occupation groups.

We began by calculating the fraction of female workers in each of the 110 3-digit occupations identified in the QP data set (based on person-year level data). Appendix Figure A7 shows the histogram of female shares across occupations. There are spikes in the distribution of female shares at close to 0 (mainly for construction and related occupations), close to 1 (mainly for personal service occupations), and around 60-70% female (for retail sales and restaurant workers). Next, we assigned each individual the average fraction of female workers in his or her current occupation in that year. Finally, we classified workers as having a "mainly female" or "mainly male" occupation depending on whether the average fraction of female workers in his or her occupation(s) is above or below the median across all occupations. This procedure classifies 83% of women and 27% of men as having mainly female occupations, and a complementary 17% of women and 73% of men as having mainly male occupations.

Columns 7-10 of Table 1 display the characteristics of the four resulting groups, limiting attention to the subsets of each group who are also in the dual connected set. Men with mainly male occupations tend to be slightly older but are less-educated than men with mainly female occupations. Despite their lower education, they also earn slightly more (+2%) than than those in mainly female occupations. In contrast, women in mainly male occupations have *higher* education than those in mainly female occupations, and earn substantially more (+20%) than those in mainly female occupations. Looking within broad occupation groups, the gender gap in wages is relatively small (8%) for workers with mainly male occupations, but is relatively large (26%) for workers in mainly female occupations.

The results from estimating equation 4 on the four groups are summarized in columns 3-6 of Table 3. As in our main analysis, we estimate the model on the set of worker and firm observations in the largest connected set for each group (without limiting attention to workers at dual-connected firms). The estimated models are similar to the models estimated for the overall samples of men and women, and yield broadly similar conclusions about the relative importance of worker and firm effects in the overall variation of wages.

Table 4c summarizes the results of our comparisons between men and women with mainly female and mainly male occupations. For reference, Row a of the Table reproduces the results from Table 4b for all men and women at dual connected firms. Row b of the Table conducts a comparison between men and women with mainly female occupations who work at firms that employ both groups. As expected given the comparisons in Table 1, the gender wage gap in this subsample is a little larger

[32] Ransom and Oaxaca (2005) present a case study of occupational segregation within a chain of U.S. grocery stores that led to substantially lower wages for female employees at the firm.

than in the overall dual connected set. The total contribution of firm-specific pay premiums to the gender gap (in column 4) is about two-thirds as large as in the overall sample, reflecting a reduction in the sorting component and especially the relative bargaining power component, which is actually "wrong signed" when evaluated using the distribution of male workers across firms. These results suggest that women in mainly female occupations receive about the same firm-specific pay premiums as their male colleagues in the same occupations. Both genders in female occupations, however, receive smaller pay premiums than men in mainly male occupations. Consequently, firm-specific pay premiums still contribute to the gender wage gap.

Row c presents the comparison between men and women with mainly male occupations who work at firms that employ both groups. Consistent with the patterns noted in our discussion of Table 1, the gender wage gap in this subsample is relatively small (13.7%). Nevertheless, a relatively large share (32%) of the gap is attributable to firm-specific pay premiums, with one decomposition (based on equation 8) attributing 20% of the gap to sorting and 12% to bargaining, and the other (based on equation 9) attributing 11% of the gap to sorting and 21% to bargaining.

6.3 Firm Effects and Profitability

So far we have focused on the differences in the firm-specific premiums paid to men and women without distinguishing between the potential sources of these premiums. As is made clear by equation (13), ψ_j^M and ψ_j^F contain components that are directly related to observable proxies for firm-specific rents as well as residual components (ν_j^M and ν_j^F) that are potentially attributable to other factors. In this section we focus more narrowly on the parts of the firm effects that are directly related to observable productivity, and ask how these differ between men and women.

Table 5 summarizes a series of models based on equation (13). Row 1 of the table presents estimates of the gender specific rent-sharing coefficients π^M and π^F, as well their ratio, estimated across all firms in our dual-connected set with linked financial data. By construction, the estimates are identical to the estimates obtained from equation (19) at the optimized value for $\hat{\tau}$. To estimate the sampling error for the estimated ratio, we note that $\hat{\pi}^F/\hat{\pi}^M$ is the two-stage least squares estimate of the parameter δ_1 from a simple model of the form:

$$\hat{\psi}^F_{J(i,t)} = \delta_0 + \delta_1 \hat{\psi}^M_{J(i,t)} + e_{J(i,t)},$$

using excess value-added per worker as an instrumental variable for $\hat{\psi}_j^M$. We therefore use the conventional standard error of the two-stage least squares estimator as our estimated standard error for the ratio. As shown in column 6 of the table, the estimated ratio is 0.88 with a standard error of 0.03. We can therefore rule out the null hypothesis of equal rent-sharing ($\pi^F = \pi^M$) in favor of the alternative that women receive a smaller share of the component of firm-wide rents that is directly related to excess value-added.

In row 2 we fit a parallel set of models for the estimated firm effects of different employee groups at firms that are included in the connected set for female workers in female occupations. As shown in columns 2 and 3, the estimated rent-sharing coefficients for all male and all female employees at these firms are very similar to the coefficients obtained over the full dual connected set (in row 1), and their ratio is identical to the ratio in row 1. Column 4 presents the coefficient estimate relating the firm effects for women in mainly female occupations to value-added at the firm. Interestingly, this is very similar to the coefficient for all female workers, and is about 88% as large as the corresponding coefficient for men (column 7).

Finally, in row 3 we fit models for the estimated firm effects of different employee groups at firms that are included in the connected set for female workers in male occupations. The estimated rent-sharing coefficients for all male workers and all female workers at this subsample of firms are a little smaller than the corresponding coefficients in row 1, and the ratio of the female to male coefficients is slightly larger (0.92 versus 0.88), but within a standard error of the ratio for the overall sample. The estimated rent-sharing coefficient for female workers in mainly male occupations (column 5) is nearly the same as the coefficient for all female workers, and is 93.3% as large as the coefficient for males at these firms (column 8), though given the sampling errors we cannot reject the null that these workers have the same rent-sharing parameter as all male workers.

To probe the robustness of the results in Table 5 we re-estimated the models including controls for industry (20 dummies), location (dummies for firms located in Lisbon or Oporto) and a quadratic in firm size (based on average total employment in all years). Estimates from these models are presented in Appendix Table A4. In brief, the addition of controls leads to a slight attenuation (on the order of 10-15%) in the estimated rent-sharing coefficients, with a slightly bigger attenuation of the coefficients for women than men. These models therefore reinforce our conclusion that women get a smaller share of rents than men, though in all cases the estimated ratios are within a standard error of the ratios in Table 5.

Given estimates of the rent-sharing coefficients π^F and π^M and the means of excess average value-added across jobs, we can estimate

$$E[\pi^M \overline{EVA}_{J(i,t)}|G(i) = M] - E[\pi^F \overline{EVA}_{J(i,t)}|G(i) = F].$$

This expression gives the contribution to the gender wage gap of the component of firm-specific pay premiums that is correlated with measured productivity. We can also decompose this contribution into bargaining and sorting channels using equations (14) and (15). In Appendix Table A5 we present estimates of the terms in these two alternative decompositions, using the overall sample of dual-connected workers. We find that the productivity-related component of the male and female firm effects account for about 80% of our overall estimates of the size of the rent premiums received by both men and women, and 80% of the impact of firm-specific premiums on the gender wage gap (see row 5 of the Table). Since excess value-added per worker is a noisy measure of the rents available at a given firm, we suspect that 80% should be interpreted as a lower bound on the share of productivity-related factors that drive the variation in firm-specific pay premiums. Applying the decompositions in equations (14) and (15) we find that differential sorting of men to high productivity firms accounts for about two-thirds of the total effect of productivity-related factors, while the lower bargaining power of women accounts for about one-third, or 1-1.5 log points of the overall gender gap.

7 Within Firm Changes in Profitability and Wages

Our analysis so far has focused on the gender-specific firm effects in equation (4) that capture *between-firm* differences in the average wages paid to men and women. As noted in Section 3.5, however, a rent-sharing model with differential bargaining power also has implications for the *within-firm* evolution of wages of male and female workers as firms experience changes in product market conditions, technology, or other factors over time. In this section we use observations from the last 4 years of our analysis sample (2006-2009) to measure the effects of changes in firm-specific profitability on the wages of male and female job stayers. In addition to exploiting a different source of variation, this analysis has the advantage of not relying on the exogenous mobility assumptions underlying the worker-firm decomposition of wages.

We begin in Table 6 with an overview of the subsamples of workers from our dual connected sample who are observed working at the same firm between 2006 and 2009, and whose employer has financial information available for each year from SABI. We focus on three specific subsamples: all workers that meet these criteria (columns 1-2); workers at firms that have some female stayers with mainly female occupations (columns 3-4); and workers at firms with some female stayers with mainly male occupations (columns 5-6). We show mean age and education of the men and women in each subsample, mean firm size and the mean fraction of female workers, mean log wages in 2006 and 2009, and the mean value of excess value-added per worker (as defined in equation 16) for employers in 2006 and 2009

The overall sample of stayers in columns 1-2 contains 280,000 men and 200,000 women employed at 33,000 firms. The men and women in this sample have about the same age, education and wages as men and women in our overall analysis sample. The gender wage gap in the sample of stayers is also very close to the gap in our dual connected sample, with a value of 22 log points in both 2006 and 2009. Real wages of both male and female stayers rise by 8 log points over the 3 year period, while excess real value-added per worker is stable. The subsample of stayers at firms with at least one female worker in a mainly female occupation (columns 3-4) includes workers at 29,000 firms (87% of the firms included in the overall stayer sample). Over 90% of the men in the overall sample of stayers are included in this subsample so the characteristics of the males in column 3 are quite similar to those in column 1. Most of the female workers at these firms are in mainly female occupations, and for simplicity we show only the characteristics of these women in column 4, which are again quite similar to the characteristics of the overall sample of female stayers. The subset of stayers at firms with at least one female in a mainly male occupation (columns 5-6) is more selective, and includes only about one-third of firms, though together these firms account for about two-thirds of all male stayers. These firms have somewhat higher excess log value-added per worker (e.g., 0.84 vs. 0.76 in 2006) and their male employees have somewhat higher wages (e.g., about 4% higher in 2006) than in the overall sample of male stayers. Female stayers with mainly male occupations have higher wages than other female workers but their wage growth is comparable to other female stayers.

Figure 6 illustrates the relationship between the changes in excess value-added between 2006 and 2009 and the corresponding changes in wages of male and female job stayers, using our broadest sample of stayers. To reduce the impact of measurement errors in the change in value-added we group firms into 20 roughly equally sized groups based on the change in excess log value-added per worker, and construct the mean log wage changes of the job stayers for each group. As shown in the graph, mean wage changes are strongly correlated with changes in excess value-added (the correlation is 0.68 for both men and women). Moreover, the relationship is noticeably flatter for women, suggesting a smaller degree of rent-sharing for female employees than males.[33]

Table 7 presents a series of models based on equation (17) that show the relationship between the change in excess value-added at a firm and the wage changes of male and female stayers. As discussed in Section 3.5, we estimate these models in two steps, first regressing wage changes over the 2006-2009 period on a quadratic in age (separately by gender), then in a second stage regressing the average regression-adjusted wage changes at each firm on the change in excess value-added at the firm, weighting by the number of workers at the firm in the gender group. Given the large variability in measured value-added, we "Winsorize" the change in excess value-added at +/- 0.50. (A parallel set of models, estimated without Winsorizing, is presented in Appendix Table A6). Columns 1 and 3 of the table show the resulting OLS estimates for all male stayers and all female stayers. For the overall sample of stayers (row 1) these estimates are 0.049 and 0.045, respectively. A concern with these OLS estimates is attenuation bias, caused by measurement error in the change in excess value-added per worker. Assuming that the measurement error is uncorrelated with other firm-specific determinants of wage changes, however, any attenuation bias will affect the male and female coefficients proportionately, so their ratio is an unbiased estimate of the relative bargaining power ratio. We show this ratio in column 5. Consistent with the pattern in Figure 5, and the estimates in Table 5, the ratio is 0.90, though the standard error of the estimate is relatively large.[34]

While the OLS estimates can be used to infer the relative bargaining power of female workers, the magnitudes of the rent-sharing coefficients are also of interest, so we show IV estimates of these coefficients, using the Winsorized change in excess value-added per worker from 2007 to 2008 as an instrument for the change from 2006 to 2009. Under the (strong) assumption that the measurement error in excess value-added in each year is serially uncorrelated, this procedure will provide consistent estimates of the rent-sharing coefficients. More realistically, with some serial correlation in the measurement errors this procedure will yield estimates that are still biased toward zero, but less so than the OLS estimates. As shown by the first-stage F-statistics reported in columns 2 and 4, the change in excess value-added from 2007 to 2008 is a powerful predictor of the change over the longer horizon, with F-statistics of around 240 (clustered by firm). The resulting IV estimates, presented in columns 2 and 4, yield a rent-sharing elasticity for men of 0.092 and for women of 0.091.

Compared to the rent-sharing coefficients from the between-firm analysis in row 1 of Table 5, even the IV estimates of the rent-sharing effects for stayers are noticeably smaller. For example, the estimated between-firm rent-sharing coefficient for men is 0.156 while the estimated within-firm

[33] A regression of the change in wages of stayers on the change in excess value-added has a coefficient of 0.071 for men (standard error 0.014) and a coefficient of 0.044 for women (standard error 0.01), excluding the top and bottom bins.

[34] We estimate this standard error from an IV regression of the firm-specific average wage change for female stayers on the corresponding average wage change for male stayers, using the change in excess value-added per worker over the period 2006-2009 as an instrument.

coefficient is 0.092. There are three plausible explanations for the discrepancy. First, we suspect that our IV procedure provides an incomplete solution to the problem of measurement error in the within-firm estimates. Second, contrary to equation (1), it may be that wages are less responsive to transitory fluctuations in rents than to permanent differences.[35] Finally, institutional features like multi-year collective bargaining contracts and lags in the individual negotiation process may slow the adjustment of wages to movements in profitability.

In row 2 of Table 7 we present models for the subset of workers employed at firms with some females in mainly female occupations. The estimates for all male and female workers at these firms (columns 1-5) are very similar to the estimates for the overall sample of stayers. The estimates for females in mainly female occupations (columns 6-8) are also very similar to those for all females at these firms, and show a relative bargaining power ratio for these women of about 0.9. Finally, in row 3 we present models for the workers employed at firms with some females in mainly male occupations. Again, the estimates for all males and females at these firms are close to the estimates for the overall sample in row 1. For women in mainly male occupations the estimated coefficients are also relatively close to the estimates for other groups of women. The estimate of their relative bargaining power (column 11) is actually larger than 1, but quite imprecise.

Overall the estimates in Table 7, while limited by the relatively short sample period over which we can observe job stayers, are supportive of the hypothesis that female workers gain less than their male co-workers when their employer becomes more profitable. Indeed, our estimate of the ratio γ^F/γ^M is centered around 0.9 for all women and women in mainly female occupations, quite close to the estimated ratio from our previous designs.

We have estimated a variety of additional models for other subgroups of male and female stayers, including workers in larger and smaller firms, workers in firms with larger and smaller fractions of female employees, and workers in firms with higher and lower within-firm wage inequality. Unfortunately, as suggested by the standard errors for the estimated ratios in Table 7, our ability to precisely estimate the relative bargaining power of women is limited, and none of the estimates of the relative ratio of female to male bargaining power are significantly different from 0.9 – the average ratio across firms.

8 Summary and Conclusions

A growing body of research argues that firm-specific wage premiums are a pervasive and economically important feature of labor market earnings. These premiums will contribute to the gender wage gap if women tend to work at firms that offer smaller premiums, or if female employees tend to earn smaller premiums than their male colleagues at higher-paying firms. Previous literature suggests that both channels may be at work in modern labor markets. On one hand, there is evidence that women gain less from inter-firm mobility than men (e.g., Loprest 1992, Hospido 2009, Del Bono and Vuri, 2011). On the other, a variety of lab- and field-based studies suggest that women are less likely to initiate

[35] Guiso et al. (2005) analyze the relationship between wages and firm profitability using Social Security earnings record for Italian workers, and find smaller impacts of short run changes of profitability than of longer-run changes.

bargaining with their employers than men, and gain less when they do bargain (see Bertrand, 2011 for a recent survey).

We have developed a simple way of jointly assessing both the sorting and bargaining mechanisms using matched employer-employee data that combines wage outcomes for workers at different firms with firm-level financial information. In essence, our approach is to estimate the full set of firm-specific wage premiums earned by men and women, then conduct a standard decomposition exercise that evaluates the effect on the gender wage gap if women received the same premiums as men at each firm (or vice versa), and alternatively if women had the same distribution across firms as men (or vice versa). We resolve the normalization problem that arises in a decomposition with dummy explanatory variables by using firm-level value-added data to define the reference group of firms that pay zero wage premiums. Our approach provides a comprehensive summary of the potential role of firm-specific pay premiums in mediating the gender wage gap.

Our wage models build on the two-way fixed effects model of Abowd, Kramarz and Margolis (1999), extended to allow for *gender-specific* firm effects. We show that the identifying assumptions required for OLS estimates of this model to recover meaningful parameter estimates are approximately satisfied in the Portuguese labor market. We also show that the wage premiums paid to men and women at a given firm are highly correlated with each other, and with a simple measure of firm-specific productivity based on value-added per worker.

We find that female employees receive about 90% of the wage premiums that men earn at more profitable firms. We also find that women are disproportionately likely to work at less productive firms paying lower premiums to both genders. Overall, we conclude that the sorting and bargaining channels explain about 20% of the gender wage gap in Portugal. Roughly two-thirds of this 20% is explained by sorting and one-third by the shortfall in relative bargaining power. As a secondary check on the relative bargaining channel, we examine the impacts of changes in firm-specific productivity on changes in the wages of males and females who remain with the firm over time. Again, we find that women receive about 90% of the wage increases enjoyed by their male co-workers.

An important question for future research is the extent to which differences in the average wage premiums paid to men and women by the same firm are due to differences in the behavior of female employees, versus differences in how employers interact with their female workers. While some of the difference may be due to a preference among women to avoid bargaining (e.g., the "nice girls don't ask" hypothesis of Babcock and Laschever, 2003), an alternative explanation is that employers treat women differently in the wage setting process. Such differential behavior is implied by monopsonist wage setting models in which firms set wages taking account of the elasticity of supply of men and women to the firm (see Manning, 2003 and Barth and Dale-Olsen, 2009). It seems plausible that the treatment of women in the wage-setting process would vary substantially from firm to firm, which may explain the emphasis in the U.S. on crafting legislation prohibiting unequal treatment *within a firm* by gender or race.

Another open question is how to explain the under-representation of women at high-wage firms. Evidence from the de-regulation of the U.S. banking sector (Ashenfelter and Hannan, 1986; Black and Strahan, 2001) suggests that firms with greater market power exercise greater discrimination against women. If this is the case, policies aimed at discrimination in hiring and firing may play an important role in mitigating some of the gender wage gap. On the supply side, men and women may place different values on non-wage aspects of a job (like proximity to home), leading to differential sorting on the wage component. Our analysis shows that the degree of differential sorting rises with age, contributing to the systematic life cycle pattern in the gender wage gap. More research on the dynamic patterns of sorting with comprehensive worker and firm data could further illuminate this channel.

Finally, our approach to combining worker-firm fixed effects models with Oaxaca-style decompositions into sorting and bargaining components is potentially applicable to other classic wage gaps in labor economics including the black/white wage gap, the immigrant/native wage gap, the experience profile of wages, and wage gaps based on various measures of intelligence – explanations for which have traditionally relied upon a market based perspective.[36] The proliferation of rich employer-employee datasets offers the opportunity to determine the extent to which these heavily studied sources of wage inequality are in fact mediated by heterogeneity across firms.

[36]See Hakanson, Lindqvist and Vlachos (2013) for a recent analysis of IQ segregation across firms, and Damas de Matos (2011) for an analysis of the impact of firms on the immigrant assimilation profile.

References

Abowd, John, Robert Creecy, and Francis Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Cornell University Department of Economics Unpublished Working Paper, March 2002.

Abowd, John, and Francis Kramarz. 1999. "The Analysis of Labor Markets Using Matched Employer-Employee Data." In Orley Ashenfelter and David Card (editors), *Handbook of Labor Economics* Vol. IIIc. Amsterdam: Elsevier, pp. 3143-3259.

Abowd, John, Francis Kramarz and David Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67(2): 251-333.

Altonji, Joseph and Rebecca Blank. 1999. "Race and Gender in the Labor Market." In Orley Ashenfelter and David Card (editors), *Handbook of Labor Economics* Vol. IIIc. Amsterdam: Elsevier, pp. 3143-3259.

Andrews, M.J., L. Gill, T. Schank and R. Upward. 2008. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? *Journal of the Royal Statistical Society* (Series A) 171 (3): 673-697.

Ashenfelter, Orley and Timothy Hannan. 1986. "Sex Discrimination and Product Market Competition: The Case of the Banking Industry." *Quarterly Journal of Economics* 101(1): 149-174.

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

Babcock, Linda, Michele Gelfand, Deborah Small and Heidi Stayn. 2006. "Gender Differences in the Propensity to Initiate Negotiations." In De Cremer, D., M. Zeelenberg, and J.K. Murnighan (editors), *Social Pyschology and Economics*. Mahwah NJ: Erlbaum Press.

Bagger, Jesper, Kenneth Sorensen and Rune Vejelin. 2012. "Wage Sorting Trends." School of Economics and Management University of Aarhus Working Paper 2012-17.

Barth, Erling and Harald Dale-Olsen. 2009. "Monopsonistic Discrimination, Worker Turnover, and the Gender Wage Gap." IZA DP No. 3930. January.

Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press. (second edition 1971).

Bertrand, Marianne. 2011. "New Perspectives on Gender." In Orley Ashenfelter and David Card (editors) *Handbook of Labor Economics*, Vol. 4b. Amsterdam: Elsevier. pp. 1543-1590.

Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94(4): 991–1013.

Bertrand, Marianne, Claudia Goldin and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2(1): 228-255.

Black, Sandra E. and Philip E. Strahan. 2001. "The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry." *American Economic Review* 91(4): 814-831.

Blau, Francine D. and Lawrence M. Kahn. 2000. "Gender Differences in Pay." *Journal of Economic*

Perspectives 14(4): 75-99.

Bowles, Hannah R., Linda Babcock and Kathleen L. McGinn. 2005. "Constraints and Triggers: Situational Mechanics of Gender in Negotiations." J*ournal of Personality and Social Psychology* 89: 951-965.

Bowles, Hannah R., Linda Babcock and Lei Lai. 2007. "Social Incentives for Sex Differences in the Propensity to Initiate Bargaining: Sometimes It Does Hurt to Ask." *Organizational Behavior and Human Decision Processes* 103: 243-273.

Burdett, Kenneth, and Dale T. Mortensen. 1998. "Equilibrium Wage Differentials and Employer Size." *International Economic Review* 39: 257-73.

Buser, Thomas, Muriel Niederle and Hessel Oosterbeek. 2013. "Gender, Competitiveness, and Career Choices." *Quarterly Journal of Economics*, forthcoming.

Card, David, Francesco Devicienti and Agata Maida. 2014. "Rent-Sharing, Holdup, and Wages: Evidence from Matched Panel Data." *Review of Economic Studies*, forthcoming.

Card, David, Joerg Heining and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128(3): 967-1015.

Cardoso, Ana Rute, Paulo Guimarães and Pedro Portugal. 2012. "Everything You Always Wanted to Know About Sex Discrimination." IZA Working Paper 7109.

Cardoso, Ana Rute and Pedro Portugal. 2005. "Contractual Wages and the Wage Cushion under Different Bargaining Settings." *Journal of Labor Economics* 23(4): 875-902.

Carlsson, Mikael, Julian Messina and Oskar Nordström Skans. 2011. "Wage Adjustment and Productivity Shocks." Sveriges Riksbank Working Paper #253.

Charles, Kerwin K. and John Guryan. 2008. "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination." *Journal of Political Economy* 116(5): 773-809.

Charles, Kerwin K. and John Guryan. 2011. "Studying Discrimination: Fundamental Challenges and Recent Progress." *Annual Review of Economics* 3(1), 479-511.

Coles, Melvyn and Dale T. Mortensen. 2012. "Equilibrium Labor Turnover, Firm Growth and Unemployment." NBER Working Paper No. 18022.

Damos de Matos, Ana, 2011. "The Careers of Immigrants." London School of Economics Unpublished Working Paper.

Del Bono, Emilia and Daniela Vuri. 2011. "Job mobility and the Gender Wage Gap in Italy," *Labour Economics* 18(1): 130-142,

Eeckhout, Jan and Philipp Kirchner. 2011. "Identifying Sorting – In Theory." *Review of Economic Studies* 78 (3): 872-906.

EPI (Economic Policy Institute). 2010. *The State of Working America* 2008-10. Washington DC: EPI.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." In Orley Ashenfelter and David Card (editors) *Handbook of Labor Economics*, volume 4a. Amsterdam: Elsevier. pp. 1-102.

Freeman, Richard B. and Jonathan Leonard. 1987. "Union Maids: Unions and the Female Workforce." In Clair Brown and Joseph Pechman, editors. *Gender in the Workplace*. Washington, DC:

Brookings Institution.

Giuliano, Laura, David I. Levine, and Jonathan Leonard. 2009. "Manager Race and the Race of New Hires." *Journal of Labor Economics* 27(4): 589–631.

Gneezy, Uri, Muriel Niederle and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* 118(3): 1049-1074.

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

Goldin, Claudia and Cecilia E. Rouse. 2000. "Orchestrating Impartiality: The Impact Of 'Blind' Auditions On Female Musicians." *American Economic Review* 90(4): 715-741.

Groshen, Erica L. 1991. "The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, orWhere You Work?" *Journal of Human Resources* 26(3): 457-472.

Guiso, Luigi, Luigi Pistaferri and Fabiano Schivardi. 2005. "Insurance within the Firm." J*ournal of Political Economy* 113(5): 1054-1087.

Hall, Robert E. and Alan B. Krueger. 2012. "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search." *American Economic Journal: Macroeconomics* 4(4): 56-67.

Hâkanson, Christina, Erik Lindqvist and Jonas Vlachos. 2013. "Firms and Skills: The Evolution of Worker Sorting." Chapter 3 in *Changes in Workplaces and Careers*, Christina Hâkanson, Doctoral Dissertation. Stockholm University.

Heckman. James J and Peter Siegelman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." In Michael Fix and Raymond Struyk, editors. *Clear and Convincing Evidence: Measurement of Discrimination in America.* Washington DC: Urban Institute Press.

Hospido, Laura. 2009. "Gender Differences in Wage Growth and Job Mobility of Young Workers in Spain." Investigaciones Economicas 33(1): 5-37.

ILO (International Labor Organization). 2011. Key Indicators of the Labour Market Database, 7th Edition. Available at www.ilo.org/empelm/pubs/WCMS_114060/lang–en/index.htm

INE (Statistics Portugal). 2012. *Estatísticas do Emprego.* 1.º Trimestre. Lisbon: Statistics Portugal.

Kramarz, Francis and Oskar Nordstrom Skans. 2013. "When Strong Ties are Strong: Networks and Youth Labor Market Entry." *Review of Economic Studies.* forthcoming.

Lang, Kevin and Jee-Yeon K. Lehmann. 2012. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature* 50(4): 959-1006.

Lentz, Rasmus and Dale Mortensen. 2010. "Labor Market Models of Worker and Firm Heterogeneity." *Annual Review of Economics* 2: 577–602.

Lopes de Melo, Rafael. 2009. "Sorting in the Labor Market: Theory and Measurement." Unpublished Working Paper (December).

Loprest, Pamela J. 1992. "Gender Differences in Wage Growth and Job Mobility." *American Economic Review* 82(2): 526-532.

Macis, Mario and Fabiano Schivardi. 2013. "Exports and Wages: Rent Sharing, Workforce Composition or Returns to Skills?" Johns Hopkins Department of Economics Unpublished Working Paper, August 2013.

Manning, Alan. 2003. *Monopsony in Motion. Imperfect Competition in Labor Markets.* Princeton NJ: Princeton University Press.

Manning, Alan and Joanna K. Swaffield. 2008 "The Gender Gap in Early-Career Wage Growth". *Economic Journal* 118(530): 983-1024.

Mare, David and Dean Hyslop. 2006. "Worker-Firm Heterogeneity and Matching: An Analysis Using Worker and Firm Fixed Effects Estimated from LEED." Statistics New Zealand Working Paper. Wellington, NZ: Statistics New Zealand.

Moscarini, Giuseppe and Fabien Postel-Vinay. 2013. "Stochastic Search Equilibrium." *Review of Economic Studies*, 80(4): 1545-1581.

Mumford, Karen and Peter N. Smith. 2008. "What Determines the Part-time and Gender Earnings Gaps in Britain: Evidence from the Workplace." *Oxford Economic Papers* 61: i56-i75.

Neumark, David, Roy J. Bank and Kyle D. Van Nort. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics* 111(3): 915-941.

Nekby, Lena. 2003. "Gender Differences in Rent Sharing and Its Implications for the Gender Wage Gap: Evidence from Sweden." *Economics Letters* 81: 403-410.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122(3): 1067–1101.

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

Oaxaca, Ronald L. and Michael R. Ransom. 1999. "Identification in Detailed Wage Decompositions." *Review of Economics and Statistics* 81(1): 154-157.

OECD, 2012. "OECD Family Database." Available at www.oecd.org/social/family/database.

Papps, Kerry L. 2012. "Spillovers and Wage Determination Within Firms." Department of Economics University of Oxford Unpublished Working Paper.

Petersen, Trond and Laurie A. Morgan. 1995. "Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap." *American Journal of Sociology* 101(2): 329-365.

Pissarides, Christopher and Jonathan Wadsworth. 1994. "On-the-job search: some empirical evidence from Britain." *European Economic Review*, 38(2), 385-401.

Postel-Vinay, Fabien and Jean-Marc Robin. 2006. "Microeconometric Search-Matching Models and Matched Employer-Employee Data." In R. Blundell, W. Newey and T. Persson, editors, *Advances in Economics and Econometrics, Theory and Applications.* Ninth World Congress, Volume 2. Cambridge: Cambridge University Press.

Ransom, Michael R. and Ronald L. Oaxaca. 2005. "Intrafirm Mobility and Sex Differences in Pay." *Industrial and Labor Relations Review* 58(2): 219-237.

Rigdon, Mary. 2012. "An Experimental Study of Gender Differences in Wage Negotiations." Rutgers University Center for Cognitive Science Unpublished Working Paper, November 2012.

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* London: Macmillan.

Royalty, Anne B. 1998. "Job-to-Job and Job-to-Nonemployment Turnover by Gender and Education Level." *Journal of Labor Economics* 16(2): 392-433.

Sandberg, Sheryl. 2013. *Lean In: Women, Work and the Will to Lead.* New York: Random

House.

Save-Soderbergh, Jenny. 2007. "Are Women Asking for Low Wages? Gender Differences in Wage Bargaining Strategies and Ensuing Bargaining Success." Swedish Institute for Social Research Working Paper 7/2007.

Skans, Oskar Nordstrom, Per-Anders Edin and Bertil Holmlund. 2008. "Wage Dispersion Between and Within Plants: Sweden 1985–2000." In Edward P. Lazear and Kathryn L. Shaw (editors). *The Structure of Wages: An International Comparison.* Chicago: University of Chicago Press.

Stuhlmacher, Alice F and Amy E. Walters. 1999. "Gender differences in negotiation outcome: A meta-analysis". *Personnel Psychology*, 52: 653– 677.

Wood, Robert G., Mary E. Corcoran and Paul N. Courant. 1993. "Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyers' Salaries." *Journal of Labor Economics* 11(2): 417-441.

Data Appendix

a. Quadros de Pessoal

The Quadros de Pessoal (QP) dataset for 2002-2009 includes over 20 million observations on 4.5 million workers. Individuals are identified over time with a unique person identifier. Firms are identified by a unique firm id. To construct our analysis sample we drop the entire history for a person if: (1) the hiring date for any job is missing or inconsistent across observations (0.6% of observations dropped); (2) the individual is observed in two consecutive years at different firms, but the hiring date for the second job is the same as the hiring date for the first job (6.9% of observations dropped); the hourly wage in any year is too high or too low (0.3% dropped); the change in the log hourly wage from one year to the next is less than -1 or greater than 1 (1.6% dropped). After these deletions we retain only person-year observations in which the worker is between the ages of 19 to 65 (1.6% of observations dropped), with at least two years of potential labor market experience (i.e., $age - education - 6 \geq 2$) (0.7% dropped) and is employed as a wage-earner (droppping 9.3% of observations). Appendix Table A1 shows the characteristics of the male and female observations in the entire QP, and our analysis sample. Overall the samples are quite similar in terms of age, education, location, mean hourly wage, and mean monthly hours of work.

b. SABI

Bureau van Dijk's SABI data base has annual data for non-financial firms including: a firm tax identifier; balance sheet information (with sales and the value of intermediate inputs); total employment; the firm's name, address, industry, shareholder capital; and date of formation. Data are available from 2000 onward, but coverage expanded substantially in 2005, and information on employment is missing for many firms prior to 2006.

c. Matching QP and SABI

The following variables are common to QP and SABI and can be used to match observations for a given firm in a given year in the two data sets: (1) location – zip code and county (*concelho*) in SABI, parish (*freguesia*) and county in QP; (2) 5 digit industry; (3) year of firm creation; (4) shareholder capital; (5) annual sales. We do not use employment in our matching procedure, but we use it as a check variable.

In QP, total sales in a given year are reported for the previous calendar year. We therefore use sales in year $t-1$ to match observations in year t. In SABI, both sales and shareholder capital are reported in thousands of euro, whereas in QP they are reported in euros. We therefore round both variables in QP to the thousands. Sales and shareholder capital are treated as missing if the reported values are zero. The zip codes reported in SABI were converted to parishes, with the exception of a few codes that cross parish boundaries and a few that appear to be non-existent codes.

We use a multi-step matching procedure which uses exact matching at each stage, and sequentially relaxes the number of variables that have to match exactly. Firms that are matched at one step are removed from both data sets, leaving unmatched observations for the next step. The steps are as follows:

1. Exact matching based on 5 variables: location, industry, year of firm creation, sales and shareholder capital. We first attempt an exact match using sales and shareholder capital for 2009 (the other variables are time-invariant), then work backwards to 2005. We initially use parish and 5-digit industry to look for exact matches. We then repeat the process using county and 3-digit industry.

2. Exact matching based on 4 variables: location, industry, and any two of: year of firm creation, annual sales, or shareholder capital. As in step 1, we initially use parish and 5-digit industry to look for exact matches, then use county and 3-digit industry.

3. Exact matching based on 3 variables: location, industry, and any one of: year of firm creation, annual sales, or shareholder capital. As in step 1, we initially use parish and 5-digit industry to look for exact matches, then use county and 3-digit industry. In this step, once a potential match was found, we compared data from QP and SABI to check the plausibility of the match. Specifically, we checked annual observations on sales and shareholder capital for 2005-2009. A match was validated only if the deviation between SABI and QP did not exceed 1% in any year for either sales or shareholder capital, or, in cases with a larger deviation in any one year, if the values in all other years were exactly the same in both data sets.

4. Exact matching based on 2 variables: location and any one of industry, year of firm creation, annual sales, or shareholder capital. As in step 3, potential matches were compared and only retained if the same criterion was met.

We matched a total of 301,417 firms between QP and SABI - representing about 80% of the firms that ever appear in SABI, and 53% of firms that appear at least once (with a worker in our analysis sample) in QP from 2002 to 2009. Of the matches, 52% were matched on all five variables, 31% were matched on four variables,12% were matched on three variables, and the remainder were matched on two varables. The match rate by firm size (based on average number of employees in QP) are as follows: 1-10 workers - 50.7%; 11-50 workers - 68.61%, 51-100 workers - 67.0%, 101-500 workers - 69.2%, over 500 workers - 61.0%.

Appendix Table A2 shows the match rates by major industry and by gender, calculated across person-year observation in our main QP analysis sample.

Figure 1: Trends in Real Hourly Wage of Men and Women

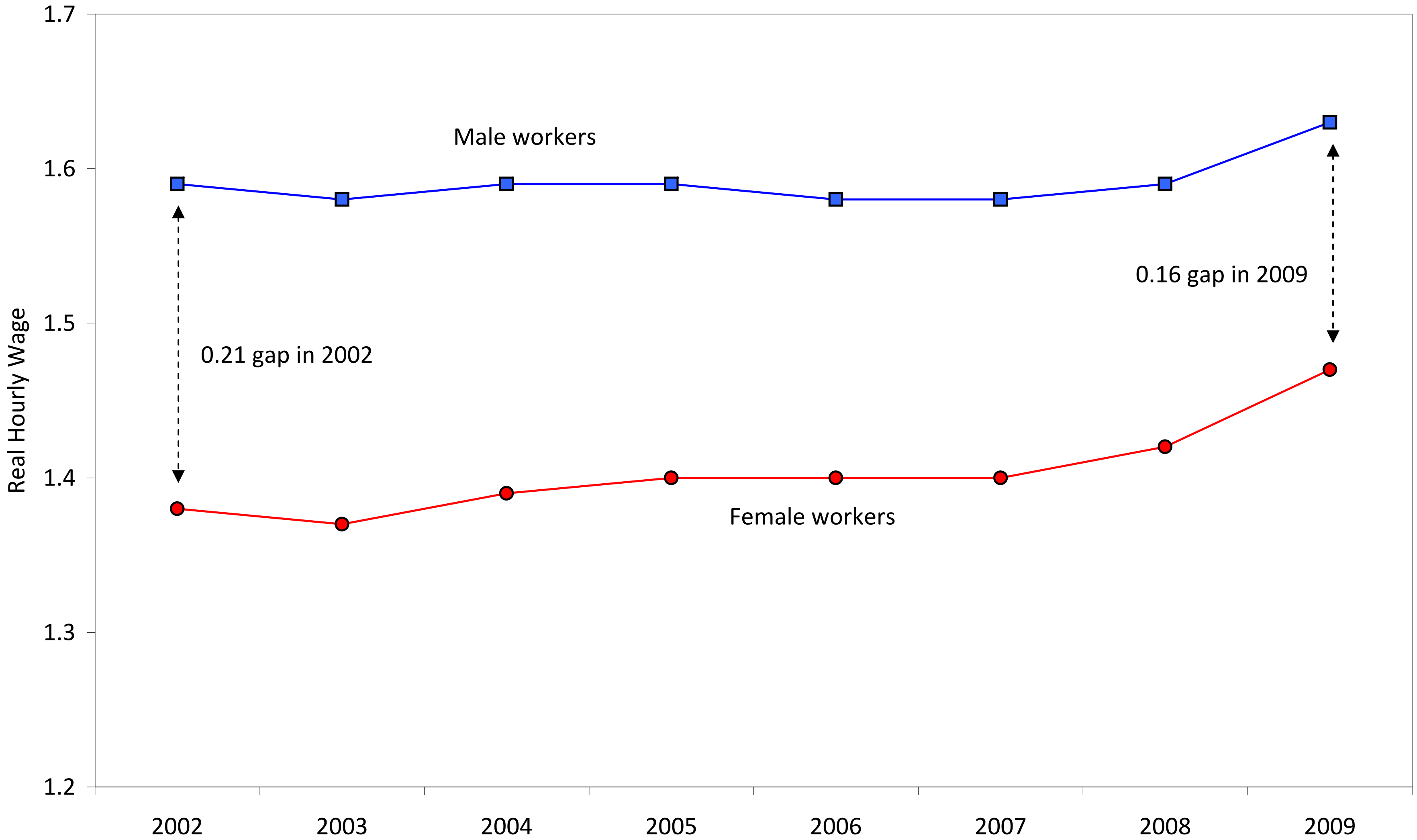
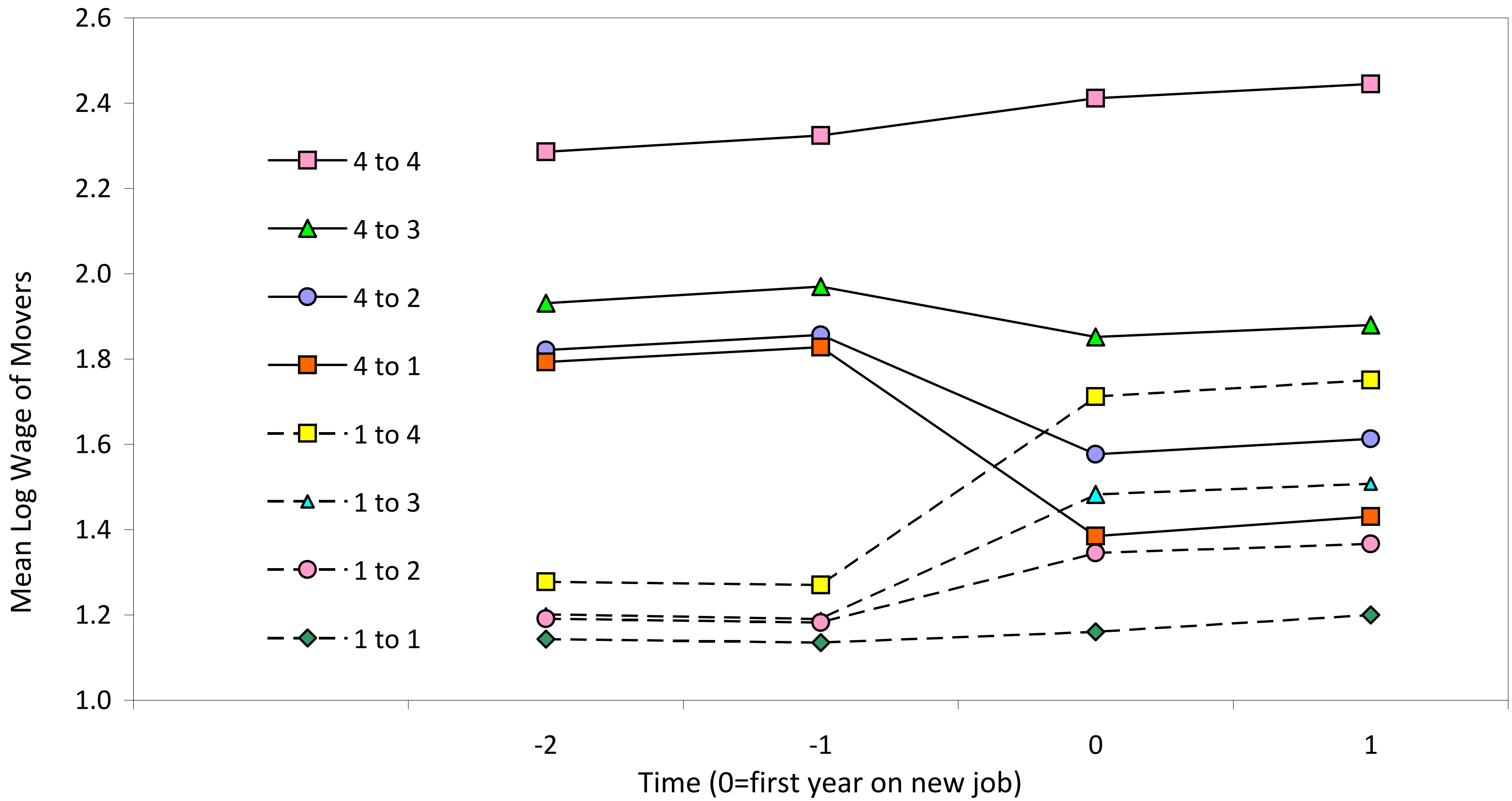
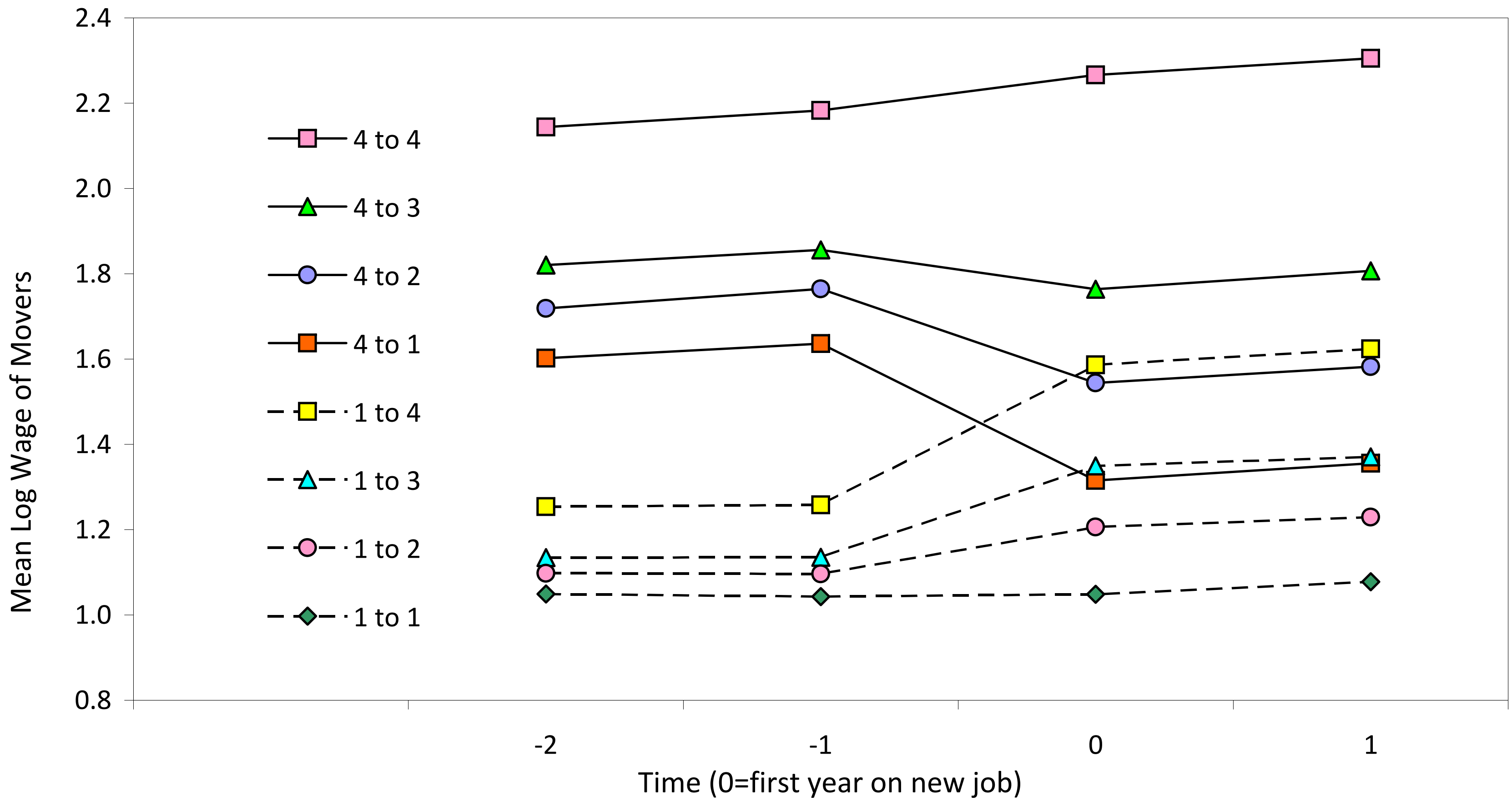


Figure 2a: Mean Wages of Male Job Changers, Classified by Quartile of Mean Co-Worker Wage at Origin and Destination Firm



Notes: figure shows mean wages of male workers at mixed-gender firms who changed jobs in 2004-2007 and held the preceding job for 2 or more years, and the new job for 2 or more years. Each job is classified into quartiles based on mean log wage of co-workers (quartiles are based on coworker wages in last year on old job and first year on new job).

Figure 2b: Mean Wages of Female Job Changers, Classified by Quartile of Mean Co-Worker Wage at Origin and Destination Firm



Notes: figure shows mean wages of female workers at mixed gender firms who changed jobs in 2004-2007 and held the preceding job for 2 or more years, and the new job for 2 or more years. Each job is classified into quartiles based on mean log wage of co-workers (quartiles are based on coworker wages in last year on old job and first year on new job).

Figure 3: Comparison of Adjusted Wage Changes of Male/Female Job Movers by Quartile of Coworker Wages of Origin and Destination Jobs

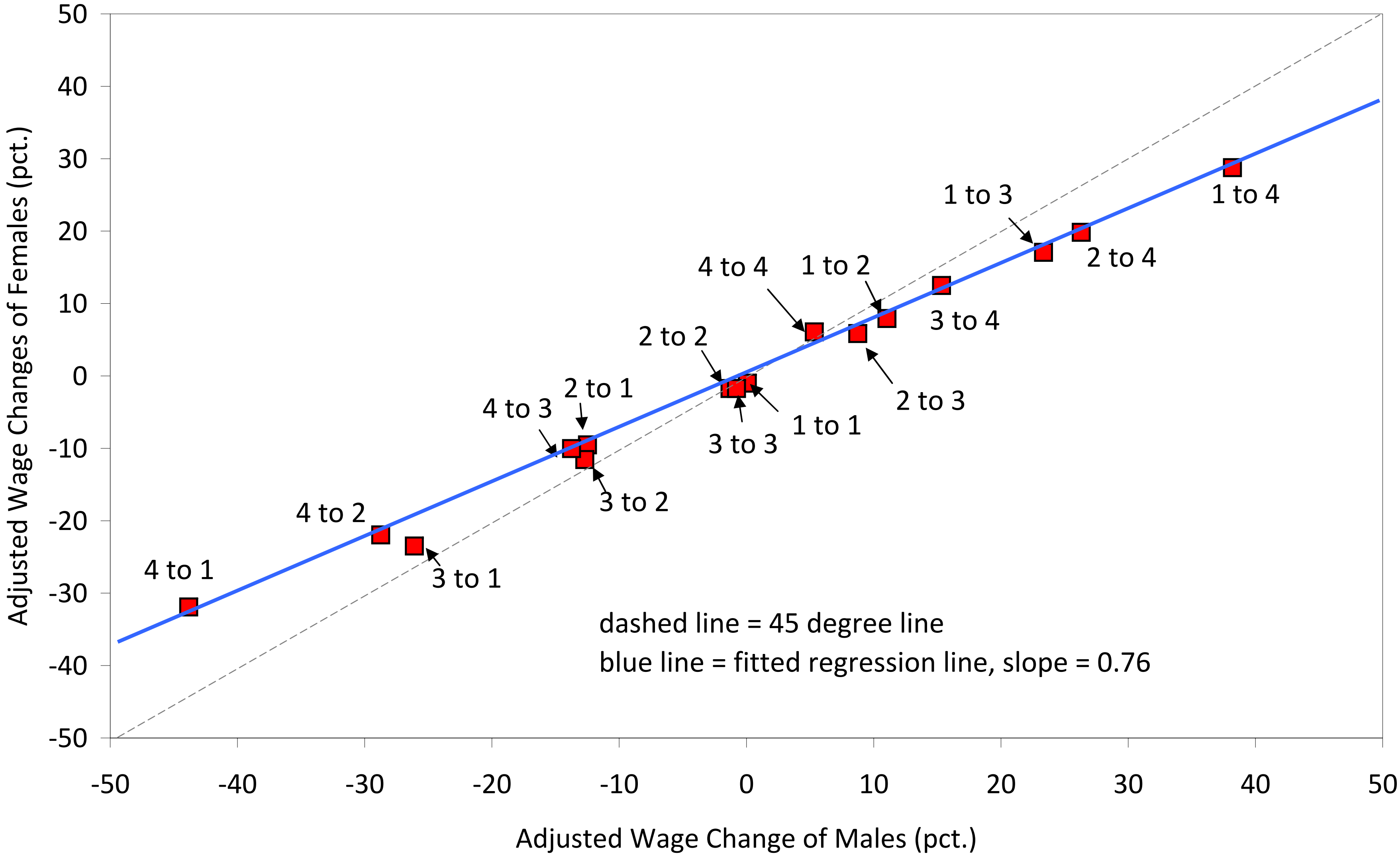


Figure 4: Firm Fixed Effects vs. Log Value Added/Worker

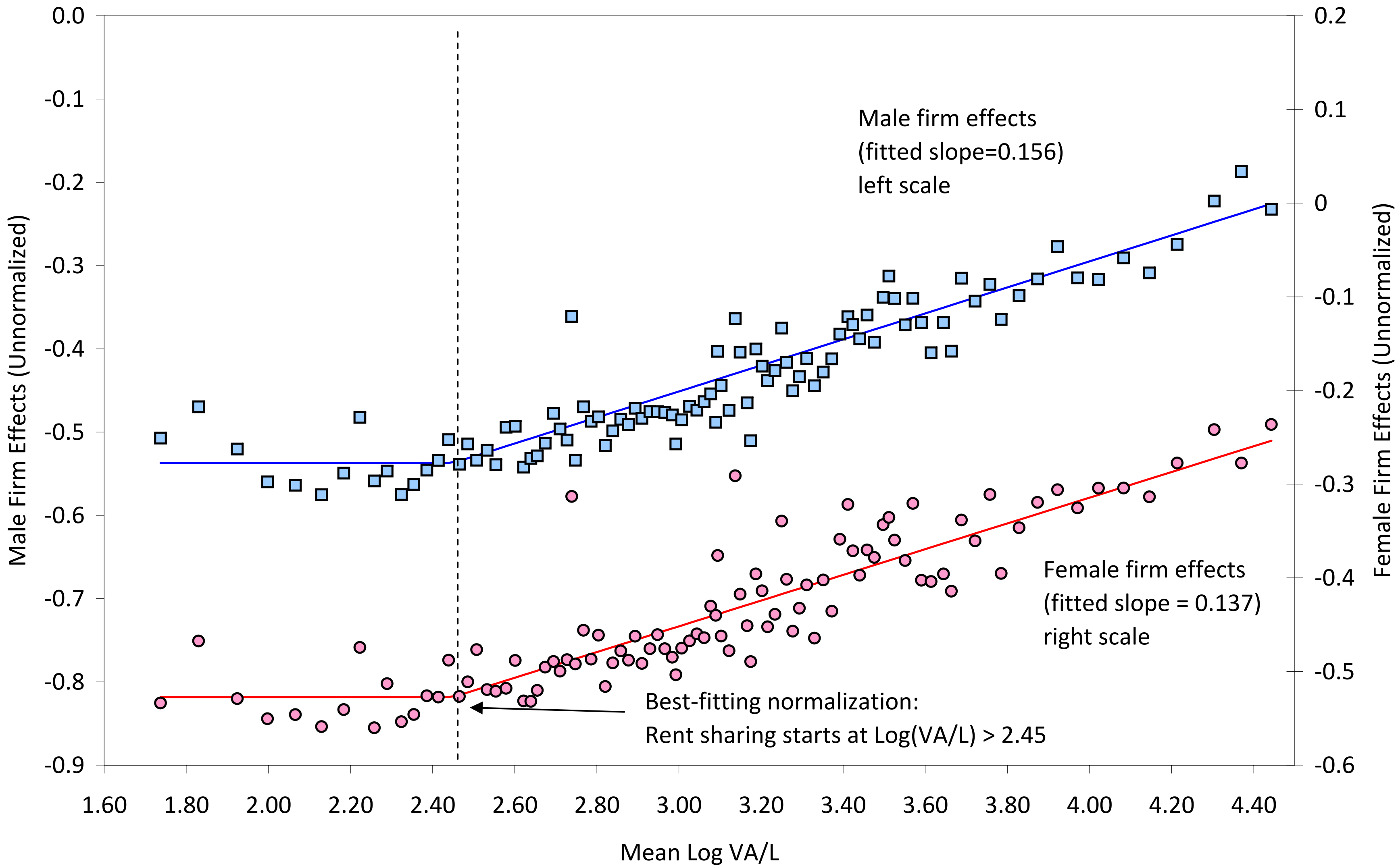


Figure 5: Estimated Firm Effects for Female and Male Workers:
Firm Groups Based on Mean Log VA/L

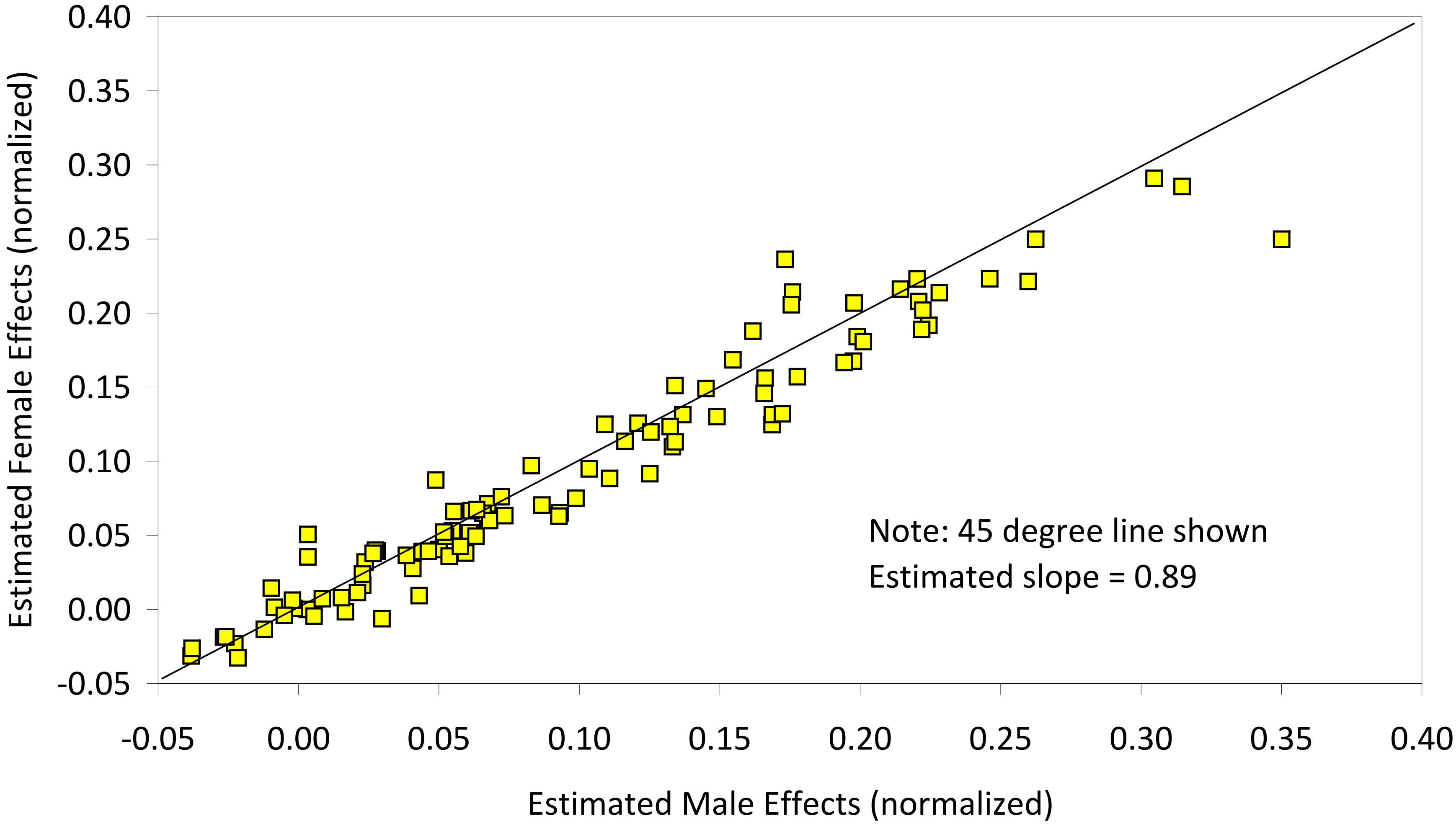
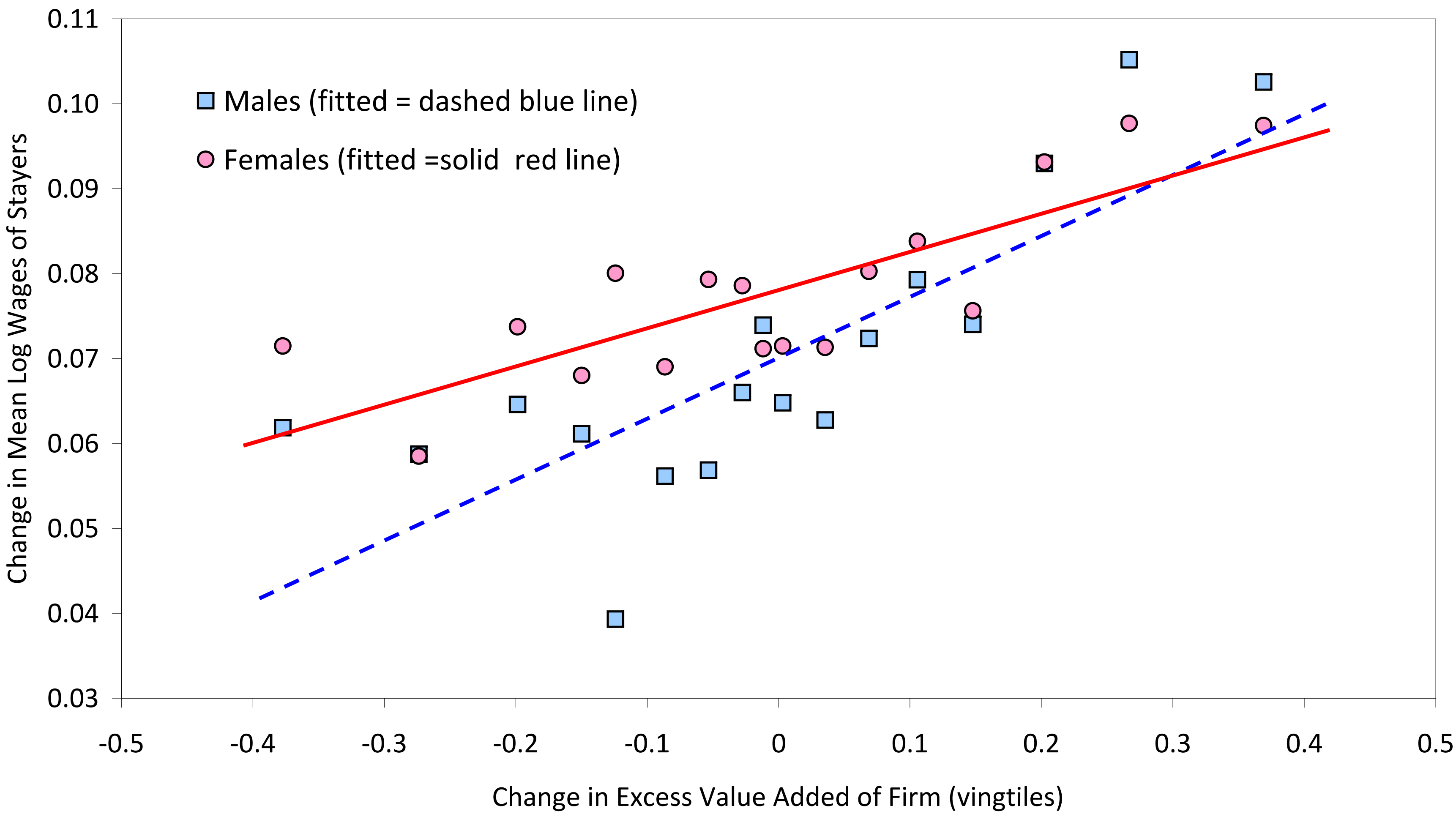


Figure 6: Changes in Excess Value Added and Changes in Wages of Stayers, 2006-2009



Note: Data for stayers are grouped into 20 cells based on changes in log value added per worker in excess of 2.45. Bottom and top vingtiles not shown.

Table 1: Descriptive Statistics for Various Samples of Employees in QP, 2002-2009

	Overall Population of QP Employees:		Connected Sets of Workers/Firms: All		Connected Sets of Workers/Firms: Dual-Connected		Dual Connected Males with:		Dual Connected Females with:		Overall Population with Av. Value Added Data	
	Males	Females	Males	Females	Males	Females	Male Occup's	Female Occup's	Male Occup's	Female Occup's	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Age:												
Mean Age	38.1	36.9	38.0	36.5	38.0	36.4	38.7	36.5	36.9	36.3	37.9	36.3
Fraction ≤ 30 years old	0.30	0.33	0.30	0.34	0.30	0.34	0.28	0.36	0.31	0.35	0.31	0.35
Fraction ≥ 50 years old	0.19	0.14	0.18	0.13	0.19	0.13	0.20	0.16	0.13	0.13	0.18	0.13
Education:												
Mean Years Schooling	8.0	8.8	8.0	8.9	8.6	9.1	8.2	9.5	9.6	9.0	7.8	8.6
Fraction with High School	0.18	0.23	0.18	0.23	0.21	0.24	0.18	0.28	0.22	0.25	0.17	0.24
Fraction with Degree	0.09	0.13	0.09	0.14	0.11	0.15	0.10	0.15	0.23	0.14	0.07	0.11
Mean Log Real Hourly Wage	1.59	1.41	1.62	1.43	1.71	1.48	1.72	1.70	1.64	1.44	1.54	1.36
(standard dev.)	(0.55)	(0.50)	(0.55)	(0.51)	(0.58)	(0.53)	(0.58)	(0.58)	(0.63)	(0.50)	(0.48)	(0.43)
Mean Monthly Hours	162.6	158.0	162.5	157.9	162.8	157.1	164.0	159.7	161.1	156.2	164.1	159.9
(standard dev.)	(24.7)	(30.1)	(24.8)	(29.9)	(24.0)	(30.5)	(22.5)	(27.2)	(25.9)	(31.5)	(23.6)	(29.4)
Fraction in Lisbon	0.35	0.35	0.36	0.37	0.42	0.40	0.42	0.43	0.42	0.40	0.32	0.34
Fraction in Oporto	0.13	0.13	0.13	0.13	0.13	0.13	0.12	0.14	0.11	0.14	0.13	0.14
Mean Firm Size (No. emp's)	730	858	804	978	1,091	1,230	1034	1234	1339	1205	500	886
Fraction Females at Firm	0.24	0.70	0.24	0.70	0.30	0.64	0.25	0.43	0.52	0.67	0.24	0.67
Mean Log VA/Worker											3.05	2.88
Number person-year obs.	9,070,492	7,226,310	8,225,752	6,334,039	6,012,521	5,012,736	4,295,906	1,716,615	949,869	4,062,867	5,786,148	4,204,851
Number of persons	2,119,687	1,747,492	1,889,366	1,505,517	1,450,288	1,247,503	1,026,308	423,980	223,896	1,023,607	1,441,626	1,082,058
Number of firms	349,692	336,239	216,459	185,086	84,720	84,720	70,614	60,061	42,527	80,477	153,998	141,889

Notes: Overall sample in columns 1-2 includes paid workers age 19-65 with potential experience ≥1. Individuals with inconsistent employment histories are excluded. Wages are measured in real (2009=100) Euros per hour. Value added is measured in thousands of real Euros per year. All statistics are calculated across person-year observations. See text for definitions of connected and dual connected sets and assignment of workers to occupations.

Table 2: Wages of Job Changes for Movers with 2+ Years of Data Before/After Job Change

Origin/ destination quartile	Number of Changes (1)	Percent of Changes (2)	Mean Log Real Wages of Movers: 2 years before (3)	1 year before (4)	1 year after (5)	2 years after (6)	3 Year Change (%) Raw (7)	Adjusted* (8)
Males								
1 to 1	13,787	43.2	1.14	1.14	1.16	1.20	5.6	0.1
1 to 2	9,139	28.7	1.19	1.18	1.35	1.37	17.6	11.0
1 to 3	6,283	19.7	1.20	1.19	1.48	1.51	30.6	23.4
1 to 4	2,682	8.4	1.28	1.27	1.71	1.75	47.3	38.2
2 to 1	7,293	21.2	1.34	1.35	1.22	1.27	-6.5	-12.5
2 to 2	12,326	35.8	1.37	1.38	1.40	1.42	5.0	-1.3
2 to 3	10,356	30.0	1.41	1.42	1.54	1.57	15.9	8.8
2 to 4	4,496	13.0	1.49	1.49	1.81	1.84	35.3	26.3
3 to 1	4,356	11.9	1.49	1.52	1.24	1.30	-19.4	-26.1
3 to 2	8,835	24.2	1.54	1.55	1.45	1.48	-5.8	-12.7
3 to 3	15,107	41.3	1.61	1.63	1.65	1.67	6.4	-0.8
3 to 4	8,246	22.6	1.73	1.75	1.94	1.97	24.7	15.3
4 to 1	1,634	5.4	1.79	1.83	1.39	1.43	-36.2	-43.8
4 to 2	3,245	10.7	1.82	1.86	1.58	1.61	-20.9	-28.7
4 to 3	6,589	21.7	1.93	1.97	1.85	1.88	-5.2	-13.7
4 to 4	18,830	62.1	2.29	2.32	2.41	2.45	15.9	5.3
Females								
1 to 1	24,130	60.9	1.05	1.04	1.05	1.08	2.9	-1.0
1 to 2	9,094	23.0	1.10	1.10	1.21	1.23	13.2	7.9
1 to 3	4,490	11.3	1.13	1.14	1.35	1.37	23.6	17.1
1 to 4	1,888	4.8	1.25	1.26	1.59	1.62	37.0	28.8
2 to 1	6,705	29.8	1.20	1.22	1.12	1.16	-4.5	-9.5
2 to 2	7,711	34.3	1.26	1.28	1.28	1.31	4.2	-1.7
2 to 3	5,495	24.5	1.33	1.35	1.44	1.46	12.6	5.8
2 to 4	2,562	11.4	1.44	1.45	1.69	1.73	29.0	19.8
3 to 1	3,283	16.7	1.38	1.40	1.15	1.20	-17.4	-23.5
3 to 2	4,762	24.2	1.42	1.45	1.34	1.37	-4.5	-11.6
3 to 3	7,245	36.8	1.51	1.53	1.54	1.56	5.3	-1.7
3 to 4	4,381	22.3	1.64	1.66	1.81	1.86	22.0	12.5
4 to 1	1,014	6.2	1.60	1.64	1.32	1.36	-24.6	-31.9
4 to 2	1,516	9.2	1.72	1.76	1.54	1.58	-13.7	-22.0
4 to 3	2,844	17.3	1.82	1.86	1.76	1.81	-1.3	-10.0
4 to 4	11,064	67.3	2.14	2.18	2.27	2.31	16.1	6.1

Notes: entries are mean log real daily wages for job changers to/from mixed-gender firms with at least 2 years of wages at the old job and the new job. Origin/destination quartiles are based on mean wages of coworkers in year before (origin) or year after (destination) job move.

*Four year wage change is regression-adjusted. Model includes dummies for age and education, and quadratic in age fully interacted with education. Models are fit by gender to job stayers.

Table 3: Summary of Estimated Two-way Fixed Effects Models for Male and Female Workers

	All Males	All Females	Wkrs. with Female Occ's. Males	Wkrs. with Female Occ's. Females	Wkrs. with Male Occ's. Males	Wkrs. with Male Occ's. Females
	(1)	(2)	(3)	(4)	(5)	(6)
Standard deviation of log wages	0.554	0.513	0.578	0.480	0.548	0.634
Number of person-year observations	8,225,752	6,334,039	1,766,441	5,220,256	6,160,204	861,451
Summary of Parameter Estimates:						
Number person effects	1,889,366	1,505,517	419,632	1,243,694	1,394,283	194,788
Number firm effects	216,459	185,086	70,170	169,100	178,332	33,633
Std. dev. of person effects (across person-yr obs.)	0.420	0.400	0.456	0.372	0.414	0.529
Std. dev. of firm effects (across person-yr obs.)	0.247	0.213	0.281	0.209	0.249	0.267
Std. dev. of Xb (across person-yr obs.)	0.069	0.059	0.078	0.058	0.065	0.063
Correlation of person/firm effects	0.167	0.152	0.036	0.120	0.169	0.055
RMSE of model	0.143	0.125	0.137	0.123	0.143	0.128
Adjusted R-squared of model	0.934	0.940	0.944	0.934	0.932	0.960
Comparison job-match effects model:						
Number of job-match effects	2,689,648	2,087,590	577,323	1,744,992	2,026,222	265,844
RMSE of match-effects model	0.128	0.113	0.128	0.112	0.129	0.119
Adjusted R-squared of match-effects model	0.946	0.951	0.953	0.946	0.944	0.965
Std. deviation of job match effect	0.062	0.054	0.048	0.052	0.061	0.046
Inequality decomposition of two-way fixed effects model:						
Share of variance of log wages due to:						
person effects	57.6	61.0	62.3	60.2	57.2	69.5
firm effects	19.9	17.2	23.5	18.9	20.6	17.7
covariance of person and firm effects	11.4	9.9	2.7	8.1	11.6	3.9
Xb and associated covariances	6.2	7.5	7.4	8.0	5.5	5.9
residual	4.9	4.4	4.1	4.8	5.1	3.0

Notes: See text. Models includes dummies for individual workers and individual firms, year dummies interacted with education dummies, and quadratic and cubic terms in age interacted with education dummies (total of 44 parameters). Comparison job-match effects models include dummies for each worker-firm job match as well as other covariates in basic model. Samples include only observations in largest connected set.

Table 4a: Contribution of Firm-based Wage Components to Male-Female Wage Gap

	Gender Group: Males (1)	Gender Group: Females (2)	Difference: Males−Females (pct. of overall gap) (3)
1. Mean log wage of group	1.715	1.481	0.234
			(100.0)
<u>Means of Estimated Firm Effects:</u>			
2. Firm Effects for Males ($\widehat{\psi}^{M}_{j(i,t)}$)	0.148	0.114	0.035
			(14.9)
3. Firm Effects for Females ($\widehat{\psi}^{F}_{j(i,t)}$)	0.145	0.099	0.047
			(19.9)
4. Within-group Difference in Mean Effects for Males and Females	0.003	0.015	
(percent of overall gap)	(1.2)	(6.3)	
5. Mean Male Firm Effect Among Men - Mean Female Firm Effect Among Women (Total contribution of firm-specific factors)			0.049
			(21.2)
6. Sample sizes	6,012,521	5,012,736	

Note: Sample includes male and female workers in "dual connected" set -- see Table 1, columns 5-6. Estimated firm effects are from models described in columns 1 and 2 of Table 3. Entries in parentheses are percents of the overall male-female wage gap in the dual connected sample that are explained by the differenced component.

Table 4b: Contribution of Firm-Level Pay Components to Gender Wage Gap at Dual Connected Firms, by Age/Education

		Means of Firm Premiums:			Decompositions			
					Sorting		Bargaining	
	Wage Gap (1)	Male Prem. Among Men (2)	Female Prem. Among Women (3)	Total Contribution of Firm Components (4)	Using M Effects (5)	Using F Effects (6)	Using M Distribution (7)	Using F Distribution (8)
All	0.234	0.148	0.099	0.049	0.035	0.047	0.003	0.015
				(21.2)	(14.9)	(19.9)	(1.2)	(6.3)
By Age Group:								
Up to age 30	0.099	0.114	0.087	0.028	0.019	0.029	-0.001	0.009
				(28.2)	(18.9)	(29.3)	1.2	(9.3)
Ages 31-40	0.228	0.156	0.111	0.045	0.029	0.040	0.004	0.016
				(19.7)	(12.6)	(17.8)	(1.9)	(7.0)
Over Age 40	0.336	0.169	0.099	0.069	0.050	0.064	0.005	0.019
				(20.6)	(15.0)	(19.1)	(1.5)	(5.6)
By Education Group:								
< High School	0.286	0.115	0.055	0.059	0.045	0.061	-0.002	0.015
				(20.8)	(15.6)	(21.4)	0.6	(5.2)
High School	0.262	0.198	0.137	0.061	0.051	0.051	0.010	0.010
				(23.3)	(19.6)	(19.5)	(3.8)	(3.7)
University	0.291	0.259	0.213	0.047	0.025	0.029	0.018	0.022
				(16.1)	(8.7)	(9.9)	(6.2)	(7.4)

Notes: see Table 4a. Numbers in parentheses represent the percent of the overall male female wage gap (in column 3) that is explained by source described in column heading.

Table 4c: Contribution of Firm-Level Pay Components to Gender Wage Gap: All Workers versus Workers in "Female" and "Male" Occupations

	Means of Firm Premiums:		Total Contribution of Firm Components	Decompositions			
				Sorting		Bargaining	
Wage Gap	Male Prem. Among Men	Female Prem. Among Women		Using M Effects	Using F Effects	Using M Distribution	Using F Distribution
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
a. All Workers at Dual Connected Firms							
(6,012,521 males and 5,012736 females at 84,720 firms)							
0.234	0.148	0.099	0.049	0.035	0.047	0.003	0.015
			(21.2)	(14.9)	(19.9)	(1.2)	(6.3)
b. Workers with "Female" Occupations at Firms that have Males and Females in "Female" Occupations							
(1,572,387 males and 3,403,802 females at 43,239 firms							
0.240	0.127	0.097	0.031	0.026	0.043	-0.012	0.005
			(12.8)	(10.8)	(17.8)	(-5.1)	(1.9)
c. Workers with "Male" Occupations at Firms that have Males and Females in "Male" Occupations							
(2,935,719 males and 801,113 females at 21,969 firms)							
0.137	0.177	0.133	0.044	0.015	0.027	0.016	0.028
			(31.9)	(11.1)	(20.0)	(11.9)	(20.8)

Notes: see notes to Table 4b. Sample in panel a. includes male and female workers who are employed at "dual connected" firms. Sample in panel b. includes male and female workers who mainly work in "female" occupations, and are employed at "dual-connected female-occupation" firms. Sample in panel c. includes male and female workers who mainly work in "male" occupations, and are employed at "dual-connected male-occupation" firms. Employees of either gender are classified into gender-occupation groups based on the female share of the occupation(s) they hold in all years. Decompositions in panel a. are based on estimated two-way fixed effects models fit to all men and all women. Decompositions in panel b. are based on estimated models fit to men and women in mainly female occupations. Decompositions in panel c. are based on estimated models fit to men and women in mainly male occupations.

Table 5: Estimated Relationship Between Estimated Firm Effects and Mean Log Value-Added per Worker

	Number Firms (1)	Regressions of Firm Effects on log(VA/L): All Males (2)	All Females (3)	Females in "Female" Occ's (4)	Females in "Male" Occ's (5)	Ratio to Men: All Females (6)	Ratio to Men: Females in "Female" Occ's (7)	Ratio to Men: Females in "Male" Occ's (8)
1. Dual connected with VA/L	47,477	0.156	0.137			0.879		
		(0.006)	(0.006)			(0.031)		
2. Dual connected, with VA/L and females in "female" occupations	42,667	0.155	0.136	0.136		0.879	0.875	
		(0.006)	(0.006)	(0.007)		(0.032)	(0.043)	
2. Dual connected, with VA/L and females in "male" occupations	14,638	0.138	0.128		0.129	0.924		0.933
		(0.008)	(0.008)		(0.009)	(0.048)		(0.049)

Notes: Columns 2-5 report coefficients of mean log value-added per worker in excess of 2.4 in regression models in which the dependent variables are the estimated firm effects for the gender/occupation group identified in the row headings. All specifications include a constant. Models are estimated at the firm level, weighted by the total number of male and female workers at the firm. Ratio estimates in columns 6-8 are obtained by IV method -- see text. Standard errors in parentheses.

Table 6: Descriptive Statistics for Job Stayers (2006-2009) at Firms with Value Added Data Available for 2006-2009

	All Firms with Male and Female Stayers		Firms w/ Female Stayers in Female Occupations		Firms w/ Female Stayers in Male Occupations	
	Males (1)	Females (2)	Males (3)	Females in Fem. Occ's (4)	Males (5)	Females in Male Occ's (6)
Mean Age	38.25	37.21	38.25	36.99	38.31	38.07
Mean Education	7.99	8.49	8.03	8.41	8.22	8.80
Mean Firm Size (workers in QP)	631	1024	668	1159	948	488
Mean Fraction of Females at Firm	0.29	0.59	0.29	0.62	0.30	0.50
Mean Log Real Hourly Wage 2006	1.62	1.40	1.63	1.37	1.68	1.52
(standard deviation)	(0.48)	(0.44)	(0.48)	(0.41)	(0.49)	(0.54)
Mean Log Real Hourly Wage 2009	1.70	1.48	1.70	1.44	1.75	1.60
(standard deviation)	(0.48)	(0.44)	(0.48)	(0.41)	(0.50)	(0.54)
Mean Log Value Added per Worker 2006	0.76	0.59	0.77	0.56	0.84	0.72
(standard deviation)	(0.50)	0.49	(0.49)	(0.48)	(0.50)	(0.50)
Mean Log Value Added per Worker 2009	0.76	0.58	0.77	0.55	0.82	0.71
(standard deviation)	(0.51)	(0.50)	(0.51)	(0.49)	(0.52)	(0.52)
Number of Workers	283,346	200,907	266,873	160,420	183,355	40,487
Number of Firms	33,104		28,893		11,820	

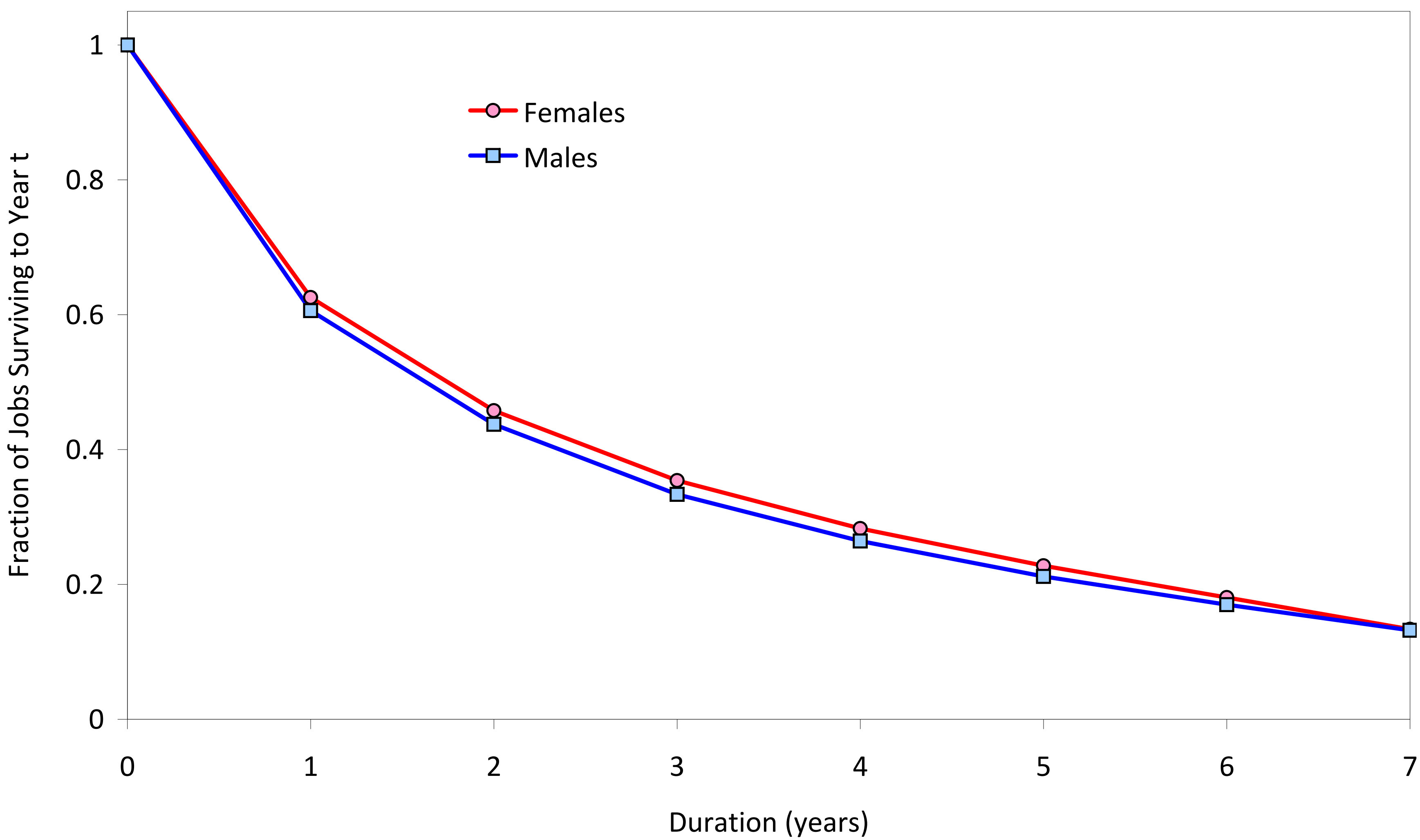
Note: Sample in columns 1-2 contains workers at dual connected firms with financial data for 2006-2009 who were employed at the firm continuously from 2006 to 2009. Sample in columns 3-4 is limited to workers at firms with at least one female stayer with a mainly female occupation. Sample in columns 5-6 is limited to workers at firms with at least one female stayer with a mainly male occupation. Age, education, firm size and fraction female refer to 2006.

Table 7: Effects of Changes in Excess Log(VA/L) on Wages of Stayers

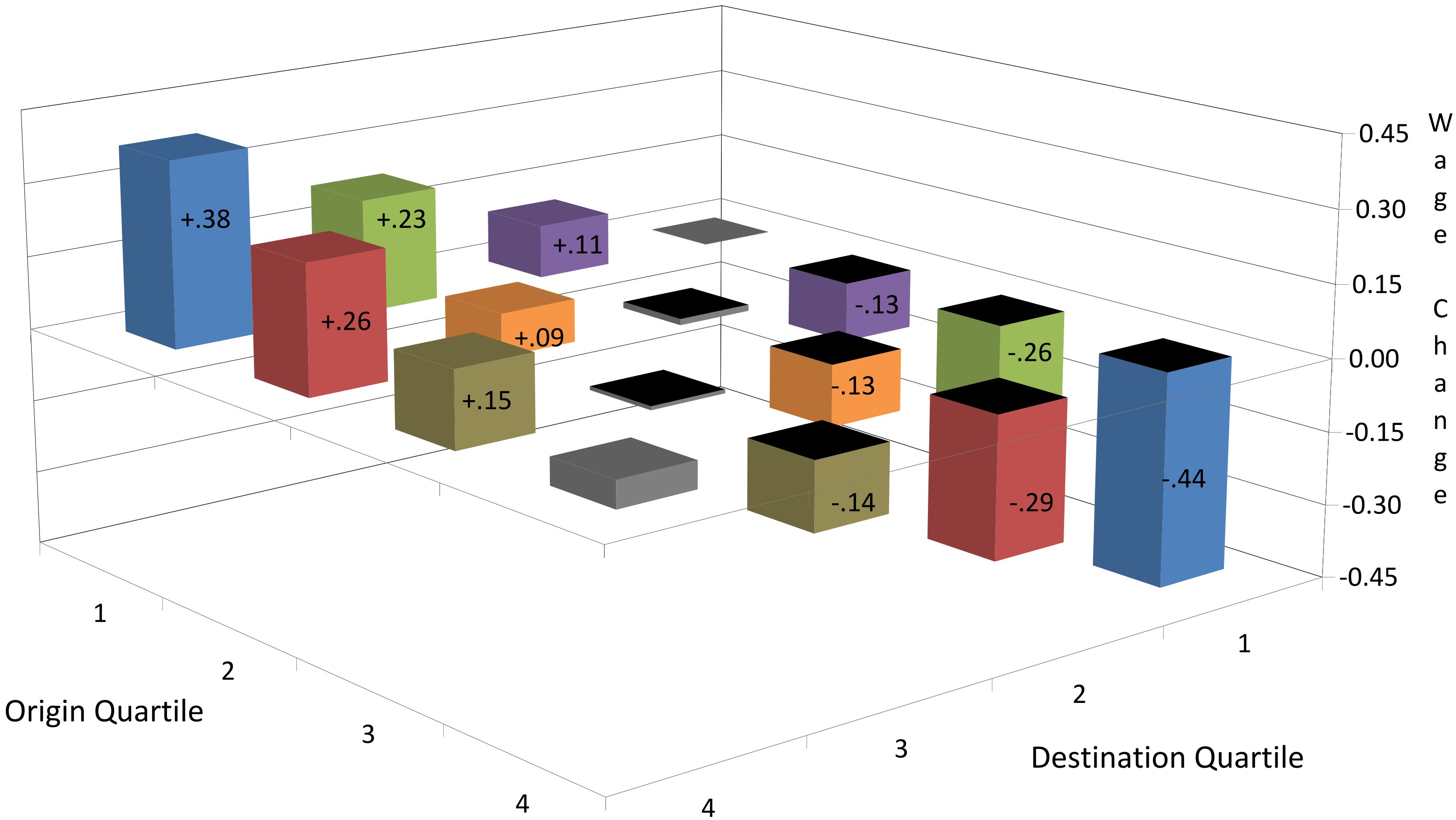
	All Male Stayers		All Female Stayers			Female Stayers in "Female" Occupations			Female Stayers in "Male" Occupations		
	OLS	IV	OLS	IV	Ratio to Men	OLS	IV	Ratio to Men	OLS	IV	Ratio to Men
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. All Firms	0.049	0.092	0.045	0.091	0.912						
(n=33,104)	(0.007)	(0.020)	(0.008)	(0.020)	(0.086)						
first stage F		231.95		231.95	45.02						
2. Firms with Females in Mainly Fem. Occ's	0.049	0.092	0.044	0.086	0.893	0.043	0.075	0.869			
(n=29,893)	(0.008)	(0.021)	(0.008)	(0.021)	(0.088)	(0.008)	(0.022)	(0.100)			
first stage F		208.80		208.80	40.58		208.80	40.58			
2. Firms with Females in Mainly Male Occ's	0.044	0.073	0.039	0.095	0.881				0.050	0.131	1.116
(n=11,820)	(0.011)	(0.028)	(0.012)	(0.029)	(0.128)				(0.013)	(0.034)	(0.204)
first stage F		113.00		113.00	16.16					113.00	16.16

Notes: Notes: See Table 6 for sample description. Dependent variable is average change in wages of male or female workers from 2006 to 2009 at a firm (regression-adjusted for quadratic in age). Table entries are coefficients of the change in excess log of value added per worker from 2006 to 2009 at the firm, , Winsorized at +/ 0.50. IV models use change in Winsorized excess log value added per worker from 2007 to 2008 as instrument. Ratios in columns 5, 8, 11 are estimated by IV, treating average change in female wages as dependent variable, average change in male wages as endogenous explanatory variable, and change in Winsorized log value added per worker from 2006 to 2009 as the instrument. Standard errors, clustered by firm, in parentheses.

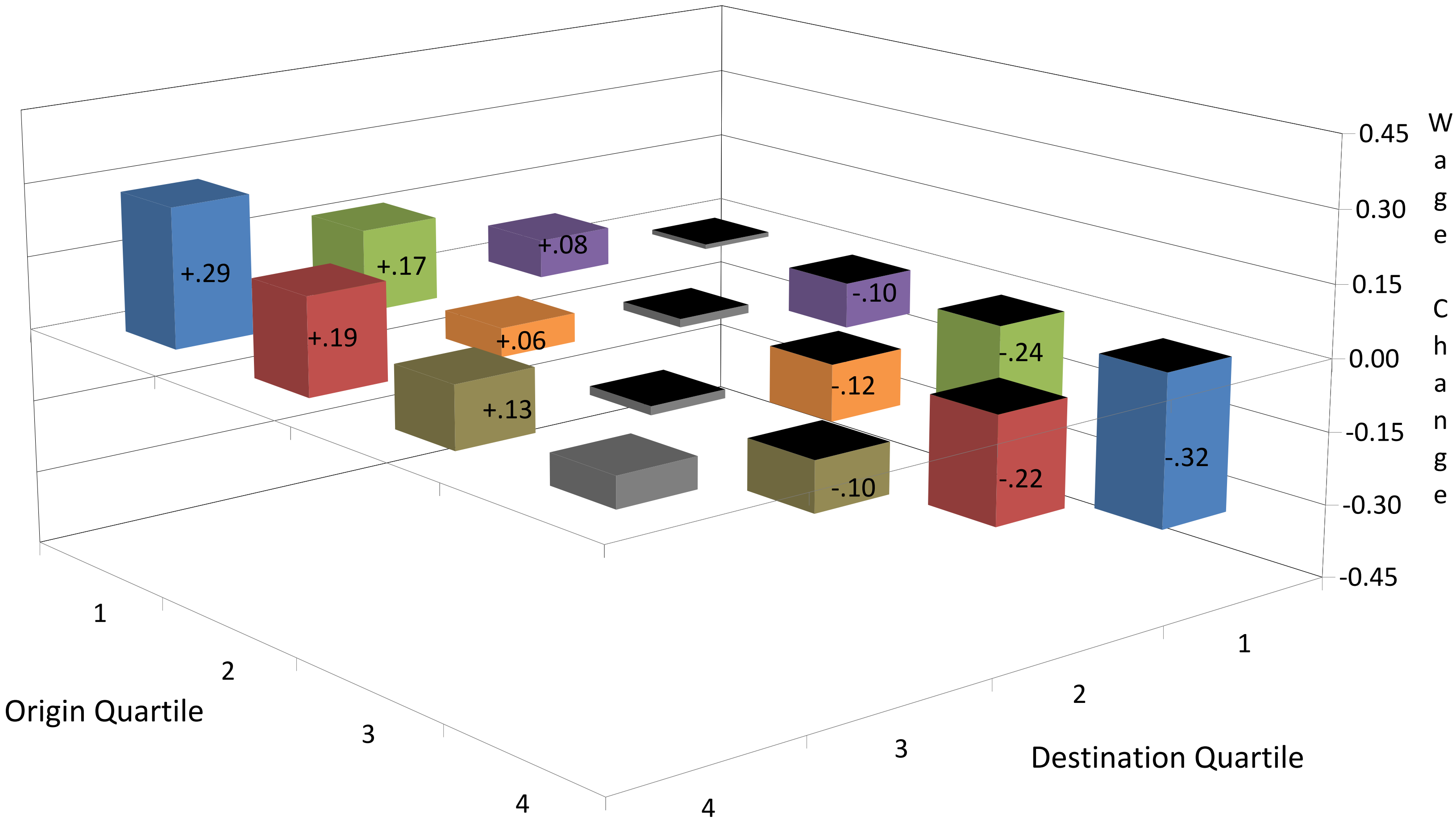
Appendix Figure 1: Job Survival Rates for New Jobs Starting 2002-2008



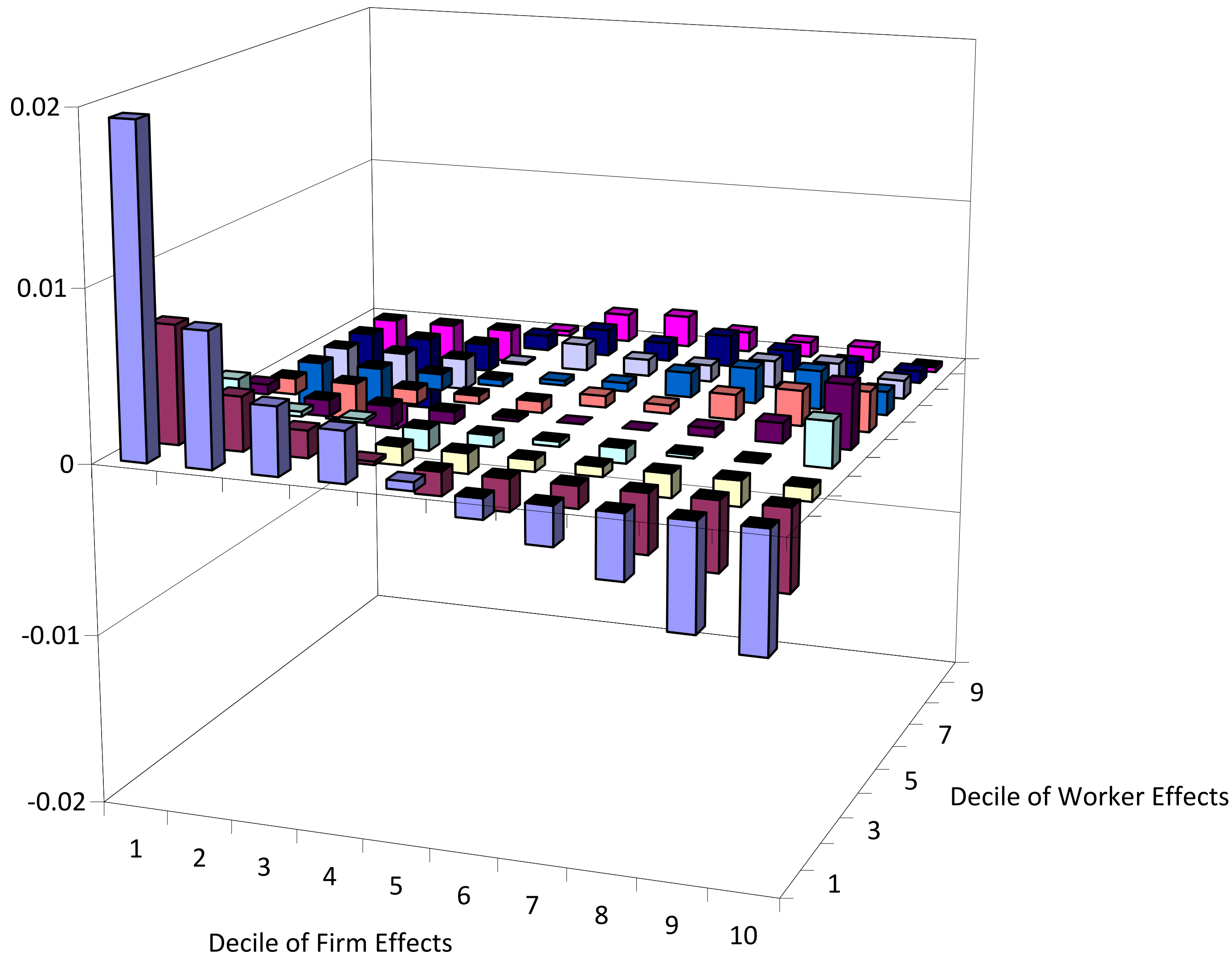
Appendix Figure A2: Regression-Adjusted Wage Changes Associated with Transitions Between Co-Worker Quartiles - Men



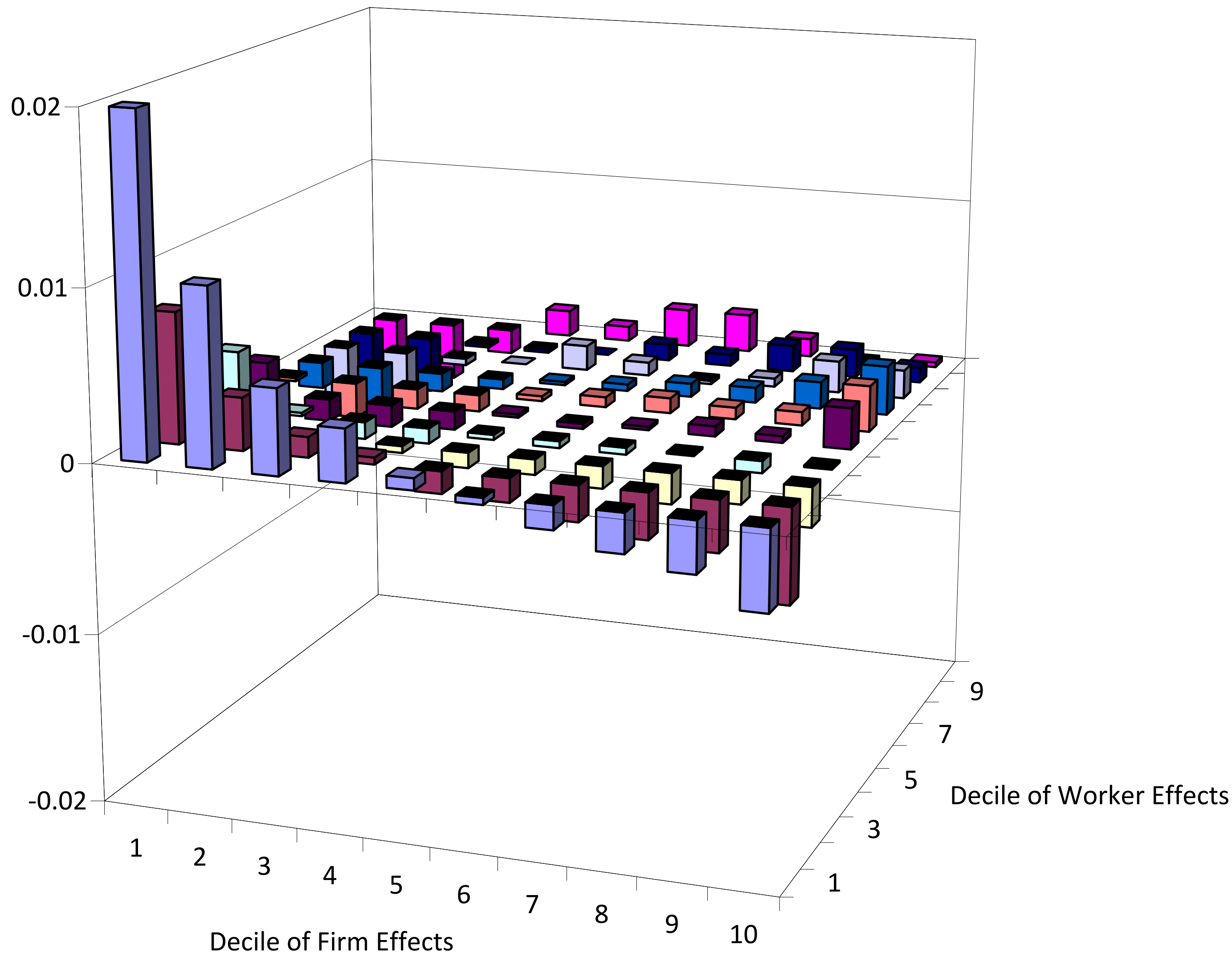
Appendix Figure A3: Regression-Adjusted Wage Changes Associated with Transitions Between Co-Worker Quartiles - Women



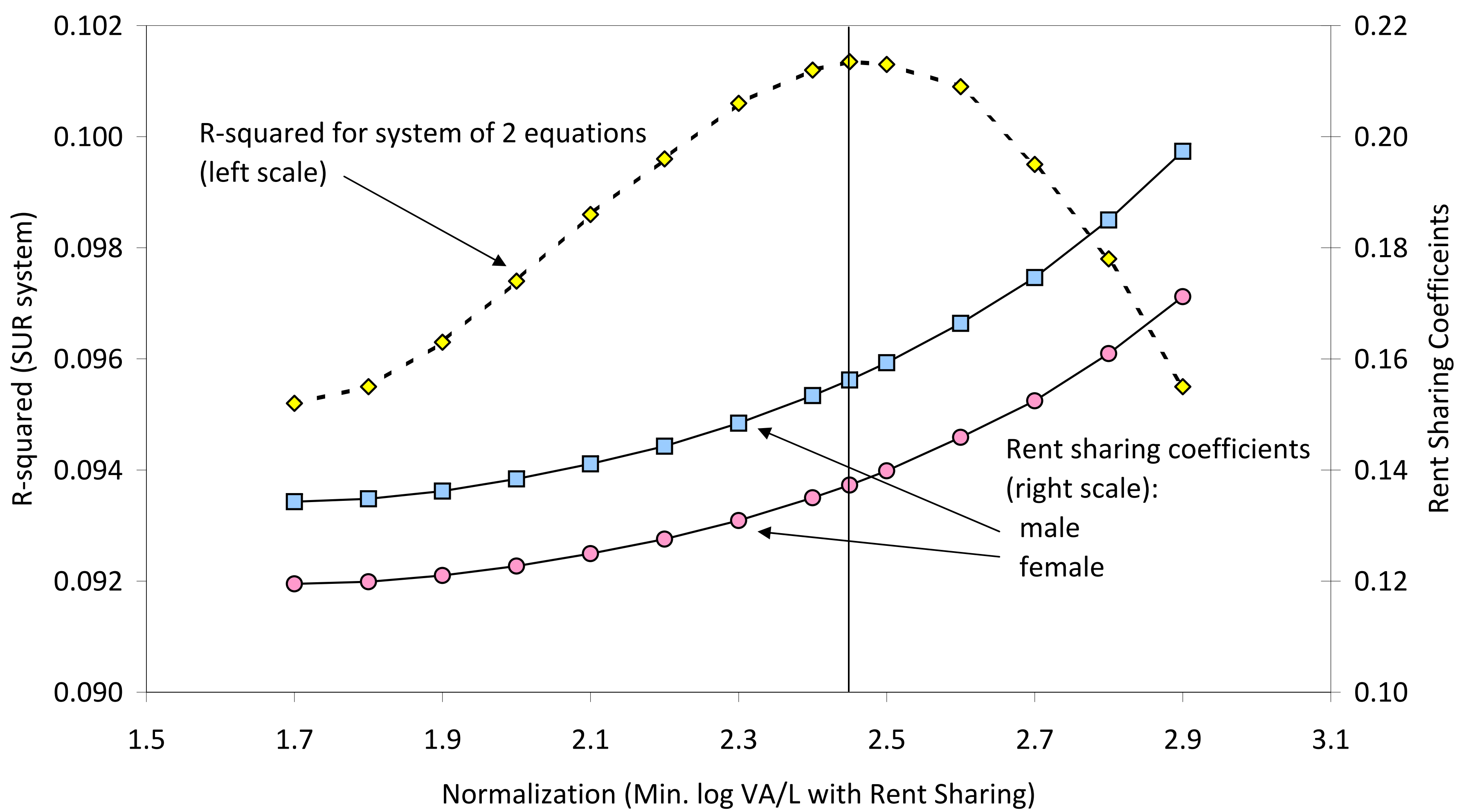
Appendix Figure A4: Mean Residuals for Males by Decile of Worker and Firm Effects



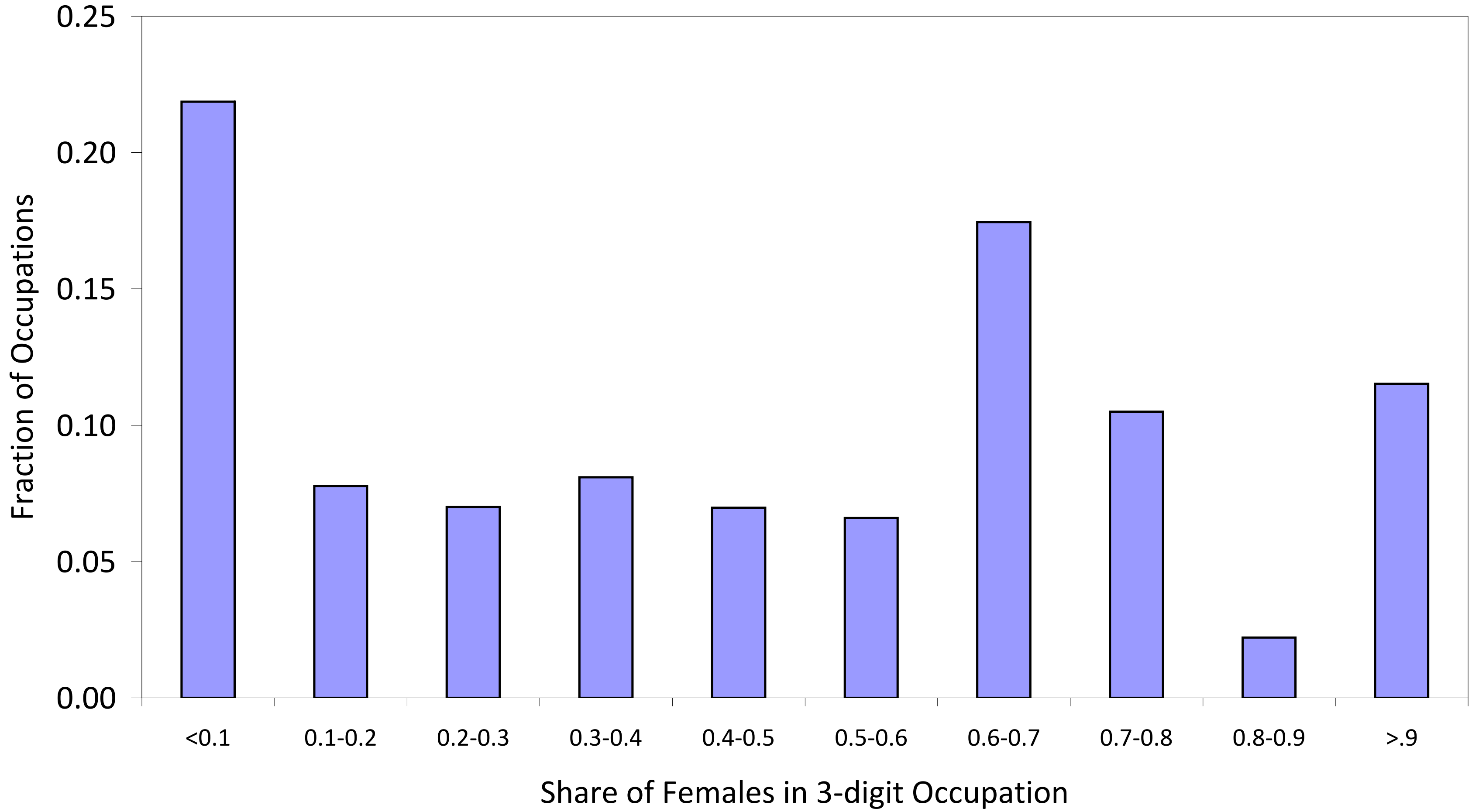
Appendix Figure A5: Mean Residuals for Females by Decile of Worker and Firm Effects



Appendix Figure A6: Goodness of Fit and Rent Sharing Coefficients for Alternative Normalizations



Appendix Figure A7: Histogram of Female Employment Shares Across Occupations



Note: based on employment-weighted female shares in 110 3-digit occupations.

Appendix Table A1: Descriptive Statistics for Overall QP and Analysis Sample

	Overall Population of Employees in QP		Analysis Sample	
	Males	Females	Males	Females
	(1)	(2)	(3)	(4)
Age:				
Mean Age	38.9	37.0	38.1	36.9
Fraction ≤ 30 years old	0.28	0.32	0.30	0.33
Fraction ≥ 50 years old	0.21	0.16	0.19	0.14
Education:				
Mean Years Schooling	8.0	8.8	8.0	8.8
Fraction with High School	0.18	0.23	0.18	0.23
Fraction with University Degree	0.10	0.14	0.09	0.13
Mean Log Real Hourly Wage	1.61	1.42	1.59	1.41
(standard dev.)	(0.58)	(0.52)	(0.55)	(0.50)
Mean Monthly Hours	161.9	156.7	162.6	158.0
(standard dev.)	(25.9)	(31.8)	(24.7)	(30.1)
Fraction in Lisbon	0.35	0.36	0.35	0.35
Fraction in Oporto	0.13	0.13	0.13	0.13
Mean Firm Size (Number employees)	668	839	730	858
Fraction Female Workers at Firm	0.25	0.66	0.24	0.70
Number person-year obs.	11,651,615	9,011,089	9,070,492	7,226,310
Number of persons	2,550,576	2,040,863	2,119,687	1,747,492
Number of firms	431,991	391,982	349,692	336,239

Notes: Overall sample in columns 1-2 includes all observations available in QP with consistent data for age, gender and education. Analysis sample in columns 3-4 excludes individuals with inconsistent employment histories. Person-year observations are also conditioned on being a paid worker in the year, age 19-65, with potential experience ≥2. Wages are measured in real (2009=100) Euros per hour. Lisbon refers to Greater Lisbon and Setubal, Oporto refers to Greater Oporto (NUTS-3 classifications).

Appendix Table A2: Matching Rates of Observations in QP to Firm Identifier in SABI

	Percent of All Observations in Industry	Percent of Observations Matched	Percent of Observations from Females	Percent Matched by Gender	
				Male	Female
	(1)	(2)	(3)	(4)	(5)
All Industries	100.0	70.9	44.3	73.3	67.8
Agriculture	1.7	52.0	40.6	50.9	53.7
Fishing	0.1	84.2	27.1	82.2	89.5
Mining	0.4	80.8	9.7	80.5	83.7
Food Products	3.5	75.2	49.1	74.9	75.6
Textiles	8.0	81.3	71.7	80.9	81.5
Wood Products	2.8	78.3	27.0	76.7	82.6
Paper	1.5	79.3	34.5	77.4	82.8
Chemicals	1.8	82.2	34.2	80.8	84.9
Other Mineral Products	2.0	81.4	29.3	81.0	82.4
Metal Fabrication	7.0	80.4	25.8	80.0	81.6
Utilities	0.8	86.5	17.7	87.3	82.7
Construction	12.4	69.5	8.3	69.0	74.1
Trade	19.5	79.4	46.6	79.8	78.9
Hotels	6.6	75.7	62.4	78.2	74.3
Transportation	5.9	71.9	22.9	72.9	68.5
Finance	2.8	27.5	44.4	27.6	27.3
Business Services	10.6	82.6	49.1	83.5	81.6
Education	2.1	42.8	76.2	44.1	42.5
Health	5.9	35.9	88.3	44.1	34.8
Recreation Services	1.0	64.2	44.2	66.4	61.5
Other	3.6	30.6	67.1	32.4	29.7

Note: All statistics are calculated across person-year observations in QP analysis sample for 2002-2009. "Matched" means that employer of person in given year can be matched to firm in SABI. Sample contains 9,070,492 person-year observations for males and 7,226,310 for females.

Appendix Table A3: Distributions of Number of Jobs Held in Sample Period, by Gender, and Mean Log Wage by Number of Jobs Held

	Distribution of Number of Jobs Held 2002-2009 (Person-year weighed)		Distribution of Number of Jobs Held 2002-2009 (Person-weighed)		Mean Log Wage of Persons, By Number of Jobs Held 2002-2009		
	Males	Females	Males	Females	Males	Females	Male-Female Gap
# Jobs	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	67.81	70.37	72.50	74.29	1.56	1.38	0.17
2	20.93	20.42	18.71	18.51	1.45	1.31	0.15
3	7.91	6.84	6.39	5.53	1.43	1.29	0.14
4	2.52	1.87	1.85	1.35	1.41	1.28	0.13
5	0.68	0.41	0.46	0.27	1.39	1.27	0.12
6	0.13	0.08	0.08	0.05	1.39	1.26	0.13
7	0.02	0.01	0.01	0.01	1.37	1.22	0.14
8	0.00	0.00	0.00	0.00	1.39	1.48	-0.09
# Obs.	9,070,492	7,226,310	2,119,687	1,747,492	2,119,687	1,747,492	--

Notes: tabulations based on overall population of male and female employees in QP data set -- see columns 1 and 2 of Table 1. There are 15 males and 7 females with 8 jobs in the sample, accounting for 120 person-year observations for men and 56 person-year observations for women.

Appendix Table A4: Additional Models for Relationship Between Firm Effects and Mean Log Value-Added per Worker

		Regressions of Firm Effects on log(VA/L)						
	Number Firms (1)	All Males (2)	All Females (3)	Females in "Female" Occ's (4)	Females in "Male" Occ's (5)	Ratio to Men: All Females (6)	Ratio to Men: Females in "Female" Occ's (7)	Ratio to Men: Females in "Male" Occ's (8)
1. Dual connected with VA/L	47,477	0.140 (0.006)	0.117 (0.007)			0.839 (0.036)		
2. Dual connected, with VA/L and females in "female" occupations	42,667	0.139 (0.006)	0.116 (0.007)	0.115 (0.008)		0.838 (0.037)	0.831 (0.045)	
2. Dual connected, with VA/L and females in "male" occupations	14,638	0.127 (0.007)	0.107 (0.010)		0.115 (0.011)	0.845 (0.055)		0.907 (0.060)

Notes: Columns 2-5 report coefficients of mean log value-added per worker in excess of 2.4 in regression models in which the dependent variables are the estimated firm effects for the gender/occupation group identified in the row headings. All specifications include 20 industry dummies, dummies for Lisbon and Porto, and quadratic in average number of employees at firm. Models are estimated at the firm level, weighted by the total number of male and female workers at the firm. Ratio estimates in columns 6-8 are obtained by IV method -- see text. Standard errors in parentheses.

Appendix Table A5: Decomposition of Male-Female Wage Gap, Based on Relation of Firm Effects to Excess Value Added

	Males (1)	Females (2)	Male-Female Difference (3)
1. Total Firm-specific Component of Wages (from Table 4b)	0.148	0.099	0.049
			(21.2)
2. Rent sharing coefficient (from row 1 of Table 5)	0.156	0.137	0.019
3. Mean "excess VA/L"	0.743	0.566	0.178
4. Firm-specific Component of Wages Attributable to Measured Productivity (= row 2 × row 3)	0.116	0.078	0.038
			(16.5)
5. Share of Total Firm-specific Component Attributable to Measured Productivity (= row 4 ÷ row 1)	0.784	0.785	0.776
Counterfactuals:			
a. Assign females the male firm effects	0.116	0.088	0.028
(sorting effect, using male coefficients)			(11.9)
b. Calculate mean female firm effect using male firm distribution	0.116	0.102	0.014
(bargaining effect, using male distribution)			(6.0)
c. Assign males the female firm effects	0.102	0.078	0.024
(sorting effect, using female coefficients)			(10.4)
d. Calculate mean male firm effect using female firm distribution	0.088	0.078	0.011
(bargaining effect, using female distribution)			(4.6)

Note: decomposition based on regression models presented columns 2 and 3 (row 1) of Table 5. See text. Entries in parentheses in column 3 represent the share of the overall male-female wage gap (0.234) that is explained by the rent sharing component under alternative counterfactuals.

Appendix Table A6: Effects of Changes in Excess Log(VA/L) on Wages of Stayers

	All Male Stayers		All Female Stayers			Female Stayers in "Female" Occupations			Female Stayers in "Male" Occupations		
	OLS	IV	OLS	IV	Ratio to Men	OLS	IV	Ratio to Men	OLS	IV	Ratio to Men
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1. All Firms	0.035	0.075	0.031	0.071	0.894						
(n=33,104)	(0.006)	(0.016)	(0.006)	(0.016)	(0.091)						
first stage F		159.26		159.26	39.31						
2. Firms with Females	0.035	0.075	0.031	0.066	0.883	0.031	0.060	0.906			
in Mainly Fem. Occ's	(0.006)	(0.017)	(0.006)	(0.016)	(0.094)	(0.006)	(0.017)	(0.109)			
(n=29,893)											
first stage F		144.48		144.48	35.28		144.48	35.28			
2. Firms with Females	0.030	0.065	0.025	0.077	0.856				0.029	0.098	0.991
in Mainly Male Occ's	(0.008)	(0.022)	(0.008)	(0.023)	(0.141)				(0.009)	(0.027)	(0.215)
(n=11,820)											
first stage F		76.74		76.74	11.56					76.74	11.56

Notes: See Table 7. Dependent variable is average change in wages of male or female workers from 2006 to 2009 at a firm (regression-adjusted for quadratic in age). Table entries are coefficients of the change in log of excess value added per worker from 2006 to 2009 at the firm. IV models use change in excess log value added per worker from 2007 to 2008 as instrument. Ratios in columns 5, 8, 11 are estimated by IV, treating average change in female wages as dependent variable, average change in male wages as endogenous explanatory variable, and change in excess log value added per worker from 2006 to 2009 as the instrument. Standard errors, clustered by firm, in parentheses.