The Wage Growth and Within-Firm Mobility of Men and Women: New Evidence and Theory *

Mary Ann Bronson† Peter Skogman Thoursie‡

September 2022

Abstract

Why do women's wages grow more slowly than men's? Using Swedish data, we answer this question in three steps. First, we analyze men's and women's lifecycle wage growth non-parametrically. We document that their annual wage growth distributions differ primarily in one respect: Women are less likely to experience infrequent, exceptionally large (right-tail) pay increases, primarily in years that they do not switch firms. These increases are persistent and move workers substantially through their firms' wage distributions, indicative of large internal promotions; they account for a striking 70% of the gender difference in wage growth by age 45. Second, we analyze the extensive margin of these right-tail pay increases. We construct an empirical measure that allows us identify and analyze comparable high-paying within-firm "promotions" across thousands of firms. We show that the cumulative gender gap by age 45 in such promotions is incurred primarily early in the lifecycle and between men and women working at the same firm, with gender differences in sorting across forms playing a minor role in accounting for the promotion gap. Most of the gap is incurred well before women begin to work part-time. Probability of promotion drops dramatically for women in the year they give birth and the year immediately following; however, gender differences in promotion, even prior to first birth and also for women who remain childless, are sizable, and reverse after age 40 to favor women. Lastly, we interpret the novel facts we document using a theoretical model of careers within firms based on Gibbons and Waldman (1999) and discuss implications for models of wage dynamics more broadly.

*We are especially grateful to Michael Waldman and Arik Levinson for helpful comments on this paper, as well as to Alex Bick, Stephane Bonhomme, Dan Cao, Pedro Carneiro, Garance Genicot, Nora Gordon, Fatih Guvenen, Rebecca Lessem, Frank Vella, and participants at various seminars and conferences.

†Georgetown University, Department of Economics, Intercultural Center 580, 37th and O Streets, N.W., Washington D.C. 20057. Email: mary.ann.bronson@gmail.com.

‡Stockholm University, Department of Economics, Stockholm. Email: peter.thoursie@ne.su.se.

1 Introduction

In most countries and skill groups, the gender gap in wage *levels* primarily reflects a gender gap in wage *growth.* In the U.S., for example, the difference in men's and women's average log hourly wages more than quadruples between ages 25 and 45. Among the college-educated, it increases from 5% at age 25 to a striking 30% at age 45.

Consequently, researchers now increasingly focus on the dynamics in the gender wage gap, documenting its expansion by age and time since first birth, along with correlates such as hours worked or firm characteristics.[1] However, the underlying wage growth patterns generating these lifecycle dynamics are still not well-understood. Consider the following simple questions about year-on-year wage growth, on which there is little direct evidence. Do women accept large hourly wage *cuts* more frequently over the lifecycle? Do their wages to grow marginally less than men's over many years? Or do women mainly fail to realize isolated years of especially high wage growth? Moreover, do these differences in growth occur within firm or during firm changes? Each described scenario widens the average wage gap with age, but calls for a different explanation and model of wage dynamics.[2]

In this paper, we provide one of the first systematic empirical characterizations of differences in men's and women's wage growth. The descriptive evidence we present is novel in two main ways. First, we analyze real annual wage growth non-parametrically. Our comparison of men's and women's wage growth distributions provides a transparent and exceptionally rich summary of the differences in their wage evolution over the lifecycle, both within and across firms. For example, it readily answers questions such as those posed above. Our second innovation is in our investigation of within-firm growth and, more specifically, within-firm mobility, which we will show to be crucial for understanding differences in men's and women's wage growth. To do this, we exploit employer-employee linkages to identify years in which workers experienced exceptionally high wage growth relative to their co-workers at the same firm.[3] This approach isolates comparable "promotions" of workers in heterogeneous skill groups and firms, and will allow us to significantly expand the existing evidence on gender differences in within-firm advancement (e.g., McCue (1996), Blau and DeVaro (2007)).

Our paper has three parts. The first two—focused, respectively, on wage growth and within-firm mobility—are empirical. In the third part, we develop a simple theoretical model to interpret the facts we document, and draw out the implications of our findings

[1] E.g., see Manning and Swaffield (2008), Bertand et al. (2010), Card et al. (2016), Albrecht et al. (2019), Kleven et al. (2019), Angelov et al. (2016), and Hotz et al. (2017), among many others.

[2] Large wage cuts, especially after childbirth, are consistent with models emphasizing skill depreciation (e.g. Adda et al. (2013)) or mid-career moves to lower-paying but more flexible firms (e.g., Hotz et al. (2017)). Marginally lower wage-growth over many years is consistent with women's work in lower-productivity industries (e.g., Hellerstein et al. (2008)). Failure to realize years of especially high wage growth may indicate fewer promotions (e.g., Blau and DeVaro (2007)) or moves to firms with high pay premiums (e.g., Card et al. (2016)).

[3] Prior studies have examined cross-sectional within-firm wage inequality (e.g., Song et al. (2019)).

for prioritizing existing explanations and models of the gender wage gap.

For our empirical analysis, we use Swedish data from 1985 to 2013, following men and women from the time they enter the labor market to age 45. Our data contains detailed demographic information, such as choice of major and the timing of all births. In addition to annual labor income, it includes a wage variable, for both hourly and salaried workers, that is administratively recorded and characterized by minimal measurement error.[4] Finally, Sweden provides a useful setting for our analysis, as the college-educated men and women we study have virtually identical labor force participation rates of 95% and 96%, respectively. This reduces concerns about selection or changes in composition of workers with age.

Our main findings are as follows. First, our non-parametric evidence indicates that men's and women's real annual wage growth distributions are, in fact, strikingly similar. For example, men and women take wage cuts at almost identical rates; they experience low to moderate wage growth of around 0% to 4% in most years; and much of their lifecycle wage growth is generated by a small number of high-growth years. Their wage growth differs mainly in one respect: Women are less likely than men to experience years of exceptionally high wage growth – that is, growth in the right tail of the distribution, on the order of 15% or more – primarily in years that they do *not* switch firms. Such right-tail, within-firm increases are highly persistent and move workers on average 17 percentiles higher in their firm's wage distribution, indicative of large, internal promotions. These facts are both novel and important. Prior research is mostly about mean gender differences in wage levels over the lifecycle. Our finding that such differences are generated primarily by women's lower incidence of right-tail wage shocks – within firms, in the form of outsize pay hikes relative to co-workers – constitutes a crucial step towards understanding the mechanisms behind men's and women's different wage dynamics.

Based on this evidence, in the second part of our analysis we focus on men's and women's within-firm mobility. To do this, we construct a variable that identifies years in which individuals' wage growth significantly exceeds (e.g., by 10+ log points) the wage growth of similar co-workers at their firm. We refer to these increases, defined formally in Section 3, as "promotions," to underline that they constitute large moves through firms' wage hierarchies.[5] This variable isolates the type of right-tail wage increase that women experience at lower rates, allowing us to study its extensive margin. More generally, however, it also constitutes an exceptionally useful tool for characterizing within-firm mobility patterns population-wide, for two reasons. First, it is tractable as it does not require knowledge of every firm's organizational hierarchy. Second, it uses a wage-based metric that is directly comparable across organizations, circumventing the typical concerns about whether reported promotions can be meaningfully compared across individuals or

[4]By contrast, constructed hourly wage variables obtained by dividing annual earnings by reported hours are known to contain substantial measurement error (Keane (2011)).

[5]Of course, some may reflect merit increases or matches to outside offers, with no corresponding job upgrade.

firms (Pergamit and Voym, 1999).

Using this measure, we first estimate that the differential incidence of such large, within-firm "promotions" for men and women accounts for a striking 70% of the gap in wage growth, even under fairly conservative assumptions about what constitutes a promotion. This fact is not simply mechanical: Promotions generate about 36%-43% of cumulative wage growth by age 45 in our population, a sizable share but significantly less than 70%. Both in percent and absolute terms, gender disparities in promotion probability are largest early in the lifecycle. Between ages 25 and 30, for every promotion awarded to a woman, her male co-worker receives 1.3 promotions. While women are more likely to work at firms with fewer opportunities for upward mobility – in line with findings by Hotz et al. (2017) and Kleven et al. (2019) – this accounts for only 10% of the promotion gap in the population we study, indicating that the gap is driven primarily by differences in promotion probability between men and women working at the *same* firm. We document also that choice of major and occupation play a minor role in accounting for women's lower promotion rates.

Three factors jointly account for nearly the entire gap in cumulative promotions between male and female co-workers by age 45. The first is forgone promotions in the year of and year after giving birth. Women's probability of promotion drops dramatically, relative to observationally similar men at the same firm. These are years when Swedish women are typically on parental leave. Those isolated years alone account for 40% of the gender difference in cumulative promotions by age 45 for individuals who ever have children. We find no evidence that these "missed" promotions are readily recovered.

Second, women's higher rate of part-time work after having children accounts for 21% of the cumulative gap by age 45 for women who ever have children. Part-time work thus plays a quantitatively significant role, but is not a dominant driver. Notably, most promotions occur early on, before Swedish women work part-time at significant rates.

Lastly, a residual gender penalty in promotion probability prior to first birth accounts for about 30% of the cumulative promotion gap for those who ever have children. It declines with age, reverses after age 40 to favor women, and is not associated with what the literature calls a "motherhood penalty" (Angelov et al. (2016), Kleven et al. (2019)). A gender penalty of comparable magnitude is also observed for women who never have children, similarly reversing after the end of their childbearing years. It is greater at small establishments and at firms where few men use their allotted parental leave.

Our investigation is data-driven and purposely not guided by one particular model. Ultimately, however, a model is useful for interpreting the large set of facts we catalogue. Given the dominant role within-firm mobility plays empirically, we build on one of the workhorse models of careers within firms, by Gibbons and Waldman (1999b), to interpret our findings. In their classic model, individuals of different abilities are promoted over time as they gain experience and become more productive. To analyze gender differences,

we add two features: a childbearing phase for workers, during which women are more likely to reduce their labor supply and take leave than men; and a cost to employers associated with such labor supply reductions, in line with evidence by Ginja et al. (2020).

Our main theoretical results can be summarized as follows. First, women – regardless of whether or not they eventually have children – are on average promoted less than men early on in their careers. The reason for this is that employers anticipate incurring future costs associated with women's labor supply reductions, and women have no credible way of signaling their childbearing intent. Second, women who have children additionally experience a large motherhood penalty in promotion probability after birth, since employers prefer not to promote workers who are currently on leave. This motherhood penalty persists in later years due to forgone human capital accumulation. Third, the "gender penalty" in promotion probability favors women after the end of their childbearing years: with uncertainty about future childbearing resolved, high-ability women initially "passed up" for promotion advance. We conclude by showing that the model matches several auxiliary facts we document, and that are ex ante not easy to explain, including gender differences in wage dynamics in periods when workers are not promoted. We then evaluate whether alternative theoretical explanations for gender differences in wage trajectories are consistent with our empirical findings.

Our analysis has important implications for several branches of the literature on the gender wage gap. First, it helps explain the emergence of the rapid and persistent hourly wage gap immediately after first birth (e.g., Kleven et al. (2019), Angelov et al. (2016)). The dramatic drop in promotions in this period generates significant, immediate wage losses relative to men, due to their sizable contribution to wage growth. These "missed" promotions are not readily recovered, generating persistence. Second, our findings indicate that while popular and scholarly emphasis on the "glass ceiling" is not unjustified — few women make it to the top — gender disparities are largest at the earliest stages of women's careers, in both percent and absolute terms. Our findings therefore strongly support recent non-academic research emphasizing "broken bottom rungs" over glass ceilings as the main obstacles in women's careers (McKinsey & Co. and LeanIn.Org (2019)).

Third, our findings answer important questions about the role of firms in women's wage dynamics: whether women experience lower wage growth because they work at low-paying or low-growth firms; whether they work in part-time "mommy-tracks" at otherwise high-growth firms; or whether much of the wage growth and promotion gap is between male and female co-workers at the same firm with overall similar characteristics. While all three play a role, we show that overwhelmingly the third is dominant. Consequently, we argue that models that primarily emphasize differences in how men and women sort and match with firms – e.g., due to preferences for workplace amenities, or being a tied mover – are unlikely to account quantitatively for the bulk of wage divergence with age,

at least for high-skilled individuals.[6] More broadly, our finding that most right-tail wage increases occur within-firm has important implications for models that load a large share of high-growth spells onto firm changes (e.g., Cahuc, Postel-Vinay, and Robin (2006)).

We conclude by highlighting several related papers we have not yet discussed. Maasoumi and Wang (2020) also characterize distributional aspects of the gender gap. While their non-parametric analysis concerns cross-sectional gender differences in earnings levels at a point in time, we focus on gender differences in wage growth over the lifecycle. In a methodologically related study, Guvenen et al. (2014) investigate annual income growth nonparametrically to analyze cyclical earnings risk. However, they characterize growth in total yearly earnings, not hourly wage. Total earnings growth is strongly affected by changes to annual hours worked and therefore exhibits different empirical patterns.

Prior empirical work on gender and within-firm mobility (see Blau and DeVaro (2007)) focuses heavily on two specific questions: whether women are promoted less than men, and whether they are paid less when promoted. As Blau and Devaro (2007) point out in their excellent review of this literature, estimates vary widely in size and sign, possibly due to frequent reliance on data from individual, non-representative firms. In data for representative populations, on the other hand, characterizing promotions consistently across firms is a challenge (Pergamit and Voym, 1999). Consequently, evidence on many features of population-wide within-firm mobility patterns is limited, e.g. on the effects of firm choice, and even on the quantitative contribution of within-firm mobility to lifecycle wage growth, as we will discuss further in Section 2. Our study thus complements and significantly expands the facts available to researchers about men's and women's different within-firm mobility patterns, a key driver of their different wage dynamics.

Lastly, our model is related to Lazear and Rosen's (1990) model of statistical discrimination (see also Milgrom and Oster (1987) and Thomas (2018)). In contrast to their model, ours focuses on time-to-birth promotion dynamics, gender vs. motherhood penalties, and wage growth also outside of promotion. Relative to previous studies, we show that a classic equilibrium model of careers within firms matches well the main facts about gender and wage growth, including the initial absence of a significant wage gap, its rapid subsequent expansion, and important non-parametric features, like the thick right tail of the wage growth distribution.

In the remainder of the paper, we first describe our data and key wage growth and within-firm mobility measures (Section 2). We then present our empirical results (Sections 3 and 4); our model, and implications for the literature (Sections 5 and 6); and a sensitivity analysis (Section 7). Section 8 concludes.

[6]Papers focusing on gender differences in firm choice and wage gains from firm changes include Card, Cardoso, and Kline (2016), Sorkin (2017), and Morchio and Moser (2020)). Hotz et al. (2017) and Kleven et al. (2019) find that women select firms with fewer opportunities for upward mobility, especially after first birth, as proxied by different measures of family friendliness.

2 Data and Construction of Key Variables

2.1 Swedish Data and Institutional Context

We rely on several administrative registers from Statistics Sweden, covering 1985-2013. Our central dataset is the LOUISE register, which contains demographic variables for the entire population of Sweden aged 16-75, including age, gender, household composition, years of post-secondary education, age at graduation, and field of major. It also provides information about parental leave pay and annual labor income, including zero income. We link this dataset with three other registers using personal and firm identifiers. Wage Structure Statistics provides information on a worker's contracted hours and contracted wages. The multi-generational register provides details about the dates of all births. Finally, the employer register provides personal identifiers of all employees, as well as certain firm characteristics, such as size, industry, and sector (private or public). Using these linked registers, we can analyze firm-specific wage distributions and other firm characteristics by gender or education.

An important feature of our data is that it records the contracted wage – full-time equivalent (FTE) monthly earnings – and not just total annual labor earnings.[7] This is useful when studying women, whose hours worked can vary significantly from year to year, especially around the time of childbirth, generating substantial variation in total labor earnings that is not directly related to hourly compensation. Obtaining a precise wage measure is often impossible using tax or social security records, which generally only record annual earnings, or in worker-firm linked survey data such as Longitudinal Employer-Household Dynamics (e.g., Barth et al. (2017).) In our data, wage information is collected once yearly for employees with positive hours in the survey month. It is collected for all public-sector employees and for all workers at firms with at least 500 employees. Firms with fewer than 500 employees are sampled each year. See Appendix C for a detailed discussion of sampling and weighting.[8]

We study college-educated individuals from cohorts born between 1960 and 1970. These are the youngest cohorts that we can follow from age 25 in 1985 until age 45 in 2013 (age 43 for the 1970 cohort). We begin following individuals from age 25 or the year after they graduate with their terminal degree, whichever comes later. We focus specifically on college-educated individuals for two reasons. First, the average labor force

[7] Qualitatively similar to an hourly wage variable, FTE monthly earnings records the contracted wage in full-time equivalent terms: i.e., the worker's earnings if he worked the standard number of full-time hours. For example, if a worker was contracted to work half-time, his FTE monthly earnings would be double his actual pay. By contrast, annual labor earnings is comprised of the product of contracted wage and hours, in addition to possible overtime pay and bonus compensation.

[8] Individuals who are currently employed at firms for which wage data is collected, but absent (e.g., on paid leave) in the survey month, do not appear in the Wage Structure Statistics. For these individuals, we interpolate wages in three ways. The first method, which we use when we report results, averages the wage from the prior and subsequent year. The other two methods assign either the prior year's wage, or the subsequent year's wage. All three methods yield similar results, and our findings are not sensitive to the choice of interpolation.

participation rate of college-educated women ages 25 to 45 is 95%, nearly as high as that of men, alleviating concerns about sample selection or changes in composition of workers over the lifecycle. Second, the increase in lifecycle gender wage differentials is more pronounced among the higher-educated, especially toward the top of the income distribution, making this a particularly interesting group to study. Correspondingly, whenever we construct firm-level variables, such as average wage growth at the firm, we also restrict our analysis only to college-educated employees at the firm.

Finally, since our focus is on the role of firms, and how wages change as workers move within and across them, we restrict our sample to individuals with degrees that are not associated almost exclusively with public sector employment in Sweden. Therefore, we exclude individuals with degrees related to teaching, medicine, and social work in the main analysis, as more than 85% of these workers are employed in the public sector. In Section 7, we provide results when all majors are included. In the paper, whenever we refer to "firms," we refer to all private sector and public sector employers.

Table 1 records summary statistics for the population of workers that we follow throughout the analysis. Women and men have similarly high labor force participation rates, exceeding 95% at all ages, although women are more likely to work part-time. Women are less likely than men to attain additional education after a bachelor's degree. More than 75% of individuals in our sample have had a child by age 45, and mean age at first birth is relatively high, at about 31.7 years for women, and 33.0 for men. About 36% of women work in the public sector, compared to 22% of men. This share does not change significantly with age for college-educated workers. On average women work at larger firms, and mean wage of college educated workers at the firms where women work is about 5% lower than at the firms where men work. Figure 1 summarizes lifecycle wage profiles for the workers in our population.

Relative to other countries, the Swedish labor market is characterized by high female labor force participation and relatively low gender wage differentials.[9] Sweden provides strong job protections for new parents and generous government-paid parental leave. New parents are entitled to 390 paid days of family leave at a 80% replacement rate (up to a cap), with an additional 90 days paid at a flat rate. While 60 days of leave are reserved for each parent, the remainder can be transferred freely between parents. During the period under study in this paper, parents have the right to work part-time (75% of full time), until their child turns eight. Prior to childbirth, part-time work is rare in Sweden, with about 3% of men and 6% of women working part-time.

[9]In 2015 the gender difference in median wages for all full-time workers in Sweden was 13.1 percent, compared to 17.9 in the U.S. See https://data.oecd.org/earnwage/gender-wage-gap.htm.

2.2 Key Variables

Our analysis focuses on three variables. The first, real annual wage growth, is simply a worker's year-on-year inflation-adjusted growth in wage (FTE monthly earnings). The other two, relative wage growth and promotion, are used to analyze within-firm mobility. We deviate methodologically from the literature by focusing on workers' movements through a firm's *wage hierarchy*, rather than up a traditional *career ladder*.[10] This approach is made possible by our data's worker-firm linkages, which allow us to observe individuals' wage increases relative to those of co-workers. With this information one can construct many possible variables to study moves through firms' wage hierarchies. We rely on two that we believe are transparent and easily interpretable, defined below.

A. Within-Firm Mobility Measures

Relative Wage Growth. The first variable is a continuous measure that compares the wage growth of an individual and the mean wage growth of other high-skilled co-workers at his or her firm in the same year.[11] Specifically, for each individual in our population who did not switch firms in the past year, we identify all other college-educated employees working at the same firm. Next, we calculate mean wage growth statistics for these co-workers by firm and year, for firms with at least ten employees. Relative wage growth is the difference between own growth and this firm-year average. The measure distinguishes between two reasons why an individual might experience high wage growth at their firm in a particular year. First, the worker may simply work at an especially high-growth firm, in which all or most employees experience large increases in compensation in a given year. Alternatively, the individual may experience substantial wage growth relative to other co-workers within the firm. The relative growth measure isolates this second factor, after accounting for systematic differences in average wage growth across firms in a given year.

Promotion. The second measure identifies years in which individuals experience especially large upward moves, to make it possible to study their extensive margin. Specifically, we transform the relative wage growth variable into a binary one, setting it equal to one whenever a worker realizes wage gains that are n log points higher than the average wage growth of his or her co-workers that year. We refer to these binary events as promotions. In all tables and figures, we also include the term "large relative wage increase" (LRWI), to remind readers of its formal definition. Throughout most of the analysis, we set n equal to 10 log points when constructing the binary measure, a fairly conservative threshold that focuses on large promotions. In Section 3 and Appendix A.6, we provide further detail about the choice of threshold n and the effect of varying it, as well as key results using alternative thresholds.

[10] Empirically, such moves are related. See e.g., Gibbons and Waldman (1999a), Prendergast (1999)).
[11] We also construct such a measure using instead median wage growth at the firm. Results are similar.

As occupational codes are available in our data, we can verify that promotions under our definition are commonly associated with changes in occupation: a change in code is about 1.8 times more likely in years that a worker is promoted. However, we do not use occupation codes to construct our promotion measure, for two reasons. First, it is difficult to distinguish between lateral and vertical moves for many changes in occupational code. Second, even detailed, four-digit codes are not sufficiently fine-grained to capture many moves up the career ladder, e.g., from analyst to project manager to division leader, CFO to CEO, or assistant to associate professor. Since a change in occupation code is neither necessary nor sufficient to indicate an upward move within the firm, we do not explicitly use this information to construct our measure.

B. Comparison to Conventional Measures

Our measures are non-standard and have, in our view, a number of advantages. To illustrate them, we describe existing approaches, and the substantial challenges inherent to population-wide analysis of within-firm mobility. Our discussion focuses on promotions as this is the dominant variable used to study within-firm mobility.

Relating our approach to existing ones would ideally begin with a definition of promotion using the "conventional" approach. However, there is no such formal definition in the existing literature.[12] Empirical studies of individual firms (e.g., Baker et al. (1994a,b)) focus on firms with mature organizational hierarchies and identify promotions based on formal job assignment. Methodologically, the advantage of this "single-firm" approach is that it can reliably isolate vertical from horizontal moves, and makes explicit what characterizes every "promotion" at the firm of interest, based on detailed institutional knowledge. However, the approach is difficult to scale, and the firm under study may not be representative. The "single-firm" approach is therefore not well-suited to provide an empirical characterization of promotion dynamics for the general population.

Studies that focus on the general population, by contrast, are representative but vague about what constitutes a promotion. Typically, they rely on self-reported promotions (e.g., McCue (1996)) or on job title data (e.g., Van der Klaauw and Da Silva (2011)). The challenge for these approaches lies in identifying promotions in a consistent fashion across thousands of heterogeneous firms. This is most evident for worker-reported promotions, the most commonly used variable, available in the PSID, NLSY, and many other datasets. Pergamit and Veum (1999) show that more than 50% of self-reported promotions in the NLSY involve no change in job position or duties, and more than one-fourth are not accompanied by a wage increase.

When characterizing promotions based on a change in job title, the key challenge is to

[12]Promotion have been modeled as a change in the technology used by the worker (Gibbons and Waldman (1999b), Bernhardt (1995)); the job's human capital intensity (Prendergast (1993)); or responsibility (Manove (1997)). By contrast, Lazear and Rosen (1981) model them as large wage increases for incentive provision.

isolate vertical moves in the firm, and distinguish them from lateral moves or other reasons for a title change.[13] A particularly instructive study by Van der Klaauw and Da Silva (2011) uses Portuguese administrative data that includes both employer-reported promotions, as well as job titles organized into hierarchical levels based on job complexity or responsibility. The authors find that only 30% of job title changes involving a change in hierarchical level are also considered promotions by the employer. Additionally, 40% of employer-reported promotions are not associated with any change in detailed job description. The measures thus capture either different types of within-firm mobility events; or, one or possibly both measures identify upward organizational moves only with significant error. However, evaluating what types of promotions are systematically omitted under each measure and how this may affect empirical results is not straightforward.

By comparison, our measure is defined by a clear metric and consistent about what it captures at every firm: large wage increases relative to one's co-workers, where large is explicitly measured. It provides a transparent benchmark for questions around population-wide within-firm mobility patterns, even if every promotion is not necessarily tied to a move up a career ladder. Our measure identifies events that are both comparable across firms and economically important. Using this variable, we can easily quantify differences in opportunities for such large promotions over the lifecycle or across firms or occupations.

Our promotion definition requires specifying a threshold. Concerns about this discretionary choice can be mitigated by reporting results for a range of thresholds, as we do in this paper. We note that definitions of promotion based on job duties or complexity also require the researcher to apply a threshold to identify a relevant change. This is often done implicitly and likely with substantial error, since an explicit threshold is difficult to specify for job responsibility, a multi-dimensional object.

We conclude with a brief discussion of relevant benchmarks in the literature. Promotions, as we will show, generate high wage growth of around 19% under our measure. This falls within the range of 8%-26% documented by Blau and DeVaro (2007) for employer-reported promotions in data about recent hires. Numerous studies analyze average gender differences in promotion rates. We do not compare our results with theirs, since they disagree about both the magnitude and sign of the differences.[14] McCue (1996) is the only study we know of that quantifies the contribution of promotions to wage growth over the lifecycle. She finds that promotions account for less than 15% of lifecycle wage growth for men and women, regardless of education. However, McCue indicates that

[13] Assessing upward vs. lateral moves using occupation codes is similarly difficult (e.g., a move from "business specialist" to "financial specialist"). Job title changes can also occur for administrative reasons, as when a firm alters part of its organizational structure, or changes how it reports some job titles in administrative data.

[14] E.g., McCue (1996), Cabral et al. (1981), and Blau and DeVaro (2007) find that women are promoted less. Booth et al. (2005) find that women have similar or higher promotion rates, while Giuliano et al. (2005) report identical rates. Estimates of the associated wage growth also differ. Blau and DeVaro (2007) find that women have higher returns to promotion, while Booth et al. (2005) find the opposite. Cabral et al. (1981), Giuliano et al. (2005), and McCue (1996) document similar returns. For a further review, see Blau and DeVaro (2007).

limitations to the self-reported PSID measure likely make this estimate a lower bound.[15] In our results, promotions account for around 40% of growth by age 45. For the other key findings we will present, it is difficult to identify relevant estimates using traditional measures. We are not aware of studies that quantify the effects of variables like firm selection, detailed educational background, or fertility events on promotion probability.

3 Empirical Analysis I: Gender Differences in Wage Growth

Why do women's wages grow more slowly than men's? In this section, we document the main differences in men's and women's wage growth patterns. We establish three stylized facts that are crucial for understanding gender differences in lifecycle wage trajectories.

3.1 Non-Parametric Analysis

Figure 2 provides a "bird's eye view" of annual wage growth in the cohort we study. It graphs the distributions of real annual wage growth for men and women between ages 25 and 45. Each point in the histogram is a person-year observation. Women's wage growth distribution (in color) is superimposed on men's distribution (outlined in black). Because the tails of the distributions are long and difficult to inspect visually, we collapse them to mass points at -0.25 and 0.25 log points.[16] As we are plotting real, rather than nominal growth, the incidence of small negative wage growth is relatively high. These observations correspond to nominal wage increases near zero.

Figure 2 features three similarities between men's and women's wage distributions, and one key difference. The first similarity is that, for both men and women, much of the distribution is concentrated at around 0 to 2% real wage growth. Thus, even though this population is high-skilled and young, annual growth is below 3% in the majority of years. The second similarity is that both genders are about equally likely to experience negative wage shocks. This indicates that, if women are switching to lower-paying firms after having children (e.g., Hotz et al. (2017)), these firm changes do not lead to a substantially greater incidence of high negative wage growth relative to men, at least for college-educated individuals.

Third, for both men and women, wage growth distributions are characterized by a large right tail. Importantly, this tail generates a considerable share of overall wage growth. Annual growth exceeding 15% accounts for only about one-fifth of observations, but around 50% of total real wage growth from age 25 to 45, which is on average 85 log points in this population. Around 90% of individuals experience at least one year

[15] PSID interviewers first ask whether an individual "changed jobs" in the past year. Only those responding affirmatively are asked whether they were fired, switched employers, promoted, etc. If individuals commonly interpret a "job change" to mean a change in employer, many promotions in this population will go unreported.

[16] For a more complete characterization of the tails, see Figure A.1 in the appendix.

of growth above 15%, and half experience at least two such years. As we show in the appendix, right-tail wage shocks are highly persistent (Table A.1).[17]

Lastly, we turn to the major gender difference in wage growth distributions. Figure 2 shows that the right tail of the distribution is significantly larger for men. Men are about 40% more likely to experience wage increases of 25 log points or more, and about 20% more likely to see their wages grow 15-24 log points. For women, this is offset by a greater number of periods with growth below 5%. This pattern constitutes the major gender difference in Figure 2 and, as will be shown in the subsequent analysis, is the key contributor to the divergence in men's and women's lifecycle wages. Men's higher probability of experiencing wage gains between 7.5-15% also contribute to wage growth differences, but play a smaller role.

The three similarities discussed above, as well as the major difference in the right tail, are summarized more formally in Table 2, which shows real wage growth at different percentiles for men and women. The higher-order moments, presented at the bottom, reflect that men's distribution is more right-skewed and characterized by a thicker tail, with a skewness and kurtosis of 0.79 and 20.1, compared to 0.58 and 18.2 for women. Men and women are equally likely to experience negative wage growth, with approximately zero real growth at the 25th percentile of the person-year wage growth distribution.

Figure 3 graphs wage growth distributions for several sub-populations to document that the patterns in Figure 2 are a systematic feature of men's and women's wage growth, and not readily explained, for example, by women's lower probability of entering high management positions at older ages or after childbirth, or by men's and women's participation in different fields. Figure 3 shows that women's lower probability of experiencing a period of right-tail wage growth is observed within all fields, at younger ages (under 30), where it is even more pronounced, and for individuals who never have children. We also observe the same pattern when we drop the first three years after graduation, to exclude observations in which individuals may simply be switching from temporary positions into more permanent roles.

To summarize, men's and women's wage growth distributions are similar in most aspects, including incidence of negative wage growth. In particular, both distributions are characterized by right-tail annual wage shocks that are relatively infrequent, and drive a substantial share of lifecycle wage growth. The primary gender difference is that women are significantly less likely to experience these years of right-tail growth (fact 1).

[17]By contrast, shocks to annual earnings exhibit high reversion (Guvenen et al. (2014)), since they are driven in large part by fluctuations in hours, e.g. following job loss/recovery. Annual earnings also include one-time bonus payments, not contained in our wage measure. Guvenen et al. find that the annual earnings growth distribution is left-skewed. This is true also in our data, while the wage growth distribution is right-skewed.

3.2 Characteristics of Right-Tail Growth

High wage growth is common during firm changes (e.g., Burdett (1978), Topel and Ward (1992), Abowd et al. (1999)) and during within-firm promotions (e.g. Doeringer and Piore (1971), Baker, Gibbs and Holmstrom (1994a,b)). In this section, we analyze how growth during firm changes and during firm tenure contribute to gender differences in the right tail and, correspondingly, to differences in overall wage growth.

A. Growth During Firm Changes vs. Within-Firm Growth

Table 3 summarizes characteristics of annual wage increases of different sizes. It features two findings. First, the majority of right-tail wage shocks occur within-firm. Column 2 shows that firm changes are more common during high-growth years than during low-growth years, as expected. Nevertheless, they constitute only 29.7% of observations exceeding 25 log points, and 21.3% of observations between 15 and 24 log points.

Second, gender differences in the incidence of right-tail wage growth are substantially more pronounced within-firm, as a comparison of columns 3 and 4 indicates. Men are about 26% more likely than women to experience a wage increase of 25+ log points when they change firms, but 47% more likely to experience such a wage increase during firm tenure. For increases of 15 to 24 log points, the gender differences are smaller, but follow a similar pattern: men are about 5% more likely to experience such growth during firm changes, and 23% more likely during firm tenure.

A decomposition of total gender differences in lifecycle wage growth confirms the finding that firm changes contribute to growth differences, but play a secondary role. We proceed with minimal assumptions, starting from the following identity governing average cumulative wage growth in log points by age a:

$$\sum_{t=26}^{a} \Delta W_t^k = \sum_{t=26}^{a} \sum_j P_t^k(j) \Delta w_t^k(j), \qquad k = m, f. \tag{1}$$

Here, j corresponds to either the immediate wage growth associated with switching firms, or to within firm-growth in subsequent years of firm tenure. We keep the notation general, as we will use this identity also in subsequent decompositions. $P_t(j)$ is the probability with which j occurs at age t, and $\Delta w_t(j)$ is the average wage growth associated with j, where growth is measured as the difference in log wages between t and $t-1$. Note that the gender differences in cumulative lifecycle wage growth by age a, in this case by age 45, is simply $G^a = \sum_{t=26}^{a} \Delta W_t^m - \sum_{t=26}^{a} \Delta W_t^f$.

The part of the gender gap G^a that is attributable to the sources of wage growth j (i.e., firm changes or within-firm growth) can be expressed as

$$g_j^a = \sum_{t=26}^{a} P_t^m(j) \Delta w_t^m(j) - \sum_{t=26}^{a} P_t^f(j) \Delta w_t^f(j), \tag{2}$$

where $\sum_j g_j^a = G^a$. The outcome of this decomposition indicates that 26.5% of differences in cumulative wage growth by age 45 are due to differences in the wage growth associated with firm changes.[18] To summarize, gender differences in both right-tail and overall wage growth are driven primarily by differences in within-firm wage growth (fact 2).

B. Average Growth at a Firm vs. Within-Firm Mobility

Lastly, we investigate the characteristics of right-tail, within-firm observations: specifically, whether they are commonly attributable to working at a firm with high average wage growth, or whether they represent years in which workers move up through a firm's wage distribution. For this purpose, we employ the two measures described in Section 2.2. The first is relative wage growth, i.e. the difference between a worker's total wage growth in a given year and the average wage growth of other high-skilled workers in the same firm and year. The second is our binary promotion variable, defined as wage gains that are n log points higher than the average wage growth at the firm that year. Table 4 summarizes these variables for wage increases of different sizes, focusing only on years in which workers did *not* switch firms.

Column 1 records, for reference, how much more likely men are to experience a wage increase of a given size, compared to women. Column 2 shows that even in years that individuals experience large wage increases, average wage growth at their firm tends to be moderate. Across categories, average firm wage growth varies only from 2.2 to 4.0%. Instead, the majority of the variation in individuals' annual wage growth is attributable to relative wage growth (column 3). This indicates that the observed gender differences in right-tail growth correspond to gender differences in the incidence of years of high within-firm mobility. Column 4 shows this explicitly. Most right-tail events correspond to years of especially high relative wage growth, in excess of 10 log points. Such large relative wage increases, or "promotions," for the baseline threshold of $n = 10$ log points, characterize 98% of observations exceeding 22.5 log points, and 84% of wage growth observations between 15 log points and 22.5 log points. By contrast, promotions characterize fewer than 2% of observations in the 0 to 7.5% range.

Finally, we conduct a decomposition similar to the previous one, again relying on equations (1) and (2). This time, instead of simply considering firm changes vs. within-firm growth, we additionally separate within-firm growth into promotion and non-promotion growth. Thus, the three categories – firm changes, promotions, and non-promotion growth – are exhaustive and mutually exclusive.

Figure 4 graphs the outcome of this decomposition, at each age a. Of the three sources of wage growth, cumulative differences in promotion-related growth are the dominant

[18] This decomposition is by design simple, non-parametric and not tied to a specific wage setting model (see Postel-Vinay and Robin (2006)). Card et al. (2016) estimate a parametric, two-way individual and firm fixed effect wage model. Different returns to firm changes account for about 20% of gender gap in wage levels according to their estimates. This would suggest that their estimates are similar to, or somewhat below ours.

driver of the increasing lifecycle gender wage differentials. At all ages, these differences in promotion-related growth are of the first order, and averaged over all ages account for 74.6% of the differences in wage growth. Differences in growth associated with switching firms account for about 26.5% of the increase in the wage gap, as quantified previously. Non-promotion growth contributes on average negatively (-1.0%), especially after age 39. This decomposition indicates that gender differences in lifecycle wage growth are driven primarily by differences in within-firm mobility (fact 3).

To summarize, men have higher wage growth for one main reason: they experience more years of right tail wage growth. This difference occurs primarily during firm tenure, not during firm changes. It is accounted for by differences in wage growth relative to other co-workers, not by differences in average wage growth at the firm. More specifically, it is driven by differences in the probability with which men and women experience years of exceptional wage growth relative to other co-workers ("promotions").

We conclude with a comment. For clarity, we reported results above only for one threshold value used to characterize the binary promotion variable. We continue to report results only for this baseline threshold, $n = 10$, which focuses the analysis on large within-firm mobility events. While the choice of threshold affects quantitative estimates, it does not affect the qualitative results and dynamics that we will document, for a large range of values for n. Further details about the choice of threshold are provided in Appendix A.6. The appendix also documents key results for alternative thresholds.

4 Empirical Analysis II: Gender Differences in Within-Firm Mobility

Based on the evidence in Section 3, we analyze next the lifecycle incidence and characteristics of promotions, as defined by discrete moves through firms' wage hierarchies. We then quantify the importance of human capital, sorting across firms, occupation, and hours worked for the "promotion gap." Lastly, we study effects of parenthood.

4.1 Within-Firm Mobility Over the Lifecycle

Table 5 summarizes the growth associated with promotions, firm changes, and interim (non-promotion) periods for men and women, and the probabilities with which they occur over the lifecycle. As expected, the table documents large differences in wage growth across the three categories. The wage gain from switching firms is around 8.4 and 9.5 log points for young women and men, and declines to 4.1-4.8% later in the lifecycle. Men and women experience low wage growth in interim, non-promotion years, between 1.1-3.3%, depending on age. By comparison, the wage gains following promotion exceed 18 log points (19.7%) at all ages, and are 1-1.5% higher for men.

Consequently, just one missed promotion corresponds to significant wage losses. For this reason, promotions also account for a significant share of lifecycle wage growth, about

36%-43% by age 45 (Figure 5). Importantly, the fact that associated wage gains exceed 18 log points for both men and women indicates that most of the gender difference in promotion-related growth in Figure 4 is driven by difference in the *probability* of experiencing a promotion, i.e. by the extensive margin. Equalizing wage gains conditional on promotion would only moderately reduce the wage growth gap attributable to promotions in Figure 4 to 69.5%, from 74.6%.

A second set of findings in Table 5 is about the probability with which promotions occur. First, promotions are most common early on in the lifecycle and become more infrequent with age. Second, women change firms at almost identical rates as men. Therefore, most of the gap in promotion probability is due to lower promotion rates, *conditional* on staying at the firm. Figure 6 documents this pattern in detail. Panel A shows that men and women switch firms at nearly identical rates over the whole lifecycle – women switch at only marginally lower rates between ages 30 and 35, when childbirth is most common in Sweden. By contrast, Panel B shows that women have significantly lower promotion rates at all ages, especially early in the lifecycle, when the promotion gap is largest both in absolute and percent terms. At ages 26 to 30, women have a 13.7% probability of experiencing a promotion in a given year, compared to 17.8% for men, about a 23% difference.

The final finding in Table 5 is that during both firm switches and non-promotion periods, the gender gap in annual growth is modest and favors men early on. It subsequently declines and, notably, reverses after age 40. This pattern clarifies why non-promotion wage growth contributed negatively at older ages to gender differentials in Figure 4. Late in the lifecycle, non-promotion periods are extremely common, and in those years women experience higher wage growth associated with such events than men. We will return to this finding in our theoretical analysis in Section 5.

The patterns in Table 5 and Figure 6 are in line with the non-parametric results. In Section 3, we showed that wage growth distributions are similar for men and women in most aspects. As we show above, men and women switch firms at almost identical rates; additionally, their wage growth conditional on changing firms or being promoted, as well as wage growth in interim years, differs modestly. By contrast, a major difference is the rate with which men and women experience promotions, in line with the differences in right-tail growth documented in Section 3. Missing just one such within-firm event implies a wage loss of more than 15%, contributing significantly to lower cumulative wage growth over the lifecycle.

Finally, we note that a key finding from Section 3 – that differences in promotion-related growth account for about 75% of gender differences in lifecycle wage growth (Figure 4) – is not simply a mechanical consequence of the role of promotions in wage growth. Promotions do account for a sizable share of cumulative wage growth at each age – about 36%-43% (Figure 5) – but nevertheless substantially less than three-quarters.

4.2 Human Capital, Firm Characteristics, Occupational Choice, and Hours Worked

Men and women may differ on average in years of post-secondary education and field of study, as well as in other potentially important factors that could affect promotion rates, such as firm characteristics. In this subsection, we analyze the role of these factors in accounting for differences in promotion probability. To describe their role succinctly, we rely on the following regression:

$$y_{ift} = \mu + \alpha \cdot female_i + X_{it}'\beta + \pi_t + \gamma_f + \varepsilon_{ift} \tag{3}$$

The outcome variable y_{ift} is an indicator corresponding to whether or not individual i in firm f and year t received a promotion. X_{it} is a set of covariates, and π_t and γ_f are year and firm fixed effects. As we add covariates of interest, we primarily describe in this subsection the change in the coefficient α, which corresponds to the gender difference in promotion rates. To illustrate why the coefficient under study changes with the addition of the various controls, we report more detailed results in Appendices A.2-A.4 and refer to them where relevant. We focus on individuals under age 35, since these are the ages at which promotions are most common. However, in Section 4.3 we will analyze key drivers of the promotion gap dynamically, by age and time to birth.

We begin by considering gender differences in human capital characteristics. Table 6 shows that the baseline gender difference in promotion rates is 4.0 percentage points, controlling only for time fixed effects (column 1). Relative to the promotion rate of men under age 35, equal to 20.1%, this gap corresponds to a difference of about twenty percent. Controlling for years of post-graduate education, field of study, and experience in column 2 increases the estimated gender difference in promotion rates to 4.4. This somewhat surprising increase is explained by the control for field major, as men in our population are more likely to have science or engineering degrees, which have lower promotion rates, compared to business or law, fields that are more common among women.[19] Therefore, once this difference in choice of major is controlled for, the gap further increases. Accounting for years of tenure at the firm has little effect (column 3), since men and women switch firms at similar rates, as we showed earlier.

Next, we consider systematic differences across firms in the probability of promotion. In Column 4, we add firm fixed effects to account for gender differences in sorting across firms with different promotion opportunities. Inclusion of firm fixed effects reduces the coefficient on female by about 10%, from 4.4 p.p. to 4.0 p.p.[20] The small decline in the

[19] In science or engineering, about 15% of men below 35 are promoted annually, compared to 22% of men in business or law.

[20] We have also re-estimated this regression, but restricting the sample to include only firms which employ at least one man and at least one woman under age 35 from the cohort we follow; or, alternatively, to firms that are no more than 90% male or 90% female. The coefficients on female are identical to the fourth decimal.

coefficient is surprising, given that we observe gender differences in sorting across firms with different promotion opportunities in the data, as we show in Appendix A.2, and in line with findings by Hotz et al. (2017) and Kleven et al. (2019) that women work at firms with potentially less career mobility. For example, at firms where fewer than 5% of workers are promoted each year according to our measure, women constitute 47% of all workers. At firms where more than 10% of workers are promoted annually, they represent only 37% of workers.

The reason why controlling for this difference in sorting does not substantially reduce the coefficient in column 4 is that the within-firm gender difference in promotion probability is observed across all firm types, with even more pronounced differences at firms with higher promotion rates. At firms where at least 10% of workers are promoted each year, women in our cohort are on average about 6.2 p.p. less likely than men to be promoted, compared to the average gap of about 4.0 p.p. In Appendix A.2, we show that equalizing men's and women's distribution across firms affects the estimated promotion gap much less than equalizing their promotion rates within-firm, in line with our regression results above. Note that firm fixed effects also control for a variety of fixed characteristics such as industry, indicating that these differences are not of first order in accounting for the gender difference in promotion probability.

Controlling for major already controls in large part for occupational choice. However, women may still be more likely to choose an occupation with, for example, better non-wage amenities and fewer opportunities for upward movement, even conditional on major. In Appendix A.3 we show that women are indeed somewhat more likely than men to work in clerical or administrative support positions, which are typically associated with flatter career progression. However, women are also more likely to work in occupations in line with their major choices such as business and law, with high promotion rates. The addition of three-digit occupation codes reduces the coefficient on female to 3.9 p.p., which is not statistically different from the baseline coefficient of 4.0.[21]

To analyze the effect of hours worked, we rely on two variables. The first variable, contracted hours, is administratively recorded, and provides detailed information about part-time work, on a scale from 1% to 100% full-time equivalent (FTE).[22] To complement this variable, which does not capture hours worked above 100% FTE, we also construct a proxy measure, which divides annual labor income by contracted wage. The drawback of this measure is that annual labor income can potentially include bonus pay or other compensation, which need not reflect higher hours worked. However, its advantage is that it can capture both overtime work, as well as time taken off for child illness or parental leave, as such leave is paid for by the government, and not reported as annual

[21]Unless otherwise indicated, we omit occupation controls in the remainder of the analysis, since many differences in occupation are outcomes of promotions, complicating the interpretation of our results. Because we account for major, adding occupation controls only has small effects on any coefficients we estimate.

[22]For example, an individual contracted at 80% FTE works 32 hours per week.

labor income. We provide further details its construction in Appendix A.4. To avoid introducing a mechanical correlation with our promotion measure, as explained in the appendix, we use a lagged version of the hours proxy. Using this measure, the average gender difference in weekly hours worked over all age groups is about 5.4 hours (Table A.4). Part-time work (35 hours or less) is most common among women ages 36-40, with about 23% of women in that age group working part-time. Women are also less likely to work 48 hours or more weekly (Table A.5), a category of weekly hours associated with higher promotion rates, in line with findings by Gicheva (2013).

Table 7 shows the effect of hours worked. In column 2, we control for both aforementioned measures of hours worked; for part-time history; and, to better capture reductions in labor supply associated with parental leave, we include an additional indicator for whether the worker has a child that was born in the current or prior year, and its interaction with an indicator for being female. These controls reduce the promotion gap to 2.1 percentage points, accounting for slightly less than half of the baseline gap in column 1. The final controls – related to childbirth in the current or prior year – are important. With hours worked and part-time history alone, the reduction in the baseline coefficient is more modest, declining to only 3.2 p.p. (column 3). In the next subsection, we examine these findings in greater detail by considering dynamic effects of parenthood and hours worked. We also consider corresponding dynamics in the residual gender promotion gap.

4.3 Dynamic Effects of Parenthood and Gender

The high-skilled Swedish women we study are highly attached to the labor force, with about 95% either working or taking parental leave at all ages between 25 and 45. However, as in other countries, their hours worked change substantially after childbirth (see Angelov et al. (2016)). From graduation until first birth, 90% of future mothers and 91% of future fathers work full-time. After childbirth, virtually all women take at least six months of parental leave, and many take more than a year. About 28% work part-time in the five years following first childbirth, compared to 6% of men. In this section, we analyze how childbirth and related labor supply reductions affect gender differences in promotion probability over the lifecycle.

A. Total Promotion Gap vs. Motherhood Penalty

To analyze the effects of childbirth, we consider two sets of estimates. These estimates are related, but capture distinct phenomena. The first set of estimates captures the *total* difference in annual promotion probability, by year relative to first birth. It constitutes a simple data description. The second set of estimates tries to isolate the difference in promotion probability specifically attributable to time to childbirth, rather than any other gender-related differences over the lifecycle. The literature refers to the latter set

of estimates as the dynamic "motherhood penalty," by year relative to first birth (e.g., Angelov et al. (2016), Kleven et al. (2019)).

To measure how the total probability of receiving a promotion changes with years k relative to first birth, where $k = -5, -4, ...10$, we run the following regression:

$$y_{it} = \mu + \sum_{k \neq -1} \theta_k D_{it}^k + \sum_k \alpha_k D_{it}^k \cdot female + X_{it}'\beta + \pi_t + \varepsilon_{it}. \tag{4}$$

D^k are a set of time-to-birth dummies, equal to one if an individual is k years from first birth.[23] The year of first birth corresponds to $k = 0$. This specification is similar to the one used in equation (3). The difference is that in lieu of a single indicator variable for female, the indicator is now interacted with time to first birth, to show how the promotion gap evolves dynamically. Thus, α_k measures the total male-female difference in promotion probability in every year relative to first birth.

To isolate only the motherhood penalty, we follow Kleven et al. (2019) and estimate an equation similar to the one above, *separately* for men and women:

$$y_{it} = \mu^g + \sum_{k \neq -1} \theta_k^g D_{it}^k + X_{it}'\beta^g + \pi_t^g + \varepsilon_{it}, \tag{5}$$

where g corresponds to gender. The coefficients θ_k^f and θ_k^m are scaled relative to the year before first birth, since $k = -1$ constitutes the omitted category. They represent, respectively, the dynamic effects of motherhood and fatherhood, after all other lifecycle factors have been controlled for. The "penalty" for mothers in each year k is simply the difference in the two parenthood effects, $\theta_k^f - \theta_k^m$, which is by construction zero in the year before first birth.

To provide an intuition for the difference between the two sets of estimates, it is easier to rewrite equation (5), which is estimated separately by gender, instead as one regression in which all variables are interacted fully with female:

$$\begin{aligned} y_{it} = \tilde{\mu} + \eta \cdot female_i + \sum_{k \neq -1} \tilde{\theta}_k D_{it}^k + \sum_{k \neq -1} \tilde{\alpha}_k D_{it}^k \cdot female + \\ + X_{it}'\tilde{\beta} + X_{it}'\tilde{\beta}^f \cdot female + \tilde{\pi}_t + \tilde{\pi}_t^f \cdot female + \varepsilon_{it}. \end{aligned} \tag{6}$$

For ease of comparison, we have kept the same notation as in equation (4), and added a tilde to the coefficients to distinguish the two models. It is now easy to see that equations (4) and (6) are identical, except that the control variables and time effects are interacted with gender. This interaction allows, for example, returns to experience or education to differ for men and women, for reasons not related to childbirth. For instance, if women in general exert less effort or are less competitive, leading to fewer

[23] Observations corresponding to more than five years before first birth are assigned to the $k = -5$ category. Observations more than 10 after first birth are assigned to the $k = 10$ category.

promotions, or if employers discriminate against all women of childbearing age, such systematic gender differences would be captured in the estimates of η, $\tilde{\beta}^f$, and $\tilde{\pi}_t^f$. With all gender differences not related directly to time to first birth now controlled for, $\tilde{\alpha}_k$ isolates the dynamic motherhood penalty by time to first birth, relative to $k = -1$, rather than just describing the total gap, as α_k from equation (4) does.

For all the above regressions, the sample includes only men and women who have ever had children, or about 75% of our original population. To focus on the within-firm promotion gap, we also control for firm fixed effects throughout the analysis. In practice, however, controlling for firm has limited effect, in line with our previous findings that most of the promotion gap is a gap between men and women at the same firm.

We begin with the total gender difference in annual promotion probability. Figure 7 graphs the estimated coefficients α_k for the baseline specification in equation 4, as well as for specifications that control for part-time history and detailed hours worked using the hours proxy variable. Figure 7's immediate and most obvious feature is that women's probability of receiving a promotion drops dramatically in the year of and immediately following first birth. In each of those two years, women are 8.4-9.0 p.p., or about 53% less likely than men to receive a promotion, as seen in Table 8. These years coincide with parental leave for the majority of women in Sweden.

Several years after first birth women continue to have lower promotion rates. Hours worked and part-time history account for part of the post-birth gap, especially six or more years after first birth. However, a majority of the gap in Figure 7 that occurs two or more years after birth is due to another large 50% reduction in promotion probability for women at time of second birth, pictured in Figure A.6.[24] About two years after second birth the promotion gap decreases dramatically and remains low, as the modal number of children in Sweden is two.

The final distinctive pattern in Figure 7 is that before first birth, women have a 2.2-3.8 p.p. (about 16%) lower probability of promotion than men. Notably, part-time controls do not affect these pre-birth estimates, since few women work part-time in these years.[25] For this reason, part-time work and part-time history only account overall for about 21% of the gap in cumulative promotions by age 45. Controlling additionally for hours worked using our proxy variable, which better captures variation in hours worked above the full-time threshold, reduces the pre-birth estimates on average by 0.8 percentage points in Figure 7, although the differences are statistically significant only in one of the years.

[24]Promotion rates, in absolute terms, decrease with age. Correspondingly, the gender differences in absolute terms also decrease with age. This explains why the promotion gap at second birth is not lower in relative terms than at first birth, but is lower in absolute terms.

[25]Low part-time rates early in the lifecycle are not surprising, as the structure of the Swedish parental leave system strongly incentivizes high hours and rapid human capital accumulation prior to birth. Workers on leave are typically compensated at an 80% replacement rate, based on their own earnings in the two years prior to birth. If births are spaced sufficiently close together, earnings prior to first birth also determine parental leave benefits after the second birth. For institutional details about the parental leave system, see Section 2.

Next, we plot the motherhood penalty in Figure 8A, together with the total male-female promotion gap for reference. As expected, motherhood penalty accounts for most of the dramatic drop in promotion probability at first birth. Additionally, the motherhood penalty is persistent and accounts for the entire promotion gap several years after first birth. Recall that the motherhood penalty is equal to the difference between the effects of motherhood and fatherhood on promotion probability, i.e. the estimates of θ_k^m and θ_k^f from equation (5). The latter estimates are graphed in Figure 8B. The figure shows that, in fact, both men and women experience a drop in promotion probability after having children. For men, the initial decline is much smaller and occurs slightly later, about one to two years after childbirth. This corresponds to the time when women in Sweden usually finish their leave and men take their "daddy-months."

The previous figures highlight two more findings. Neither Panels A nor B of Figure 8 show any substantial uptick in promotion probability for women following the dramatic decline around the year of birth. This suggests that the "missed" promotions after first birth are not readily recovered. Additionally, several years prior to first birth, the motherhood penalty is either slightly positive or approximately zero. Thus, the motherhood penalty accounts for only a portion of the total promotion gap, 49.7% as graphed in Figure 8A.[26] This also confirms results from Table 7, which showed that hours worked, part-time history, and immediate effects of birth events – the key variables associated with a motherhood penalty – account for only about half of the promotion gap.

B. Dynamic Gender Penalty, by Time to First Birth

The motherhood penalty accounts for different shares of the total promotion gap in different years. Consequently, residual gender differences in promotion probability also exhibit a distinctive dynamic pattern. Figure 9A plots the difference between the total promotion gap and the motherhood penalty. The series captures the gender penalty in promotion by time to first birth, net of the motherhood penalty and after controlling for detailed human capital characteristics and firm fixed effects. Thus, it illustrates the lifecycle incidence of the unaccounted-for promotion gap. It points to three important patterns. First, the gender penalty is quantitatively large several years before birth (both in absolute and percent terms). Second, it declines over time, going to zero about 6 to 7 years after first birth, when women are on average 37 to 38 years old. Third, it eventually reverses (becomes positive) 10 years after first birth, when women are 41.

For individuals who never have children, the same dynamic pattern is observed. Panel B plots the coefficients on interactions between indicators for *female* and *age* in our standard regression, with all controls for human capital, firm fixed effects, and hours

[26] This estimate should be interpreted with care, as it relies on the standard normalization of the motherhood penalty to zero in the year prior to first birth. See Appendix A.5 for a discussion. Normalization does not affect the dynamics or qualitative results we document in this section, or any quantitative estimates based on the total promotion gap by time to birth.

worked. For childless individuals, the residual gender penalty is also initially negative and decreases with age. It goes to zero between ages 36 and 40 and similarly reverses (favors women) after age 40. Thus, sizable dynamic gender penalties are observed at younger ages for both women who ever children and those who remain childless, but reverse around the end of women's childbearing years.

Summary. We conclude with a summary of the key empirical facts about wage growth and within-firm mobility. We have shown that the overwhelming majority of gender differences in lifecycle wage growth are driven by differences in the incidence of promotions – persistent, sizable moves through firms' wage hierarchies. The gap in annual promotion probability is largest early in the career and more akin to "broken bottom rungs" (McKinsey and LeanIn.Org (2019)) than to "glass ceilings," which have been traditionally emphasized in the academic literature (Albrecht et al. (2003)). The gap is also first and foremost a within-firm gap, between observationally similar male and female co-workers.

What accounts for this gender gap in promotion between co-workers? Our estimates point to the following decomposition. Women's "missed" promotions in the year of and year immediately following birth events are the biggest factor. They account for 40% of the cumulative promotion gap by age 45, for women who ever have children. These are years when Swedish women are typically on parental leave. This finding implies that a large share of the gap in wage growth by 45 is incurred over a strikingly short period of time. It also indicates that much of the drop in women's hourly wages relative to men's shortly after birth (e.g. Kleven et al. (2019)) is explained by the fact that in those years, women fail to get the exceedingly large wage increases that drive an important share of wage growth. This generates a rapid and persistent gender divergence in wages. We find no evidence that women can readily recover these "missed" promotions.

Part-time work accounts for 21% of the cumulative gap by 45. By comparison, the promotion gap incurred prior to first birth accounts for more, about 30% of the cumulative promotion difference by 45. One reason why part-time work does not play a more prominent role is that promotion probabilities are highest at young ages, well before Swedish women enter part-time work at significant rates. Prior studies found that part-time work significantly increases gender earnings differentials. Our results do not contradict this: Lower hours worked, at the very least, reduce annual earnings mechanically. However, as a driver of lifecycle differences in hourly wage growth and, specifically, promotion, the role of part-time work is non-negligible, but decidedly secondary.

Finally, we observe a quantitatively large residual gender gap in promotion. This gap is dynamic: it is greatest early in the lifecycle, but reverses to favor women after the end of their childbearing years, around age 40. Its magnitude and dynamics are similar for women who become mothers and those who remain childless. One interesting feature of this reversal is that it mirrors our earlier findings that women's wage growth in years when they are not promoted or change firms is initially lower, but exceeds men's after 40.

We consider many of these key findings together in the theoretical framework we develop in the next section.

5 A Model of Gender Differences in Careers Within Firms

To help interpret the main facts of the paper, we develop a theoretical model of men's and women's careers within firms. We build on one of the workhorse models in this literature, by Gibbons and Waldman (1999b). Several classic models of careers can potentially generate a key fact documented in Section 3, that workers commonly experience low annual growth, with periodic large wage increases within-firm.[27] We build on Gibbons and Waldman's model, as it generates many insights with simple math and rich intuition.

To keep our exposition brief, we focus on the simplest version of their model, with full information. Readers familiar with their model will recall that this full information version generates moderate wage growth at time of promotion. This wage growth can be amplified either by incorporating learning by the employer about the worker, as is developed in the same paper by Gibbons and Waldman;[28] or by incorporating compensation for effort associated with different jobs, as we do. The addition of either feature complicates our exposition, without providing further intuition for our findings. Consequently, we present the simplest version of the model below, but provide proofs in Appendix B for the augmented model, which generates large wage increases associated with promotions, in line with the data. All results hold identically for the augmented model.

5.1 Benchmark Model

Before introducing gender, we describe the benchmark model by Gibbons and Waldman (1999b), henceforth GW'99. It is identical to the full information setting in GW'99, with the exception that they consider discrete worker types, while we consider a continuum.

General Environment. Firms are identical, there is free entry into production, and workers can change firms costlessly. Workers and firms are risk-neutral and have a time discount rate of zero. A measure one of workers, indexed by i, work for $T+1$ periods and are characterized by innate ability $\theta_i \in [0,1]$, which has a uniform distribution. Workers accumulate labor market experience x_{it}. Innate ability and experience together determine the worker's *effective* ability η_{it} in period t:

$$\eta_{it} = \theta_i f(x_{it}),$$

with $f'(\cdot) > 0$, and $f(0) = 0$. Both innate and effective ability are observed by firms.

[27] Examples of such models include Gibbons and Waldman (1999b), Lazear and Rosen (1981), and Harris and Holmstrom (1982), among others reviewed in Gibbons and Waldman (1999a).

[28] This feature can be even further strengthened by incorporating private information by the employer, as in Waldman (1984). See also Gibbons and Waldman (1999a).

Worker output. Firms consist of $J+1$ jobs, and take labor as the only input. The output of a worker assigned to job j is linear in η_{it} and equals

$$y_{ijt} = d_j + c_j \eta_{it},$$

where for all $j = 0, 1, ..., J$, parameters c_j and d_j are positive, $c_{j+1} > c_j$, and $d_{j+1} < d_j$.

The job at which a worker will be most productive depends on his or her effective ability. A worker with no experience (who therefore has η_{it}=0) is most productive in job 0, since $d_0 > d_1 > ... > d_J$. Workers' experience increases by one unit after each period of work. As a result, their effective ability increases over time, and they may become more productive at other jobs. The effective ability at which a worker is equally productive at jobs 0 and 1 solves $d_0 + c_0\eta_{it} = d_1 + c_1\eta_{it}$. We denote this solution as $\overline{\eta}^1$, since it defines the threshold value above which a worker starts to be more productive at job 1. Similarly, $\overline{\eta}^2$ solves $d_1 + c_1\eta_{it} = d_2 + c_2\eta_{it}$, and so on, for the remaining jobs. It is assumed that parameter values for c_j and d_j are such that $\overline{\eta}^1 < \overline{\eta}^2 < ... < \overline{\eta}^J$, so that higher level jobs require higher effective ability. Efficient assignment implies that workers are assigned to job 0 if $\eta_{it} < \overline{\eta}^1$, to job 1 if $\overline{\eta}^1 \leq \eta_{it} < \overline{\eta}^2$, and so on.

Promotions. In period 0, workers enter the labor market with no experience and are all hired optimally into job 0. Promotions to higher jobs are possible starting in period 1. Firms make promotion decisions at the start of the period. Individuals with the highest innate ability, θ_i, will be the first to be promoted as they accumulate experience, since their effective ability increases most rapidly over time. This is illustrated in Figure 11A.

Threshold values for promotion, defined by effective ability, can be restated in terms of innate ability and years of experience. This feature is useful for characterizing the share of all individuals who are promoted in any given period. We will let $\overline{\theta}^j_\tau$ refer to the minimum value of innate ability required to be promoted to job j, for someone with $\tau \geq 1$ years of experience, where $\overline{\theta}^j_\tau = \frac{\overline{\eta}^j}{f(\tau)}$. Innate ability thresholds have two properties. They are higher for higher-level jobs, holding years of experience fixed: $\overline{\theta}^J_\tau > ... > \overline{\theta}^2_\tau > \overline{\theta}^1_\tau$. Additionally, innate ability required for promotion to a given job j is lower for those with more experience, i.e. $\overline{\theta}^j_1 > \overline{\theta}^j_2 > ... > \overline{\theta}^j_T$ (see Figure 11B).

Equilibrium. In this frictionless setting, firms make zero profit but optimize the efficiency of output. Assignments to job tasks for all workers in all periods are efficient, and workers are paid a per-period wage $w_{ijt} = d_j + c_j\eta_{it}$. Assuming a function for human capital accumulation $f(x_{it})$ such that it is efficient for individuals with the highest innate ability, $\theta_i = 1$, to be promoted exactly once each period, it also follows that workers, on average, move up the career ladder as they accumulate experience, as GW'99 show.[29]

Parameters. We consider a restricted set of parameters for job technologies and human capital accumulation, c_j, d_j and $f(x)$, to focus the analysis on the most insightful

[29] GW'99 also prove that this model generates serial correlation in promotions and other important wage dynamics. As we focus on a new set of findings, we refer the interested reader to their original study.

cases, similar to GW'99. We assume that the parameterizations satisfy the following conditions. First, the highest ability individuals are promoted exactly once in each period, if they work every period. Second, the lowest ability individuals are never promoted.

5.2 Model with Gender and Birth-Related Labor Supply Reductions

To study gender differences, we extend the model in two ways. First, we introduce a lifecycle for workers with a limited childbearing period. Second, we incorporate a cost to employers associated with worker labor supply reductions after birth, taken mostly by women. The latter feature is motivated by evidence from Sweden that firm wage bills increase substantially during workers' parental leaves (Ginja, Karimi, and Xiao (2020)).

A. Description of Model Features

Lifecycle Structure. The worker's lifecycle has three phases. In the pre-childbearing phase, probability of having a child is low. We assume the probability is zero, but this can be readily relaxed. Next, in the prime childbearing phase individuals have a strictly positive probability of birth. In the third and final phase individuals are no longer fertile, and no births can occur. We set the number of periods to $T = 3$, as shown in Figure 12, so that each phase lasts for one period.[30] In period 0, when workers enter the labor market and before promotions can occur, we also assume a zero probability of birth.

Births affect labor supply. In period 2, share p_f of women and share p_m of men have children and do not work, with $p_f > p_m$. We normalize $p_m = 0$ without loss of generality.[31] The assumption that women reduce their labor supply to zero in the period they give birth is motivated by the institutional context in Sweden, which has a generous government-paid parental leave program, with women typically taking around one year of leave or more. We make two further simplifications that we later relax. In all periods other than $t = 2$, workers supply a unit of labor. Additionally, the probability of having a child and taking a leave is uncorrelated with ability.

Figure 12 describes the timing in each period. At the start of the period, firms observe whether a worker gives birth and will take a leave that period. Next, firms make their promotion decision, and then production takes place. Firms know the probabilities with which births occur in each period, but (prior to period 2) they do not know which workers will eventually have children, and individuals cannot credibly signal their intentions.

Employer costs. We assume that firms incur a cost when an employee is on leave.

Assumption 1 *Output of a worker assigned to job task j is $y_{it} = d_j + c_j\eta_{it}$ if the worker works, and $-k_j$ otherwise, where $k_{j+1} > k_j > 0$.*

[30] More periods could be added to each phase, but this adds complexity without changing the main insights.

[31] One can think of p_f and p_m as a composite probability: the probability of birth times the probability of taking time off, conditional on birth, where we normalize the latter to zero for men, and to one for women.

Our production technology abstracts from complementarities between workers, knowledge hierarchies, slot constraints, or hiring or training costs for temporary replacement workers. Assumption 1 is therefore introduced to capture two ideas: (1) that employers incur costs when a worker assigned to a job is on leave, even when firms do not pay for the leave (Ginja et al. (2020)), and (2) that a worker's absence is more costly when the employee is a manager who affects the productivity of many workers, vs. a rank-and-file worker (Garicano and Rossi-Hansberg (2006)). In line with Swedish laws, firms cannot demote or fire workers based on their current or anticipated childbearing or leave-taking decisions, or write long-term contracts contingent on labor supply behavior after childbirth.

B. Main Result: Promotion Dynamics

Proposition 1 *Promotion dynamics differ for men and women as follows:*

i. Women initially experience a "gender penalty:" in period 1, they are promoted less frequently than men.

ii. Women who give birth also experience a large "motherhood penalty" relative to men or childless women: promotion rates drop dramatically for them in period 2.

iii. Finally, the "gender penalty" reverses, favoring women. However, women with children continue to experience a persistent "motherhood penalty" in period 3.

Proof. See Appendix B.2.

We provide a sketch of the proof and intuition by considering gender differences in the effective ability required for promotion. In the benchmark model, promotion thresholds were time-invariant:

$$\overline{\eta}^j = \frac{d_{j-1} - d_j}{c_j - c_{j-1}}. \tag{7}$$

Men are still promoted according to these thresholds, since nothing changes for them relative to the benchmark model.

By contrast, women's labor supply now varies over the lifecycle, as does their promotion threshold. Employers in period 1 get an immediate return from promoting a high-ability woman from job 0 to job 1, where she will be more productive. However, they anticipate incurring cost $k_1 > k_0$ in the following period if she has a child and takes leave. Consequently, in period 1 the promotion threshold for women, which we will call $\overline{\eta}^*$, is greater than the corresponding threshold for men, $\overline{\eta}^1$. In the appendix, we show that

$$\overline{\eta}^* = \overline{\eta}^1 + p_f \frac{k_1 - k_0}{c_1 - c_0} > \overline{\eta}^1. \tag{8}$$

The inequality follows from the fact that $p_f > 0$, $k_1 > k_0$ and $c_1 > c_0$. Strictly fewer ($\overline{\theta}_1^* - \overline{\theta}_1^1 > 0$) women are promoted to job 1 (the only possible promotion that period), where $\overline{\theta}_1^* - \overline{\theta}_1^1$ is the difference in the corresponding innate ability thresholds for women and men.

The intuition for the dramatic drop in promotion rates for mothers in period 2 — the second part of the proposition — is straightforward. Decisions about childbirth are revealed at the start of the period. For women who have a child and go on leave, the probability of being promoted drops to zero, since employers would only incur a higher cost $k_{j+1} > k_j$ that period upon promoting them. By contrast, both men and and women who do not have children that period are promoted with strictly positive probabilities.

Finally, once there is no more uncertainty about future childbearing, equation (7) determines job assignment for women. For childless women, this occurs after the start of period 2. They are now promoted at the same rate as equally able men; additionally, those women who were previously "passed up" are also promoted, yielding *higher* overall promotion rates and reversing the gender penalty.

For women with children, the gender penalty similarly reverses,but in period 3. Their promotion thresholds are now the same as men's, and those who were "passed up" in period 1 (and on leave in period 2) are now promoted. However, they also experience a persistent motherhood penalty, due to foregone experience, as we show in the appendix. Figure 13A illustrates the dynamic motherhood and gender penalties in promotion probability generated by the model, for a sample parameterization.

C. Additional Wage Dynamics

Workers' wages grow also outside of promotions due to human capital accumulation, since it increases effective ability η_{it}. Appendix B.4 derives wage functions for men and women in each period.[32] Men's wages are identical to the benchmark model: $w^m_{ijt} = d_j + c_j\eta_{it}$. However, women's wage functions vary by period:

Period 0:	$w^f_{ij0} = d_1 + c_1\eta_{i0}$	
Period 1:	$w^f_{ij1} = d_j + c_j\eta_{i1} - p_f k_j$	
Period 2:	$w^f_{ij2} = d_j + c_j\eta_{i2}$	if childless, on govt.-paid leave otherwise
Period 3:	$w^f_{ij3} = d_j + c_j\eta_{i3}$	

Using these wage functions, we can evaluate wage growth outside of promotions. In Section 4.1, we documented that early in the lifecycle, women's wage growth is lower than men's in periods when they are not promoted or when they change firms; however, this difference reverses after age 40 (Table 5). We can now evaluate how the model rationalizes this reversal.

The wage functions imply that, initially, all men and women are hired into job 0 at

[32]For the wage results presented in this subsection, we require an additional assumption that individuals change employers from period 0 to period 1 with probability $\epsilon > 0$. The assumption resolves an indeterminacy issue, as discussed in Appendix B.4. The assumption can be motivated by high job switching rates at young ages. Alternatively, one can impose the assumption that individuals change employers with probability $\epsilon > 0$ in every period, with the same result.

the same wage. In period 1, however, employers pass on to women the expected period 2 costs, $p_f k_j$, in the form of lower wages, since they cannot fire or demote workers who take leave. If they did not pass on those costs, firms would expect to make negative profits. Thus, even among those not promoted from period 0 to period 1 (as well as those promoted), women's growth is strictly lower than for men of the same ability.

Once women's childbearing years end, employers no longer anticipate incurring cost k_j, and a female worker's value to the firm rises. In period 3 (period 2 for childless women), their wages are bid back up in the market. Specifically, wages grow by an additional $p_f k_{j_{-1}}$, where j_{-1} refers to the job held in the prior period. This can lead to a reversal in the wage growth gap between men and women who were not promoted, as observed in the data. Figure 13B illustrates this result for one sample parameterization.

Figure 13B shows that the model's predictions for wage growth during firm changes are similar. Consider an exogenous separation rate $\epsilon > 0$ at the end of each period, leading to a firm change. In Appendix B.4, we prove that under this assumption, women's wage growth during firm changes is again strictly lower than men's in period 1. Additionally, the wage growth gap can reverse in period 3. The same mechanism drives this result: women's market value is lower than men's early on, but increases again once there is no more uncertainty about future childbearing.

To summarize, the model generates a gender gap in promotion probability that reverses; a motherhood penalty in the year of birth and its subsequent persistence; and empirically observed gender differences in wage growth even in non-promotion periods. Assumption 1 is central to these results. Our next proposition tests this key mechanism of the model.

D. A Testable Implication

Proposition 2 *The larger the difference $k_1 - k_0$, the larger the gender difference in promotion probabilities in period 1.*

Proof. See Appendix B.3.

The proposition characterizes a unique testable implication of the model. It says that when the cost of having a promoted worker on leave (relative to an entry level worker) is higher, the pre-birth gender gap is also higher. This follows directly from equation (8), since $\overline{\eta}^*$ is increasing in $k_1 - k_0$. Testing the proposition empirically requires data on costs k_j to employers, which we do not have.[33] However, using information available to us about firms, we can proxy for these costs in two ways.

Our first approach looks at establishment size. The idea behind this test is that an establishment of 1000 employees, with dozens of managers and formalized, on-site

[33]Parental leave-taking has been shown to generate costs for employers, including increases in firm wage bills (Ginja et al. (2020)) and reductions in productivity (Friedrich and Hackmann (2017)). However, existing studies do not quantify whether these costs are greater for higher-level employees.

human resources departments, should be more likely to have established processes for finding or training temporary replacements for managers, redistributing responsibilities and smoothing any disruptions to productivity during a manager's leave than at an establishment with, say, 20 employees. Therefore, the prediction of the model is that one should observe a smaller pre-birth penalty in promotion rates for women at larger establishments. Here the cost of a manager on leave is assumed to vary more across establishments or firms than the cost of an entry-level employee on leave, the idea being that many potential workers can fill a low-level position or be trained for it.

The second approach exploits a feature of the Swedish parental leave system, that a certain amount of paid parental leave (one month prior to 2002, and two months after) is reserved for each parent. It cannot be transferred between the couple, and is forfeited if it is not used. If we observe that men in a given workplace systematically do not use their allotted "daddy months," this may indicate that the firm finds it particularly costly for workers to take time off. Thus, the second test uses information about the share of college-educated men at the firm who do not use their allotted daddy month(s) within the first two years after birth. We note that one possible caveat to this test is that low use of daddy months could capture higher costs not just for k_1, but also k_0. However, men in Sweden tend to have their first birth in their mid-thirties, when many promotions have already occurred. Therefore, the daddy month proxy should generally capture behaviors of older men who are already in some promoted position. We thus expect a larger pre-birth promotion gap at firms where few men take their leave months.

Figure 10 graphs the gender promotion gap at establishments of different sizes, for employees who are under 35 and currently without children, to approximate the population of individuals who are working during period 1 of our model. In line with our prediction, the promotion gap decreases monotonically with establishment size. At the smallest establishments (under 32 employees), the pre-birth promotion gap is 4.0 percentage points, or about 18%. It declines steadily to approximately zero at the largest establishments (more than 717 employees).

To test the same prediction using our alternative proxy based on "daddy months," we construct a binary variable equal to one if the majority (more than 50%) of men who had a child while working at the firm did *not* take their dedicated parental leave months. To ensure consistency, we classify firms only based on observations from the post-2002 period, when the second daddy month was introduced. Approximately 25% of individuals work at such a "low uptake" firm.[34] Table 9 shows that the promotion gap between childless men and women under 35 was greater at firms where men tend not to take their allotted daddy months, as predicted, by about 1.4 percentage points. The gap at high-uptake (i.e., "low-penalty") firms is 2.7 percentage points, or approximately 13.5%. However, at firms where uptake is low, it is 4.1 percentage points, or about 18.6%.

[34]We use a binary variable, rather than quintiles, as there is little variance in this variable across some quintiles. For approximately half of observations, the uptake rate of daddy months is between 60% and 70%.

E. Relaxing Assumptions and Incorporating Part-Time Work

Throughout this section we made several simplifications to keep the analysis tractable. In particular, we excluded the possibility of part-time work. Our model can be generalized by incorporating this margin, and allowing k_j to be a function of hours worked, rather than to jump discretely at zero working hours. We show in Appendix B.6 that under such an extension the motherhood penalty in period 3 is further amplified, while the reversal of the gender penalty is preserved.

Other simplifications included the assumptions that births in period 1 occur with zero probability, and that ability and probability of childbirth are uncorrelated. We relax these assumptions in Appendix B.6 and prove that this does not affect the results.

F. Counterfactual Predictions

Our model does generate at least one counterfactual prediction: that after the end of the childbearing period, childless women not only experience higher promotion rates, but converge completely with men in wages and job assignments. A likely reason for this is that our skill accumulation process is too simple: it depends only on number of years worked, whereas existing studies indicate that skill accumulation depends itself on job assignment (Lise and Postel-Vinay (2019)). Consequently, early lifecycle promotion differences do not have sufficiently persistent effects in our model. Incorporating this mechanism would permit evaluation of the long-run effects of early lifecycle gender penalties.[35]

6 Implications for Existing Explanations

Thus far we have shown that an equilibrium model of careers within firms can reconcile many of the facts we document. A useful conclusion to our analysis is to ask whether other existing explanations can also rationalize the data, and to evaluate the implications of our findings for frameworks currently used to analyze gender wage differentials.

Models of sorting and matching. Firm heterogeneity, which we abstract from in our model, is clearly important, as men and women sort differently across firms.[36] Our interpretation of the facts, however, is that models focused primarily on firm choice and sorting (e.g., Morchio and Moser (2020), Hotz et al. (2017)) are unlikely to account for the majority of the gender wage divergence with age. First, differences in the immediate wage gains associated with firm changes account for only about a quarter of the differences in total wage growth by age 45; second, differences in sorting across firms with different promotion opportunities account for only about 10% of the observed promotion gap. These magnitudes are not negligible, but within-firm differences—between male and female co-workers—strongly dominate. It is still possible that women are more poorly

[35] Gibbons and Waldman (2006) develop an extension along these lines to study cohort effects.

[36] E.g., Hellerstein et al. (2008), Card et al. (2016), Hotz et al. (2017), Sorkin (2017), Barth et al. (2017).

matched with employers in ways we do not directly observe, for example because married women or mothers are more likely to be tied movers. Or, women may have lower firm-specific wage elasticities (Manning (2011)), for the same reasons. This could explain women's lower promotion rates relative to otherwise observationally similar men at the same firm. However, the lifecycle patterns we document generally do not support such an explanation. Rather, promotion gaps are greatest early in the lifecycle, when women are least likely to be married or cohabit, and to fall into this tied mover category.

Finally, this class of models is not designed to generate the thick right tail of the within-firm wage growth distribution that we observe, and does not target this data moment (e.g., Bagger et al. (2014)).[37] Indeed, the fact that most right-tail wage increases occur within-firm has important implications for models that typically load a large share of wage growth onto firm changes (see Cahuc, Postel-Vinay, and Robin (2006)).

Models of human capital accumulation. A second set of models focuses on the links between women's lower labor supply, occupational choice, and skill accumulation and depreciation (e.g. Adda et al. (2017), Bronson (2015), Mincer and Polachek (1956)). These models generally attribute large post-birth penalties to skill depreciation. They can also plausibly rationalize the dynamics in the gender promotion gap that we document if young women, anticipating future labor supply reductions, have less incentive to accumulate human capital and to work high hours prior to birth, leading to fewer promotions. At older ages, in turn, concavity on the human capital accumulation function could imply that returns to additional human capital accumulated after 40 could be higher for women, plausibly leading to higher later-age wage growth.

When prioritizing alternative theories this explanation is promising, especially if our analysis misses some differences in hours worked. However, it has certain drawbacks. One limitation is that the explanation cannot account for gender penalties in promotion for those who never have children. While many women who remain childless may still anticipate having children, one would nevertheless expect the estimated gender penalty to be on average smaller in magnitude. However, this is not what we observe in the data. Second, our results indicate that women experience lower promotion rates than men at *all* levels of hours worked, even among women and men who work part-time. Third, anticipated leave-taking need not decrease Swedish women's incentive to accumulate human capital or work high hours prior to birth. Paid leave in Sweden compensates workers according to their earnings in their previous two years, at an 80% replacement rate, generating significant incentives to work high hours prior to birth.[38]

Finally, similar to models of matching and sorting, these models do not study the non-

[37] We thank Jean-Marc Robin in part for this observation. This is true even for models with on-the-job search and matches to outside offers, as more than 75% of wage increases exceeding 20% occur within firm.

[38] For an individual that expects to take leave in the near future, each additional hour worked generates 1.8 times the wage earnings, which far exceeds the typical estimated returns to human capital accumulation on the intensive margin (e.g., Imai and Keane (2004), Keane (2011)). See discussion in Appendix B.6.

parametric wage growth distribution, and therefore cannot speak to these facts. They are also generally silent about implications for within- vs. across-firm growth.

Behavioral explanations and negotiation. Behavioral theories (Bertrand (2011)) focus on explaining wage differences not related to motherhood. They include theories about systematic gender differences in productivity (e.g. Azmat and Ferrer (2017)), propensity to negotiate (Babcock and Laschever (2009)) or propensity to behave competitively (e.g., Niederle and Vesterlund (2007)). Such differences could generate gender gaps in promotions favoring men, and in theory could affect right-tail pay increases in particular. However, these explanations would have be augmented to rationalize the reversal of the gender penalty after age 40. If lower propensity to compete or negotiate is systematically characteristic of women, one would expect a negative gender penalty over the entire lifecycle, which is not what we observe in the data.[39]

An explanation that is also related to negotiation, but uses instead tools from on-the-job search models, is that women's set of wage offers from outside firms is inferior to men's, as proposed by Booth et al. (2003). In this case, women would have less bargaining power to negotiate for wage increases or promotions. This explanation could account for the dynamic gender penalties we document in this paper if (1) outside firms give women inferior offers relative to men early in the lifecycle, and (2) improve their offers to female workers once they are in their early 40s. However, it is difficult to rationalize why employers would behave this way, unless firms anticipate a possible cost associated with employing women of childbearing age, as in our model.

Careers Within Firms. Relative to other explanations, we have argued that a simple equilibrium model of careers within firms performs exceptionally well in capturing the main features of the data. It is well-suited to generate key features of the non-parametric wage growth distributions we documented in Section 4, including a large right tail that differs for men and women.[40] It generates the virtual absence of a wage gap between men and women upon labor market entry, and its rapid expansion; the observed dynamic motherhood and gender penalties in promotion probability; and observed wage dynamics outside of promotion. This does not rule out the contribution also of other explanations, but strongly supports calls by researchers (e.g., Goldin (2014)) to apply tools from organizational and personnel economics into analyses of gender and wage dynamics. Notable exceptions aside (e.g., Cullen (2020)), empirical and theoretical investigations focused on mechanisms driving gender differences in within-firm advancement are still scarce.

Our model is testable and suggests useful avenues for further research. One prediction is that conditional on ability, young women are promoted at lower rates than men. This

[39] A possible caveat is that the selection of working women could change with age—if the least productive or competitive women increasingly drop out of the labor force, one might observe a relative increase in promotion rates for women at older ages. However, this is ruled out by our setting, as 95% of women in our high skill population are in the labor force at all ages.

[40] The model does not generate the negative wage growth observed in the empirical distribution from Section 4. This can be easily rectified by incorporating learning about ability, as GW'99 also do in an extension.

matches evidence from Blau and DeVaro (2007) that among recent hires, women are promoted at lower rates conditional on performance evaluation. Another interesting prediction is that gender gaps should be small for promotions to positions where labor supply reductions are not much more costly. For example, one should observe that women are promoted as quickly as men from, say, analyst to senior analyst, but then more slowly to project manager. Finally, our model focuses on homogeneous firms, but gender promotion gaps vary by firm, as we show, raising questions about the management practices that affect this variation. These are promising investigations to help answer why women receive fewer high-paying promotions, especially early in their careers.

7 Limitations and Sensitivity

7.1 Limitations

Our analysis has limitations. First, we cannot quantify what share of promotions using our measure correspond to moves up an organizational career ladder. This does not impact our results, but affects their interpretation. Nevertheless, we can observe that individuals are about 1.8 times more likely to change occupation codes in years they are promoted, indicating a strong association. Additionally, in public sector establishments the gender difference in promotion probability is only marginally smaller at 3.7 percentage points. Such establishments generally have regulated wage schedules and limited flexibility to match external wage offers, for example.[41]

Another qualification is that we set the threshold for promotion to be identical for men and women. This is, we believe, the clearest and most transparent approach. It is possible, however, that women are paid much less for identical moves up a career ladder, and are therefore less likely to show up in our promotion measure. This should not affect our qualitative results, but could lead us to overstate gender differences in the probability of experiencing large upward career moves, and understate differences in associated wage growth. Additionally, our study focuses specifically on "large" promotions. Gender differences in promotion rates may be smaller for low-compensated moves, especially those that primarily affect job title, but not job tasks. Indeed, such a prediction follows from our model if costs associated with labor supply reductions differ only marginally across two positions. Combining our wage-based measure with detailed data on organizational structure for a subset of firms can shed additional light on these issues.

Third, existing studies find that moves up the career ladder not only increase wage or base pay, which we study, but also bonus compensation (e.g., Ekinci, Kauhanen, and Waldman (2015)). Through this channel, gender differences in promotion probability could further amplify gender differences in total annual earnings. Finally, our findings

[41] Existing evidence also indicates that relative wage increases of the type we document normally coincide with a change in job position. See Doeringer and Piore (1971) and, most recently, Bayer and Kuhn (2019).

are for a specific institutional and policy environment. Gender and motherhood penalties may differ in countries such as Sweden and the U.S., depending on how costly lengthy parental leaves are to employers, as compared to high rates of job separation after birth. We leave an investigation of this important question for future work.

7.2 Sensitivity

Our main analysis sets the threshold n equal to 10 log points, for the relative wage growth that defines a promotion. Thus, we focus on sizable moves within the firm. In Appendix A.6, we analyze how our results change with alternative choices for n. By construction, as n increases, the promotion rates decline. Table A.6 shows that the share of lifecycle wage differentials explained are fairly similar as one varies n from 7.5 to 12.5, ranging from 83% to 73%. Setting n to 15 reduces the annual promotion probability to just 0.11 for young workers. Nevertheless, the share of lifecycle gender wage differentials accounted for is still quite high, at 64%. The results are comparable when the measure is based on median (rather than average) wage growth of co-workers. Overall, all definitions in Table A.6 yield similar results, for values of n within a reasonable range.

Second, as discussed in Section 2, we restricted our analysis to college-educated individuals with degrees that are not associated almost exclusively with public sector employment in Sweden. In Appendix A.7, we consider the results when all majors are included. The omitted public sector majors, of which women make up a large share (e.g., teaching and nursing), are associated with lower promotion rates than in the baseline population. Since the employers for these majors are commonly municipalities, workers in this population also experience substantially fewer changes in employer. As a result, the share of gender differences in lifecycle wage growth accounted for by promotions in the full population declines, to 58%, while differences in growth associated with firm changes increase, as one would expect. However, even in the omitted group the estimated promotion gap is quite high, at 3.0 p.p. As Table A.7 shows, for the full population of college graduates, the estimated promotion gap under age 35 is about 3.7 p.p., compared to 4.0 p.p. in the baseline population. Similarly, Figure A.5 shows the same non-parametric patterns in wage growth are observed in the full population.

8 Conclusion

This paper is the first to provide extensive empirical evidence on gender differences in individual wage growth. Two important features of our Swedish data – the availability of an administratively recorded wage variable and worker-firm linkages – allow us to study annual wage growth and within-firm wage mobility with minimal measurement error. We highlight three takeaways from our analysis, which have important implications for policy and for models of wage dynamics.

First, our non-parametric analysis indicates that men's wages are higher than women's at age 45 primarily for one reason: isolated years of exceptionally high wage growth are more prevalent for men, especially within firm. Such increases generate less than half of lifecycle wage growth, but account for the overwhelming majority of gender differences in growth. This fact constitutes a crucial step for understanding the gender difference in wage dynamics. Theories addressing differences in men's and women's wage dynamics should be able to generate this important feature of the data.

Second, the cumulative gender gap by age 45 in promotions – that is, in exceptional pay increases relative to other co-workers – is incurred primarily early in the lifecycle and between men and women working at the same firm. Moreover, only about 21% of it is driven by part-time work. These findings are notable in light of a literature that has commonly emphasized glass ceilings, gender differences in sorting across firms, and part-time work as key career impediments. They indicate that gender differences in "big" promotions early in the career are a crucial area for investigation. Our findings also have policy implications. For example, one reason why sorting does not play a more prominent role is that gender differences in promotion probability are even more amplified at firms with high rates of promotion. This indicates that shifting more women to such firms would not, by itself, be likely to substantially increase their overall rates of promotion. Additionally, we find that much of the promotion gap is incurred over a strikingly short period of time – in the year women have a child and year immediately following – with implications for parental leave policy design.

Third, using a simple theoretical model, we argue that the observed wage growth patterns – including a large residual gender penalty in promotion that reverses after age 40 – are consistent with costs to firms associated with employee labor supply reductions, and employer uncertainty about women's future childbearing. On the other hand, they are difficult to reconcile through the lens of several common explanations in the literature. The facts we document also lead us conclude that models focused primarily on search and matching behaviors are unlikely to account for the majority of the gender divergence in wages with age, in line with recent findings by Morchio and Moser (2020).

One key question is whether, in countries such as the U.S., where lengthy job-protected maternity leaves are rare but firm separations for women after childbirth are common, a gender wage growth decomposition would exhibit different patterns. Data on employer costs associated with child-related work interruptions – both in the form of worker leave-taking and earlier job separation – would further speak to this question. However, at least one large-scale data collection finds a similar early career gender promotion gap in the U.S., with the largest differences at the first promotion to a management position (McKinsey and LeanIn.Org (2019)). This suggests that the patterns we document are likely to be observed in a large number of countries, and under diverse policy environments.

References

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

Adda, Jérôme, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy* 125 (2): 293–337.

Albrecht, James, Anders Björklund, and Susan Vroman. 2003. "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics* 21 (1): 145–177.

Albrecht, James, Mary Ann Bronson, Peter Skogman Thoursie, and Susan Vroman. 2018. "The Career Dynamics of High-Skilled Women and Men: Evidence from Sweden." *European Economic Review* 105:83–102.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics*, vol. 34.

Azmat, Ghazala, and Rosa Ferrer. 2017. "Gender Gaps in Performance: Evidence from Young Lawyers." *Journal of Political Economy* 125 (5): 1306–1355.

Babcock, Linda, and Sara Laschever. 2009. *Women Don't ask: Negotiation and the Gender Divide*. Princeton University Press.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *The American Economic Review* 104 (6): 1551–1596.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994a. "The Internal Economics of the Firm: Evidence from Personnel Data." *The Quarterly Journal of Economics* 109 (4): 881–919.

———. 1994b. "The Wage Policy of a Firm." *The Quarterly Journal of Economics* 109 (4): 921–955.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti. 2017. "The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data." *NBER Working Paper No. 23381.*

Bayer, Christian, and Moritz Kuhn. 2019. "Which Ladder to Climb? Decomposing Life Cycle Wage Dynamics."

Bernhardt, Dan. 1995. "Strategic Promotion and Compensation." *The Review of Economic Studies* 62 (2): 315–339.

Bertrand, Marianne. 2011. "New Perspectives on Gender." *Handbook of Labor Economics* 4:1543–1590.

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:pp. 228–255.

Bidwell, Matthew, and Ethan Mollick. 2015. "Shifts and ladders: Comparing the Role of Internal and External Mobility in Managerial Careers." *Organization Science* 26 (6): 1629–1645.

Blau, Francine, and Jed DeVaro. 2007. "New Evidence of Gender Differences in Promotion Rates: An Empirical Analysis of a Sample of New Hires." *Industrial Relations* 46 (3): 511–530.

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

Booth, Alison, Marco Francesconi, and Jeff Frank. 2003. "A Sticky Floors Model of Promotion, Pay and Gender." *European Economic Review* 47:295–322.

Burdett, Kenneth. 1978. "A Theory of Employee Job Search and Quit Rates." *The American Economic Review* 68 (1): 212–220.

Cabral, Robert, Marianne A Ferber, and Carole A Green. 1981. "Men and Women in Fiduciary Institutions: A Study of Sex Differences in Career Development." *The Review of Economics and Statistics*, pp. 573–580.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2): 323–364.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics* 131 (2): pp. 633–686.

Cullen, Zoë B, and Ricardo Perez-Truglia. 2020. "The Old Boys' Club: Schmoozing and the Gender Gap." *NBER Working Paper No. 26530*.

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

Doeringer, Peter B., and Michael J. Piore. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, MA: Heath Lexington Books.

Ekinci, Emre, Antti Kauhanen, and Michael Waldman. 2015. "Bonuses and Promotion Tournaments: Theory and Evidence." *The Economic Journal*.

Friedrich, Benjamin U, and Martin B Hackmann. 2017. "The Returns to Nursing: Evidence from a Parental Leave Program." Technical Report.

Garicano, Luis, and Esteban Rossi-Hansberg. 2006. "The Knowledge Economy at the Turn of the Twentieth Century: the Emergence of Hierarchies." *Journal of the European Economic Association* 4 (2-3): 396–403.

Gibbons, Robert, John Roberts, et al. 2013. "Economic Theories of Incentives in Organizations." *Handbook of Organizational Economics*, pp. 56–99.

Gibbons, Robert, and Michael Waldman. 1999a. "Careers in Organizations: Theory and Evidence." *Handbook of Labor Economics* 3:2373–2437.

———. 1999b. "A Theory of Wage and Promotion Dynamics Inside Firms." *The Quarterly Journal of Economics* 114 (4): 1321–1358.

Gicheva, Dora. 2013. "Working Long Hours and Early Career Outcomes in the High-end Labor Market." *Journal of Labor Economics* 31 (4): 785–824.

Ginja, Rita, Arizo Karimi, and Pengpeng Xiao. 2020. "Employer Responses to Family Leave Programs." *Manuscript*.

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

Goldin, Claudia, and Lawrence F Katz. 2011. "The Cost of Workplace Flexibility for High-Powered Professionals." *The Annals of the American Academy of Political and Social Science* 638 (1): 45–67.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, Erling Barth, et al. 2017. "The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census." *American Economic Review: Papers and Proceedings* 107 (5): 110–14.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy* 122 (3): 621–660.

Harris, Milton, and Bengt Holmstrom. 1982. "A Theory of Wage Dynamics." *The Review of Economic Studies* 49 (3): 315–333.

Hellerstein, Judith, David Neumark, and Melissa McInerney. 2008. "Changes in Workplace Segregation in the United States between 1990 and 2000: Evidence from Matched Employer-Employee Data." In *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, 163–195.

Hotz, V Joseph, Per Johansson, and Arizo Karimi. 2017. "Parenthood, Family Friendly Workplaces, and the Gender Gaps in Early Work Careers." *Manuscript*.

Imai, Susumu, and Michael P Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review* 45 (2): 601–641.

Javdani, Mohsen, and Andrew McGee. 2015. "Moving up or Falling Behind? Gender, Promotions, and Wages in Canada." *IZA Discussion Paper No. 9380*.

Johnston, David W, and Wang-Sheng Lee. 2012. "Climbing the Job ladder: New Evidence of Gender Inequity." *Industrial Relations* 51 (1): 129–151.

Keane, Michael P. 2011. "Labor Supply and Taxes: A survey." *Journal of Economic Literature* 49 (4): 961–1075.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209.

Lazear, Edward P. 1995. *Personnel Economics*. Volume 1993. MIT press.

Lazear, Edward P, and Paul Oyer. 2004. "Internal and External Labor markets: A Personnel Economics Approach." *Labour Economics* 11 (5): 527–554.

Lazear, Edward P, and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89 (5): 841–864.

Lise, Jeremy, and Fabien Postel-Vinay. 2019. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *Manuscript*.

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

Manning, Alan. 2011. "Imperfect Competition in the Labor Market." In *Handbook of Labor Economics*, Volume 4, 973–1041.

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

Manove, Michael. 1997. "Job Responsibility, Pay and Promotion." *The Economic Journal* 107 (440): 85–103.

McCue, Kristin. 1996. "Promotions and Wage Growth." *Journal of Labor Economics* 14 (2): 175–209.

McKinsey, & Co., and LeanIn.Org. 2019. Women in the Workplace 2019. `https://wiw-report.s3.amazonaws.com/Women_in_the_Workplace_2019.pdf`. [Accessed: 2019-11-27].

Milgrom, Paul, and Sharon Oster. 1987. "Job Discrimination, Market Forces, and the Invisibility Hypothesis." *The Quarterly Journal of Economics* 102 (3): 453–476.

Morchio, Iacopo, and Christian Moser. 2020. "The Gender Pay Gap: Micro Sources and Macro Consequences." *Manuscript*.

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

Pergamit, Michael R, and Jonathan R Veum. 1999. "What is a Promotion?" *ILR Review* 52 (4): 581–601.

Postel-Vinay, Fabien, and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6): 2295–2350.

———. 2006. "Microeconometric Search-Matching Models and Matched Employer-Employee Data." *The Proceedings of the 9th World Congress of the Econometric Society*. 877–907.

Prendergast, Canice. 1993. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *The Quarterly Journal of Economics* 108 (2): 523–534.

———. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming Up Inequality." *The Quarterly Journal of Economics* 134 (1): 1–50.

Sorkin, Isaac. 2017. "The Role of firms in gender earnings inequality: evidence from the United States." *American Economic Review* 107 (5): 384–87.

Thomas, Mallika. 2018. "The Impact of Mandated Maternity Benefits on the Gender Differential in Promotions: Examining the Role of Adverse Selection." *Manuscript*.

Topel, Robert, and Michael Ward. 1992. "Job Mobility and the Careers of Young Men." *The Quarterly Journal of Economics* 107 (2): 439–479.

Van der Klaauw, Bas, and António Dias Da Silva. 2011. "Wage Dynamics and Promotions Inside and Between Firms." *Journal of Population Economics* 24 (4): 1513–1548.

Waldfogel, Jane. 1998. "Understanding the "Family Gap" in Pay for Women with Children." *The Journal of Economic Perspectives* 12 (1): 137–156.

Waldman, Michael. 1984. "Job Assignments, Signalling, and Efficiency." *The RAND Journal of Economics* 15 (2): 255–267.

Table 1: Summary Statistics, 1960-1970 Cohort

	Men	Women
Labor Force Participation:		
Ages 25-29	0.97	0.97
Ages 30-34	0.96	0.95
Ages 35-44	0.96	0.95
Ages 40-45	0.95	0.95
Educational Attainment:		
Bachelor's	0.43	0.54
Master's, Ph.D., or Professional	0.57	0.46
Children and Fertility		
Mean Age at First Birth	32.95	31.72
Had a child by 45	0.75	0.81
Mean # of Children, Conditional on Having Children	2.25	2.19
Workplace Characteristics		
Share in Public Sector, Ages 25-30	0.23	0.36
Share in Public Sector, Ages 40-45	0.21	0.37
Average Log Firm Size	6.07	6.45
Average Log Wage at Firm	10.18	10.13
Individuals	60,353	42,602
Individual-Year Observations	958,322	686,917
Individual-Year Obs., incl. Educated Co-Workers at Firm	39,193,218	39,037,944

Notes: Data comes from multiple matched Swedish administrative data registers covering years 1985 to 2013, for college educated individuals ages 25 to 45, born between 1960 and 1970. See Section 2 for additional details.

Table 2: Percentiles and Higher-Order Moments of Real Annual Wage Growth Distribution

Real Annual Wage Growth	Men	Women	Difference (M-W)
Percentiles (in log points):			
10th	-2.4	-2.2	-0.2
25th	0.3	0.3	0.0
50th	3.1	2.9	0.2
75th	8.1	7.0	1.1
90th	15.7	13.8	2.0
99th	37.3	32.7	4.7
Tail Characteristics:			
Kurtosis	20.1	18.2	1.9
Skewness	0.79	0.58	0.21

Table 3: Real Annual Wage Growth During Firm Changes vs. Within-Firm

Size of Real Annual Wage Increase:	% Wage Increases Occurring During Firm Changes	Likelihood of Experiencing Wage Increase: Men to Women	
		During Firm Change	During Firm Tenure
<0	12.1%	0.99	1.06
0 - 0.05	8.1%	0.94	0.90
0.05 - 0.10	11.2%	0.98	1.01
0.10 - 0.15	15.5%	1.00	1.13
0.15 - 0.25	21.3%	1.05	1.23
0.25+	29.7%	1.26	1.47

Notes: The first column categorizes by size all individual-year observations of real annual wage growth. The second column summarizes the share of observations corresponding to firm changes over the past year. To calculate the ratios in the third column, we construct the distribution of wage increases during firm changes, separately for men and women. For example, 27.8% and 29.5% of men's and women's firm changes, respectively, are associated with growth of 0 to 5 log points. We then take the ratio of these two numbers (e.g., is 0.94, as shown). The final column does the same for growth during firm tenure.

Table 4: Within-Firm Growth: Average Growth at Firm, Relative Wage Growth, and Promotions

Size of Real Annual Wage Increase	Likelihood, Men to Women (1)	Average Wage Growth at Firm (2)	Relative Wage Growth (3)	Share That Are Promotions (LRWI) (4)
0 to 0.075	0.92	0.022	0.009	0.02
0.075 to 0.15	1.09	0.035	0.071	0.24
0.15 to 0.225	1.21	0.040	0.141	0.81
0.225 to 0.25+	1.44	0.025	0.263	0.98

Notes: The table categorizes by size all individual-year observations of wage growth that occur during firm tenure. In column (1), we construct the distribution of within-firm wage increases separately for men and women, and then take the male-female ratio. Columns (2) and (3) summarize two components of each observed wage increase: average wage growth at the individual's firm, calculated based on college-educated co-workers' growth at the same firm in the same year; and relative wage growth, equal to the individual's total wage increase minus average wage growth at the firm. In column (4) a promotion, or large relative wage increase, is a wage gain that is 10 p.p. higher than the average wage growth at the firm. See text for details.

Table 5: Probability of Wage Growth Event and Associated Real Wage Growth, By Type

	Firm Change		Promotion (LRWI)		Non-Promotion	
	Men	Women	Men	Women	Men	Women
Annual Probability						
Ages 26-30	0.295	0.291	0.178	0.137	0.527	0.572
Ages 31-35	0.247	0.234	0.140	0.107	0.613	0.660
Ages 36-40	0.192	0.186	0.097	0.079	0.712	0.735
Ages 41-45	0.145	0.147	0.069	0.063	0.785	0.791
Annual Wage Growth						
Ages 26-30	0.095	0.084	0.199	0.186	0.033	0.031
Ages 31-35	0.091	0.083	0.215	0.203	0.034	0.031
Ages 36-40	0.065	0.059	0.211	0.195	0.020	0.022
Ages 41-45	0.041	0.047	0.200	0.182	0.011	0.016

Notes: Each year, the three categories above correspond to mutually exclusive and exhaustive events. A promotion, or large relative wage increase (LRWI), is a wage gain that is n percentage points higher than the average wage growth at the firm that year, in years that the worker did not switch firms. In this and all subsequent tables and figures, promotions are defined using a threshold of $n = 10$. For results for alternative thresholds and a related discussion, see section 7.2 and Appendix A.

Table 6: Gender Difference in Annual Probability of Promotion, Ages 26 to 35

	(1)	(2)	(3)	(4)	(5)
Female	-0.039***	-0.044***	-0.044***	-0.040***	-.039***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Human Capital	No	Yes	Yes	Yes	Yes
Tenure at Firm	No	No	Yes	Yes	Yes
Firm FE	No	No	No	Yes	Yes
Occupation FE	No	No	No	No	Yes

*** Significant at 1% level. Notes: See note in Table 5. In the regressions above, the outcome variable is an indicator corresponding to whether or not an individual experienced a promotion (LRWI) in a given year. All specifications control for year fixed effects. Human capital controls include indicators for age, years of higher education, field of major, as well as a quadratic in actual years of labor market experience. $N = 190,404$. Regression includes only years with no firm change.

Table 7: Gender Difference in Probability of Promotion Ages 26 to 35: Controls for Hours Worked

Dep. Variable: Probability of Promotion (LRWI)	(1)	(2)	(3)
Female	-0.040***	-0.021***	-0.032***
	(0.002)	(0.003)	(0.003)
Human Capital Controls & Firm Fixed Effects	Yes	Yes	Yes
Part-Time Work, Part-Time History, & Hours Worked	No	Yes	Yes
Controls for Year of & Year Following Birth	No	Yes	No

*** Significant at 1% level. Notes: See notes in Table 5 and Table 6. All specifications also control for year fixed effects.

Table 8: Annual Promotion Probability and Promotion Gap, by Years Relative to First Birth

	Promotion Gap (1)	Promotion Probability (2)	% Difference (3)
5 to 1 years before first birth	-0.032	0.198	-16%
Year of and first year after birth	-0.087	0.163	-53%
2 to 10 years after first birth	-0.015	0.124	-12%

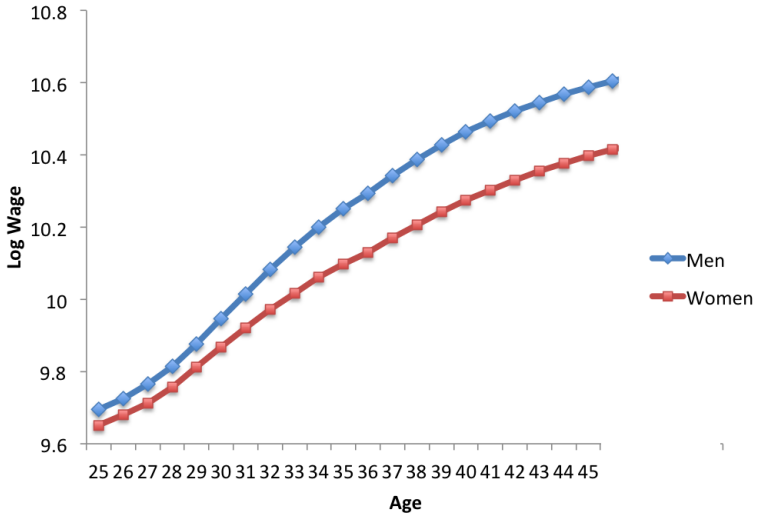
Notes: See note in Table 5. Column (1) records the gender difference in annual promotion (LRWI) probability. Column (2) records men's annual promotion rates. Column (3) is the ratio of column (1) to column (2). Calculations are for individuals who ever had children, in years they did not switch firms.

Table 9: Promotion Probability Prior to Childbirth or for Childless Individuals Under 35, by Men's Uptake of "Daddy Months" at the Firm

Dep. Variable: Probability of Promotion (LRWI)	Coefficient	Std. Error
Low Daddy Month Take-Up Firm × Female	-0.014**	(0.007)
Low Daddy Month Take-Up Firm	0.021***	(0.005)
Female	-0.027***	(0.003)

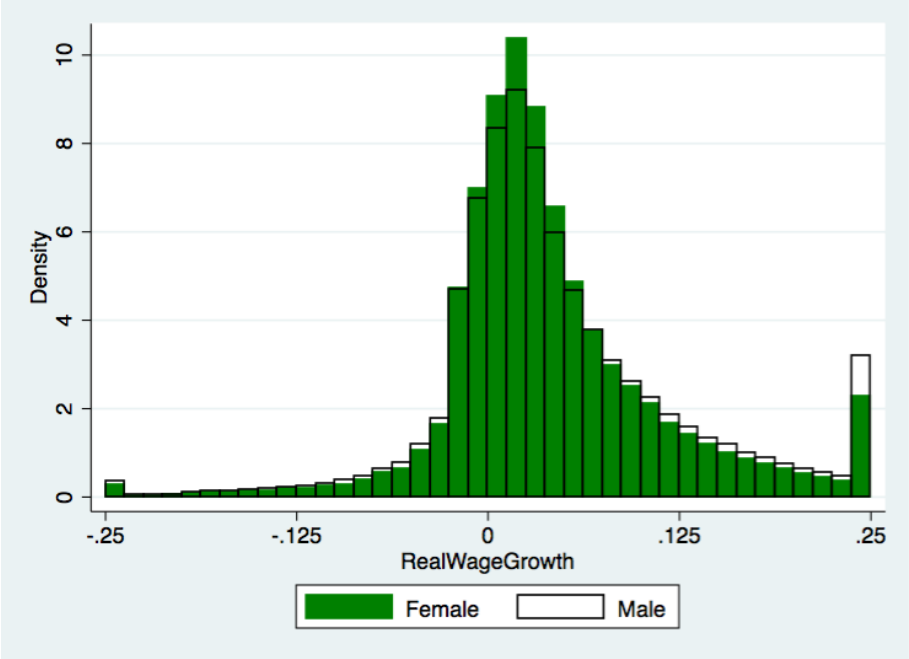
** Significant at 5% level. *** Significant at 1% level. Notes: See note in Table 5. Regression controls include indicators for year, age, years of higher education, and field of major, as well as a quadratic in years of experience. A firm is considered to have low uptake of daddy months if fewer than half of male employees at the firm use their 60 days of allotted parental leave within the first two years after the birth of their child. Sample includes all individuals under 35 who have not yet had a first birth or who never have children.

Figure 1: Lifecycle Wage Profiles, Men and Women with College Education



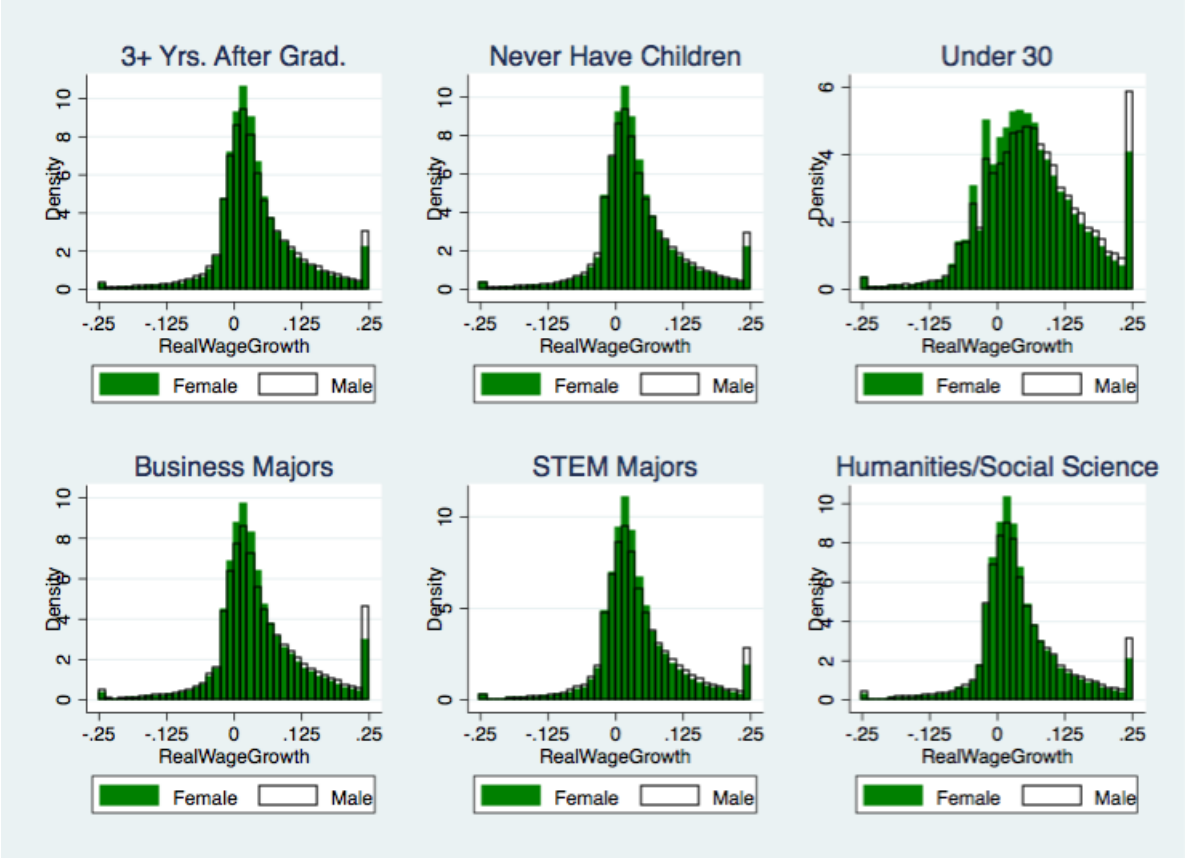
Notes: Profiles are for college-educated individuals from the 1960-1970 birth cohorts in Sweden. For additional population details, see Section 2. Source: Statistics Sweden.

Figure 2: Distribution of Real Annual Wage Growth, Ages 25 to 45



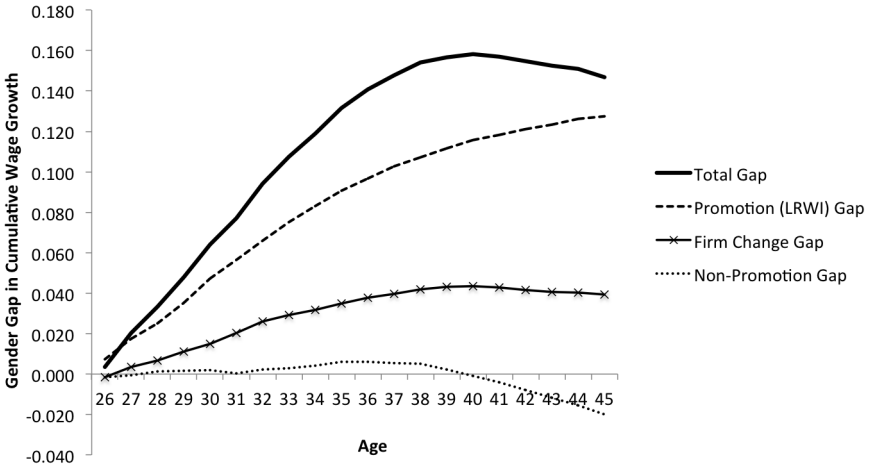
Notes: The histogram pictures individual-year level observations of real annual wage growth, for individuals ages 25 to 45. The tails of the distribution are collapsed to mass points, at -0.25 and 0.25.

Figure 3: Distribution of Real Annual Wage Growth, Ages 25 to 45



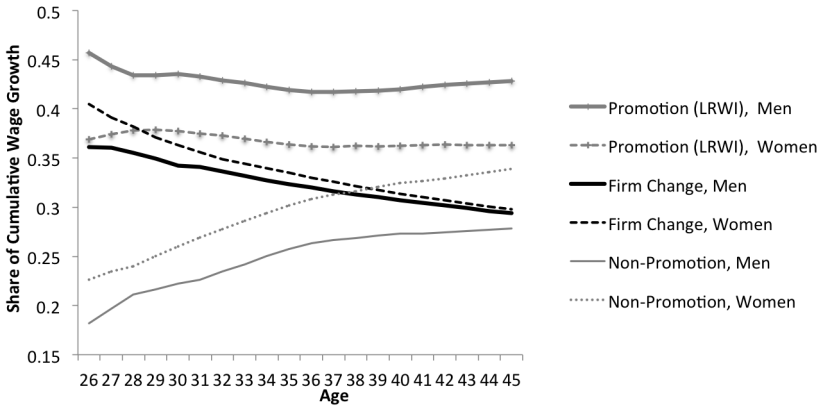
Notes: See note in Figure 2.

Figure 4: Decomposition of Gender Difference in Cumulative Wage Growth Since Age 25



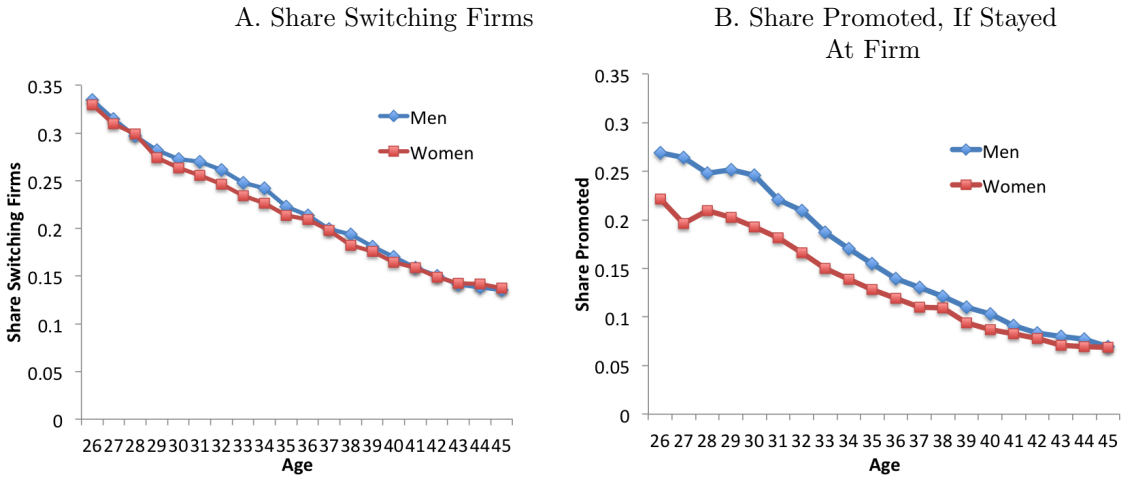
Notes: See note in Table 5. The "total gap" is the total cumulative gender difference in wage growth, since age 25. The remaining series refer to cumulative gender differences in wage growth associated with each type of growth event. Together, the three lower series add up to the total gap.

Figure 5: Decomposition of Lifecycle Wage Growth



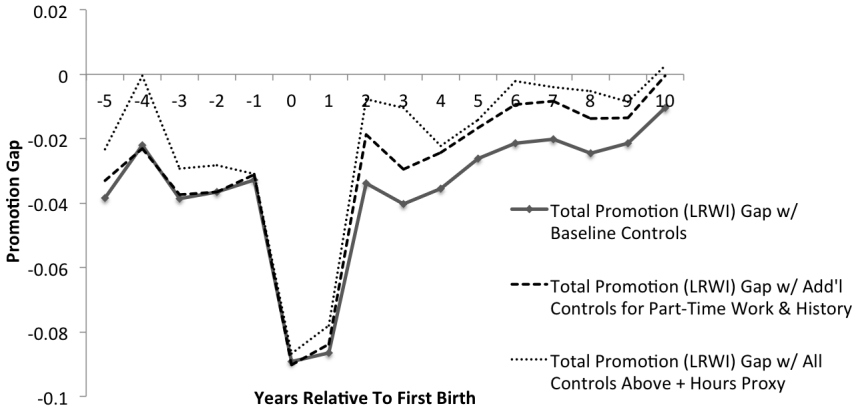
Notes: See note in Table 5. The figure records the contribution of each type of wage growth event to total cumulative wage growth over the lifecycle, since age 25. At each age, the three solid (dotted) series add up to 1, for men (women).

Figure 6: Share Switching Firms and Share Promoted, by Age



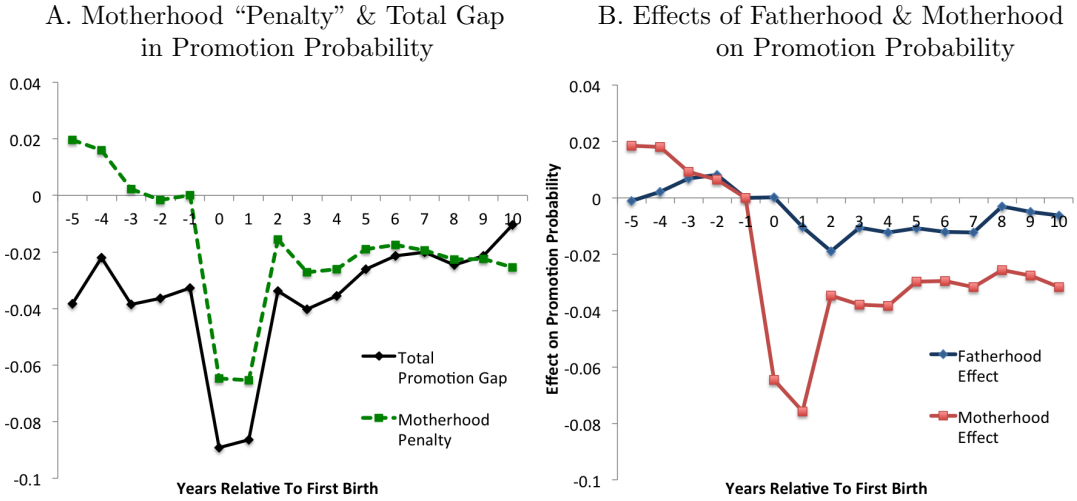
Notes: See note in Table 5. Panel A graphs the share switching firms out of all workers. Panel B graphs the share experiencing a promotion (LRWI), out of those who did not switch firms in the current period.

Figure 7: Gender Difference in Annual Promotion Probability, by Years Relative to First Birth



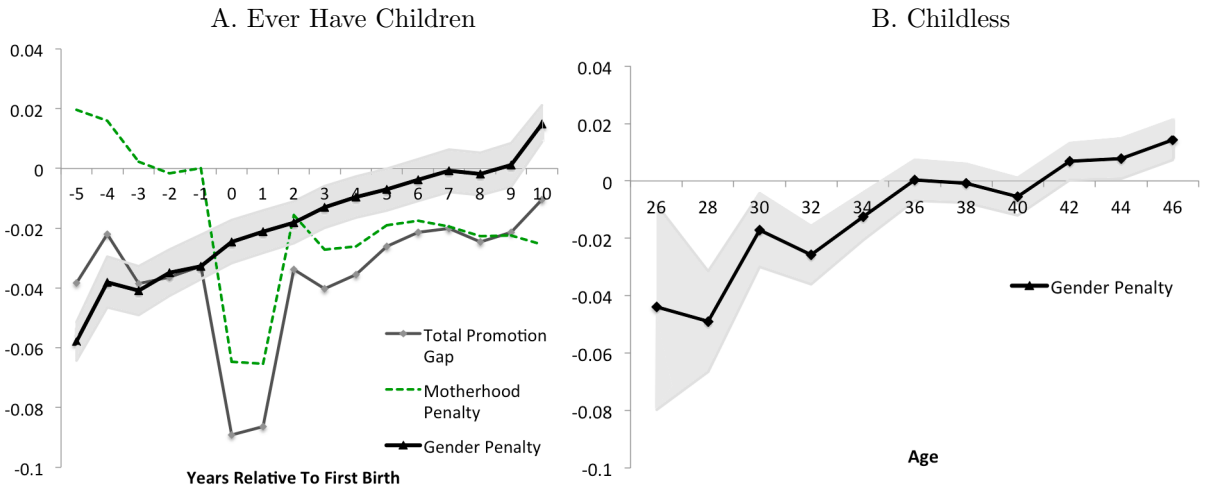
Notes: See note in Table 5. Confidence intervals omitted for clarity. Table A.8 shows standard errors for all coefficients. The series plot coefficients on the interaction between indicators for female and for number of years relative to first birth, from equation (4), with an indicator for whether an individual received a promotion (LRWI) in a given year as the outcome variable. Baseline controls include indicators for year, age, years of higher education, field of major, firm, and a quadratic in years of experience. For further information about the controls for part-time hours and part-time history, as well the detailed hours proxy variable, see Section 4.2 and Appendix A.4. Sample includes all individuals in years that they did not switch firms.

Figure 8: Motherhood Penalty, Total Gap in Annual Promotion Probability, and Effects of Parenthood, by Years Relative to First Birth



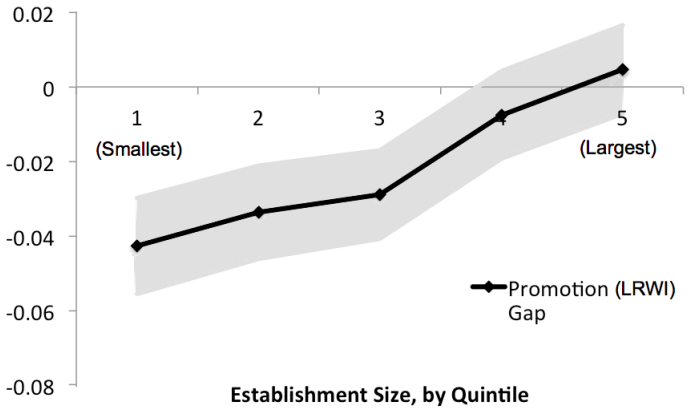
Notes: Confidence intervals omitted for clarity. See Table A.8 for standard errors and the note in Table 5. Both series in Panel A correspond to regression coefficients on the interaction terms between indicators for female and for number of years relative to first birth. The difference between the two series in Panel A is that the specification for the motherhood penalty also interacts all other control variables (except firm fixed effects) with gender, while the specification for the total promotion gap does not (see equations (4) and (6)). Controls include indicators for year, age, years of higher education, field of major, firm, and a quadratic in years of experience. Panel B decomposes the "motherhood penalty" in Panel A into separate motherhood and fatherhood effects (see equation (5)). Specifically, the "motherhood penalty" in Panel A is equal to the difference between the motherhood and fatherhood effects that are graphed in Panel B. See main text for details. Sample includes all individuals in years that they did not switch firms.

Figure 9: Gender Penalty in Annual Promotion Probability, by Time to First Birth (Individuals with Children) and by Age (Childless Individuals)



Notes: See note in Table 5. Shaded areas represent 95% confidence intervals. In Panel A, the "gender penalty" is the difference between the total gender gap in annual promotion (LRWI) probability and the motherhood penalty in annual promotion, which are graphed in the same figure. In Panel B, "childless" refers to individuals who never have children by age 45. The gender penalty in promotion probability for childless individuals corresponds to regression coefficients on interactions between female and age. Controls in Panel B include indicators for age, years of education, field of study, and firm, as well as a quadratic for years of experience, years of tenure, hours worked, and part-time history. For controls in Panel A, see Figure 8. Sample includes all individuals in years that they did not switch firms.

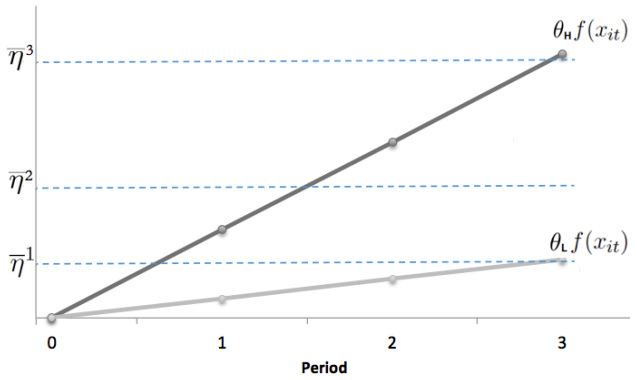
Figure 10: Gender Difference in Annual Promotion Probability Prior to Childbirth or for Childless Individuals Under 35, by Establishment Size



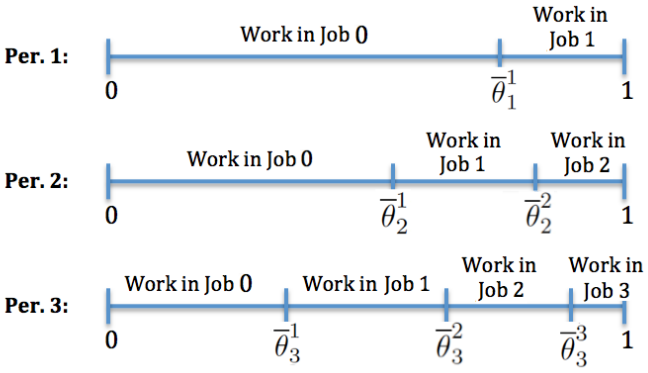
Notes: See note in Table 5. Shaded areas represent 95% confidence intervals. The above estimates correspond to coefficients on interaction terms between female and quintile of establishment size, with promotion (LRWI) in a given year as the outcome variable. Sample includes all individuals under 35 who have not yet had a first birth or who never have children. Regression controls include indicators for year, age, years of higher education, field of major, establishment size quintile, a quadratic in years of experience, and log firm size. The promotion gap represents the coefficients on an interaction between female and establishment size. The quintiles of establishment size vary from 32 employees or less (first quintile), to 717 employees or more (top quintile).

Figure 11: Promotion of Individuals in Model: Example

A. Promotion Over Time for High and Low θ



B. Share Assigned to Jobs, by Period



Notes: Panel A illustrates effective ability and promotion over time for two individuals who work every period, one with high innate ability θ_H, and one with low innate ability, θ_L. Cut-offs $\overline{\eta}^1, \overline{\eta}^2, \overline{\eta}^3$ determine when individuals are promoted. In the example above, the high innate ability individual is promoted exactly once each period. The low innate ability individual is promoted for the first time only in period 3. Panel B illustrates innate ability cut-offs and job assignments by period as individuals accumulate experience, again for those who work every period.

Figure 12: Lifecycle Structure and Timing of Model

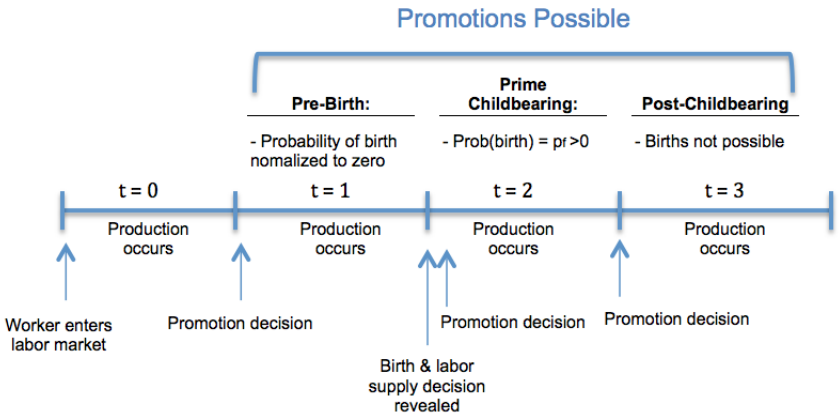
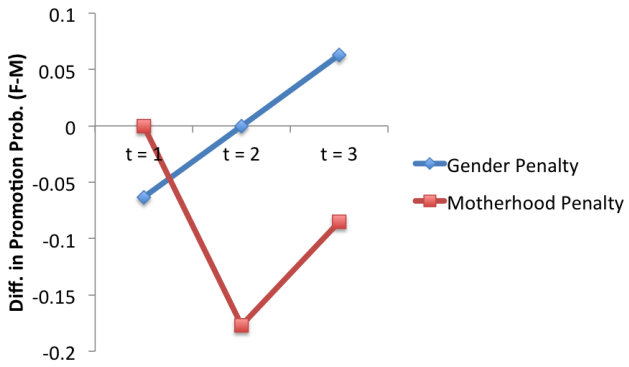
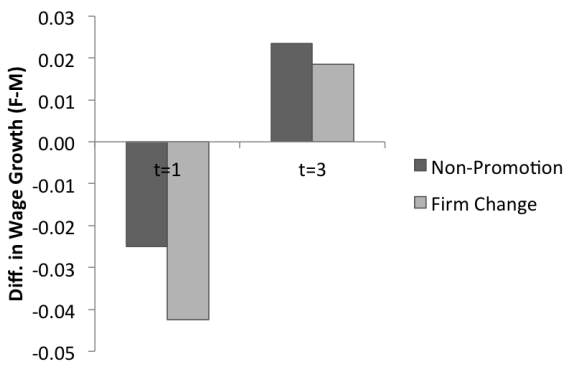


Figure 13: Model Dynamics: Simulation

A. Promotion Gap (F-M)



B. Conditional Wage Growth (F-M)



Notes: For the simulations above, the technology parameters are $(d_0, d_1, d_2, d_3) = (1, 0.8, 0.6, 0.4)$, $(c_0, c_1, c_2, c_3) = (0.5, 1, 1.3, 1.5)$. The human capital accumulation function corresponds to $(f(1), f(2), f(3)) = (0.95, 1.1, 1.3)$; $p_f = 0.8$; and $k_0 = 0.1$, $k_1 = 0.15$. In panel B, conditional wage growth is calculated for all individuals who were either not promoted or switched jobs, across all jobs.

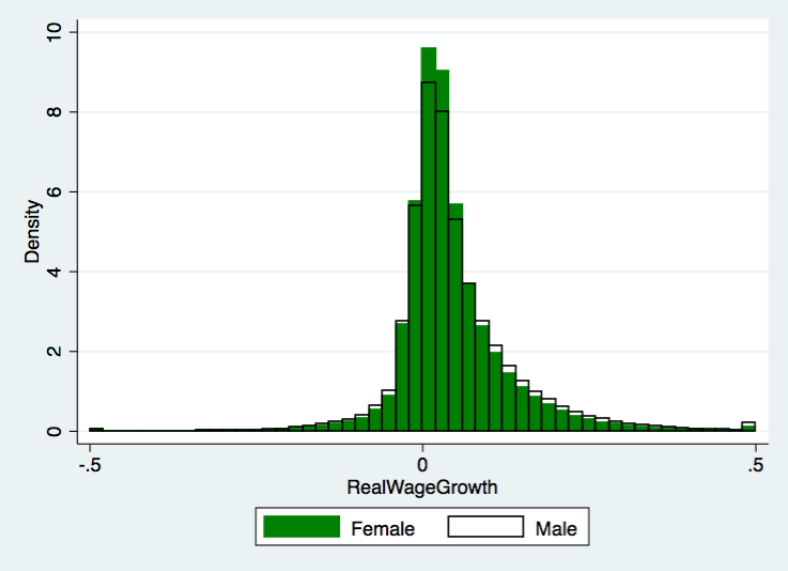
FOR ONLINE PUBLICATION

Appendix A Empirical Analysis

A.1 Wage Growth Analysis

This section supplements the analysis in Section 3. Figure A.1 shows the majority of the tail behavior for men and women. It is identical to Figure 2, except that the tails are collapsed at 0.5 and -0.5 log points. Next, Table A.1 displays the relationship between one-year and five-year wage growth, to document the persistence in wage shocks. As the table shows, real annual wage growth shocks are highly persistent. For example, 91% of individuals who experienced a 20 log point shock from year t to $t+1$ have wage levels at $t+5$ that are also at least 20 log points higher relative to year t. For all sizes of wage shocks, mean reversion is low.

Figure A.1: Distribution of Real Annual Wage Growth, Ages 25 to 45



Notes: The histogram tabulates individual-year level observations of real annual wage growth, for individuals ages 25 to 45. The tails of the distribution are collapsed to mass points, at -0.5 and 0.5.

Table A.1: Persistence of Wage Shocks: 1-yr vs. 5-year Wage Growth

		$\ln w_{t+1} - \ln w_t$				
		-0.2	-0.1	0.0	0.1	0.2
$\ln w_{t+5} - \ln w_t$	-0.1	**0.51**	**0.17**	0.02	0.01	0.01
	0.0	**0.28**	**0.49**	**0.26**	0.05	0.02
	0.1	0.08	0.15	**0.32**	0.16	0.06
	0.2	0.06	0.09	0.19	**0.26**	0.13
	0.3	0.03	0.05	0.10	**0.20**	**0.20**
	0.4	0.02	0.03	0.06	0.14	0.19
	0.5	0.02	0.03	0.06	0.19	**0.39**

Finally, Table A.2 shows that for most individuals, a small number of high-growth

years generate a large portion of lifecycle wage growth, in line with the distributional evidence. About 80% of individuals achieve half of their wage growth between ages 25 and 45 in just three (not necessarily consecutive) years.

Table A.2: Concentration of Lifecycle Wage Growth

Share of individuals who achieved, during three yrs. of greatest wage growth:	
50% of lifecycle wage growth	0.80
60% of lifecycle wage growth	0.55
70% of lifecycle wage growth	0.34
50% of lifecycle wage growth (excluding first three years after graduation)	0.81

Notes: In the calculations above, the three years of greatest wage growth need not be consecutive. Sample includes only the 21,222 individuals for whom we observe wages in all years after graduation.

A.2 Sorting Across Firms

In this section, we provide supplementary figures and graphs that complement the analysis of gender differences in sorting across firms in Section 4.2. We provide further detail about whether women are (1) more likely to work at firms at which there are few promotion opportunities, or (2) promoted less than men at firms the same promotion opportunities.

To analyze this question, we first construct a variable measuring promotion opportunities at the firm that is consistent with the analysis throughout the paper. Specifically, we calculate the average share of high-skilled employees who are promoted annually at each firm, using the same definition of promotion as described in Section 2.2. In Figure A.2, we order firms by this measure on the x-axis. Panel A plots the share female across firms with different promotion opportunities. For reference, it also plots the distribution of high-skill workers across these firms, since high-promotion firms are less common. Figure A.2 shows that women on average work at firms with fewer opportunities for promotion, both in our cohort and among high-skill workers overall. In firms with the fewest opportunities, women and men represent a roughly equal share of high-skill workers. However, at firms that are in the upper half of the distribution for the yearly share of workers promoted, women represent around 35%-43% the firm's high-skilled employees. These differences can be interpreted as gender differences in "sorting" across firms.

Panel B plots the probability of being promoted for men and women in our cohort, conditional on the promotion opportunities at their firm. Women have a lower probability of being promoted across all firm types, with more pronounced gender differences at firms with more promotion opportunities. On average, women each year are about 3.9 p.p. (20.9%) less likely to get promoted. However, at firms where at least 10% of workers are promoted each year, women in our cohort are on average about 6.2 p.p. (21.5%) less likely than men to get promoted. These differences can be interpreted as the "within-firm" differences in promotion probability.

To analyze the importance of sorting vs. within-firm gender differences in promotion, in Section 4 we compare estimates from regressions with and without firm fixed effects. To complement this, we conduct a simple decomposition exercise below. In particular, we consider what the implied gender gap in promotion rates would be if (1) women were

distributed across firms as men are (i.e., no differences in sorting), with the corresponding promotion rates for women at those firms; or, (2) women had identical probabilities of promotion as men at their current firms. As Table A.3 shows, assigning men's distribution across firms to women only reduces the gap in promotion rates from 3.9 p.p. to 3.1 p.p. By contrast, assigning men's probability of promotion to women at their current firms reduces the gap in promotion rates by 75%, to 1.0 p.p. The estimates from the simple decomposition exercise indicate that about a quarter of the promotion gap is accounted for by sorting, which is similar to the results from the fixed effects estimates, albeit slightly higher. However, unlike the fixed effects analysis, this decomposition does not control for any covariates. Both results indicate, however, that gender differences in promotion rates of men and women at the *same* firm – rather than differences in sorting across firms – are the primary driver of gender differences in promotion probability in this population.

Figure A.2: Share Female and Share Promoted At Firm, By Opportunities for Promotion (LRWI)

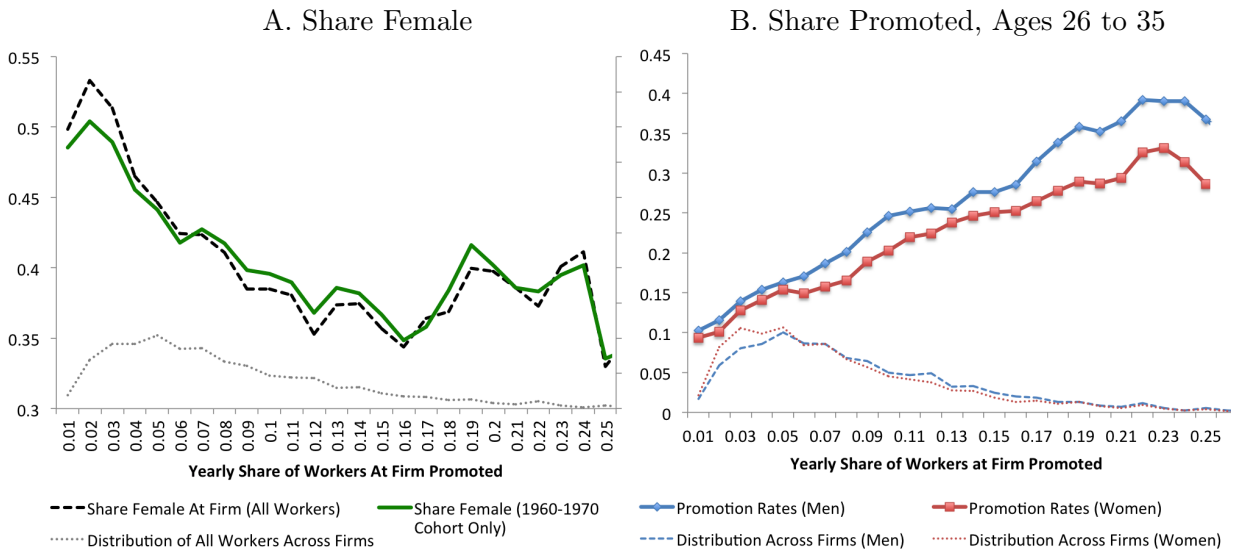


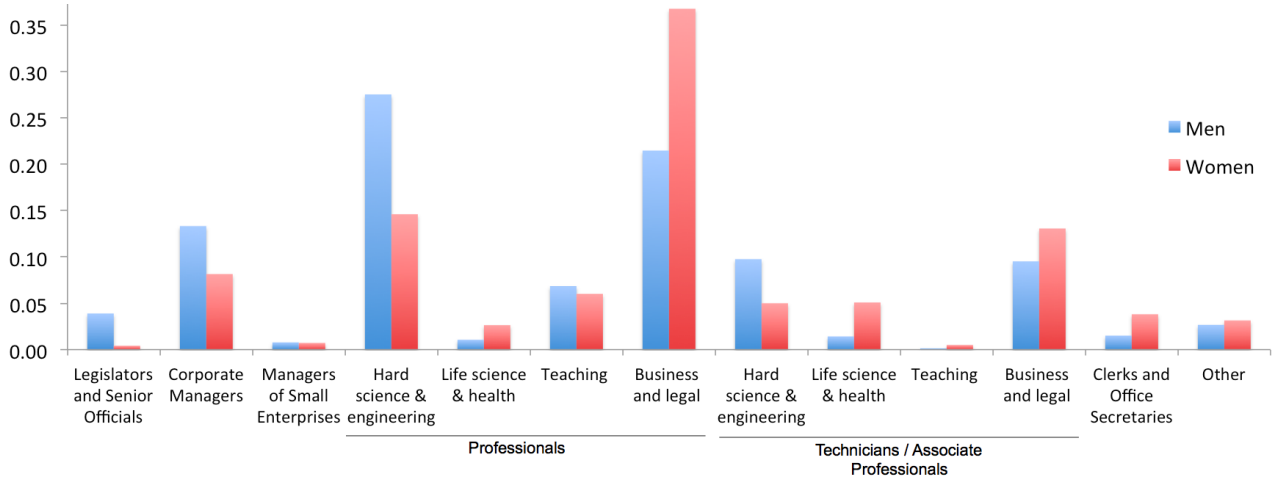
Table A.3: Importance of Cross-Firm vs. Within-Firm Differences, Ages 26 to 35

	Gap	Share of Gap Explained
Gender Gap In Promotion (LRWI) Rates	3.88	-
Counterfactual: Same Distribution Across Firms	3.05	0.214
Counterfactual: Same Promotion Rate Within Firms	0.97	0.749

Notes: See note in Table 5. In the first counterfactual, women are re-assigned to have the same distribution across firms as men. In the second counterfactual, women are assigned the same average promotion rate as men at their firm.

A.3 Occupational Distribution of Men and Women

Figure A.3: Occupational Distribution of Men and Women, Ages 25-45



Notes: The two-digit occupation codes above consist of the following three-digit occupations. *Legislators and senior officials:* Legislators and senior government officials; Senior officials of special-interest organizations. *Corporate managers:* Directors and chief executives; Production and operations managers; Other specialist managers. *Managers of small enterprises:* Managers of small enterprises.
Hard science and engineering professionals: Physicists, chemists and related professionals; Mathematicians and statisticians; Mathematicians; Computing professionals; Architects, engineers and related professionals. *Life science and health professionals:* Life science professionals; Health professionals (except nursing); Nursing and midwifery professionals. *Teaching professionals:* College, university and higher education teaching professionals; Secondary education teaching professionals; Primary education teaching professionals; Special education teaching professionals; Other teaching professionals. *Business, Legal, and Other professionals:* Business professionals; Legal professionals; Archivists, librarians and related information professionals; Social science and linguistics professionals (except social work professionals); Writers and creative or performing artists; Religious professionals; Public service administrative professionals; Administrative professionals of special-interest organizations; Psychologists, social work and related professionals.
Hard science and engineering associate professionals: Physical and engineering science technicians; Computer associate professionals; Optical and electronic equipment operators; Ship and aircraft controllers and technicians; Safety and quality inspectors. *Life science and health associate professionals:* Agronomy and forestry technicians; Health associate professionals (except nursing); Nursing associate professionals; Life science technicians. *Teaching associate professionals:* Pre-primary education teaching associate professionals; Other teaching associate professionals. *Business, Legal, and Other associate professionals:* Finance and sales associate professionals; Business services agents and trade brokers; Administrative associate professionals; Customs, tax and related government associate professionals; Police officers and detectives; Social work associate professionals; Artistic, entertainment and sports associate professionals; Religious associate professionals. *Office clerks:* Office secretaries and data entry operators; Numerical clerks; Stores and transport clerks; Library and filing clerks; Mail carriers and sorting clerks; Other office clerks. *Customer services clerks:* Cashiers, tellers and related clerks; Client information clerks.
All other two-digit categories represent a small share of college-educated workers. They include the following two-digit occupations: *Personal and protective services workers; Models, salespersons and demonstrators; Skilled agricultural and fishery workers; Extraction and building trades workers; Metal, machinery and related trades workers; Precision, handicraft, craft printing and related trades workers; Other craft and related trades workers; Machine operators and assemblers; Drivers and mobile-plant operators; Sales and services elementary occupations; Agricultural and fishery laborers; Laborers in mining, construction, manufacturing and transport; Armed forces.*

A.4 Hours Worked and Construction of Proxy Measure

In this section, we discuss the construction of our proxy hours measure, and provide supplementary evidence about gender difference in hours worked.

The proxy measure is constructed by dividing annual labor income, which we observe for the calendar year, by contracted wage, which we observe in the yearly survey month, typically September. This contracted wage measure is the same one used to construct the promotion variable, which compares wages in Septembers of consecutive years. One issue with using the constructed hours variable to analyze the relationship between current year hours and current year promotions is that one will, on average, underestimate the relationship between the two variables. The reason is that if a promotion occurred, for example, in August of the current year, then total annual labor income will reflect lower

wages from January to July, and higher income only from August to December. Dividing this annual income by the high wage recorded in September will lead us to infer that hours worked were lower in the current year than they truly were. This would be true for all individuals promoted after January of the current year.

One alternative is to use hours worked from the previous period. In fact, conceptually this is desirable, as the personnel economics literature suggest that promotions are awarded for past effort and on-the-job learning (Gibbons and Waldman (1999a)). However, this has a similar, although opposite problem. Suppose a promotion occurred in October 2000, which would be recorded as a promotion only in 2001, when we observe it in September of that year. In this case, dividing year 2000 annual labor income – which will already partially reflect the promotion – by the wage from September 2000 would lead us to infer that hours worked were higher than they truly were. In this case, we would overestimate the true relationship between hours worked and promotion.

To avoid introducing a mechanical correlation between promotion in year t and hours worked in year t or $t-1$, we therefore use a twice lagged measure of hours worked whenever we rely on the proxy hours measure, i.e. hours in year $t-2$. Whenever we use the proxy hours measure, we therefore restrict the sample to individuals who have at least two full years of tenure on the job, to ensure that our measure captures hours worked at the current firm, not a previous firm. In practice, however, this restriction does not affect any of the results.

In Table A.4, we summarize the proxy hours measure, by age, and compare it to the contracted hours measure. The contracted hours measure and the proxy measure capture similar patterns. As expected, the proxy hours measure is somewhat higher for men than contracted hours, since contracted hours do not capture work above full-time. For women, the proxy combines both the effect of hours worked above full-time, as well as time away for parental leave, and therefore is lower in most periods.

Finally, Table A.5 documents the promotion rate at different levels of hours worked using the proxy measure (column 1), and the share of men and women working at those hours (columns 2 and 3). As the Table shows, the relationship between promotion rate and hours worked is roughly flat below 41 hours worked, and positive above 41 hours worked. As women are less likely than men to work in categories with the highest weekly hours, this clearly contributes to the documented differences in promotion probability. However, column 4 of the table shows that women have lower promotion rates, for any level of hours worked, complicating the simple interpretation that women's lower promotion rates are entirely a consequence of lower hours worked. Column 4 records the coefficient on female from equation (3), estimated separately for each category of weekly hours. The results indicate that for all ranges of hours worked in the prior year, we observe a substantial gender difference in current period promotion rates, of about 3 percentage points.

Table A.4: Share Working Part Time and Average Weekly Hours Worked, By Age

	Part-Time		Hours (Contracted)		Hours (Proxy)	
	Men	Women	Men	Women	Men	Women
Ages 26-30	0.04	0.11	39.1	38.0	41.9	38.6
Ages 31-35	0.04	0.19	39.2	37.1	41.9	34.9
Ages 36-40	0.05	0.23	39.1	36.7	40.5	33.7
Ages 41-45	0.04	0.18	39.2	37.2	40.6	36.0

Table A.5: Probability of Promotion Ages 26 to 35, By Average Weekly Hours Worked

Weekly hours:	Promotion Rate (1)	Share of Men (2)	Share of Women (3)	Promotion Gap (4)
20 to 32	0.147	0.051	0.140	-0.027***
32 to 39	0.164	0.123	0.139	-0.036***
39 to 41	0.159	0.341	0.339	-0.032***
41 to 48	0.199	0.321	0.212	-0.031***
48 or more	0.244	0.145	0.053	-0.031***

Notes: See note in Table 5. The promotion rate is calculated for men, and is conditional on not having switched firms. All hours calculations use the proxy hours measure. Column 4 of Table A.5 records the coefficient on *female* from equation (3), estimated separately for different categories of hours, and controlling for additional variation in hours worked within category.

A.5 Normalization of the Motherhood Penalty

This section briefly describes the required normalization for the motherhood penalty, and how the choice of normalization affects any qualitative results in Section 4.3.

In the existing literature (e.g., Kleven et al. (2019), Angelov et al. (2016)), it is standard to normalize the motherhood penalty to zero in the year prior to first birth ($k = -1$). In other words, the year prior to first birth is the omitted category in the regression. Because of the required normalization, the term "dynamic motherhood penalty" is used, since only changes in the penalty from year to year have a direct interpretation. By contrast, estimates of the total promotion gap by year to first birth are not normalized and the magnitudes are directly interpretable.

A change in the normalization shifts the series of estimated coefficients for the motherhood penalty by a constant. Crucially, however, all dynamics for the motherhood penalty are preserved, as are the dynamics for the implied gender penalty (i.e., the total promotion gap minus the motherhood penalty). Of course, the precise point at which the gender penalty graphed in Figure 9A crosses the x-axis will depend on the specific normalization adopted for the motherhood penalty. Under the alternative assumption that the motherhood penalty accounts for 100% of the promotion gap in the year after first birth, the gender penalty would become positive and statistically significant about five years earlier, when women are on average 36.

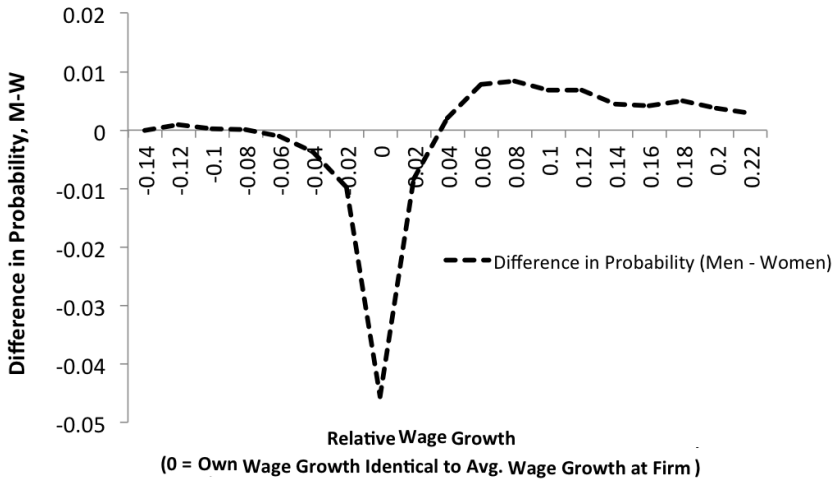
A.6 Alternative Definitions of Promotions

In Section 2.2, we discuss our wage-based measure of promotion. According to our definition, a promotion occurs when the wage growth of a worker is n percentage points higher than the average annual wage growth of college-educated co-workers at the firm

in the same year. In this section, we discuss the choice of threshold, $n = 10$, used as a baseline, and what happens as it is varied. We also discuss results using alternative thresholds.

Figure B.1 provides information about the choice of threshold, and what happens as this threshold for defining a promotion is varied. The x-axis in the figure corresponds to relative wage growth, and the dotted series in the figure corresponds to differences in the probability with which men and women experience such relative wage growth. As the figure shows, women are substantially more likely to experience zero relative growth – i.e., to experience average wage growth at the firm – while men are more likely to experience wage growth that is at least four percentage points higher than the firm average. Precisely at the point where $n = 10$, the cut-off for our baseline measure, the average wage growth is approximately 12.3%. If the threshold n is reduced below 10 (but above $n = 4$), more lower-growth observations will be recorded as promotions, with a larger gender difference in cumulative number of promotions, since men are still substantially more likely to experience wage growth that is at least 4 log points above the firm's mean. The opposite is true when n is increased above 10.

Figure A.4: Gender Difference in Probability of Experiencing Level of Relative Wage Growth



Next, in table A.6, we consider alternative thresholds for n, setting n equal to 7.5, 12.5, and 15. We compare this to our baseline results, when $n = 10$. Additionally, we construct the promotion measure using median (instead of mean) wage growth of college-educated co-workers, again setting n equal to 10. Finally, for reference, we also define promotion as any real wage gain that exceeds 10%. Table A.6 summarizes the results, showing that both qualitatively and quantitatively, the main results are similar for a relatively wide range of values for n. The same is true also for the remaining results of the paper (available upon request).

Table A.6: Main Results Using Alternative Measures Promotion

	Promotion Gap	Promotion Rate (M)	Wage Gain M	Wage Gain F	% Explained by Promotion
n = 7.5	-0.047***	0.28	0.18	0.17	0.83
	(0.002)				
n = 10	-0.040***	0.20	0.20	0.19	0.75
	(0.002)				
n = 12.5	-0.032***	0.15	0.24	0.22	0.71
	(0.002)				
n= 15	-0.026***	0.11	0.26	0.25	0.64
	(0.002)				
Median-based,	-0.042***	0.22	0.21	0.20	0.80
n = 10	(0.002)				
Absolute wage growth,	-0.047***	0.29	0.19	0.18	0.87
10+ percent	(0.002)				

Notes: See note in Table 5. In the first column, controls include indicators for year, age, years of higher education, field of major, a quadratic in years of experience, and firm fixed effects. The sample for the regression in column (1) and for the promotion rate calculation in figure (2) consists of individuals ages 26 to 35 in years when a firm switch did not occur. In columns (3) and (4) the wage gain is the annual wage growth associated with a promotion. Column (5) calculates the share of the gender differences in lifecycle wage growth explained by promotion-related growth, using the decomposition from equation (2).

A.7 Public Sector Majors

In Table A.7, we consider results for the full population of college graduates in the 1960-1970 cohort. The full population of graduates includes the baseline population analyzed throughout the paper, as well individuals with majors associated almost entirely with public sector employment, omitted in the analysis. These include all majors related to teaching, medicine and social work. Column (1) in Panel A shows summary results for the baseline population. Column (2) shows results for just the omitted population, and column (3) provides results for all graduates from the 1960-1970 cohorts. As Panel A shows, even among individuals with predominantly public sector majors, the total wage gap by age 45 is quite large, at 0.19, identical to the baseline population. The overall gender wage difference for all graduates when the two groups are combined is even higher than in the baseline group, at 0.25. The reason for this is that average wages in the omitted group, which has more women, are significantly lower than in the baseline group.

Next, Panel B compares decomposition results for the baseline vs. full population. The share of gender differences in lifecycle wage growth explained by firm changes is higher in the full population. This is the consequence of two facts. First, women are far more represented among public sector majors. Second, those majors are associated with substantially fewer firm changes: among both men and women ages 25-45, about 13%-14% of those with public sector majors change employers annually, compared to 20% in the baseline population we study. Additionally, the importance of gender differences in non-promotion growth increases modestly in the full population, since wage growth in non-promotion periods is lower in the omitted majors. Correspondingly, the share of gender differences in wage growth explained by differences promotion-related growth decreases to 58%.

Table A.7: Summary Results: Baseline vs. Full Population

Panel A: Summary Statistics

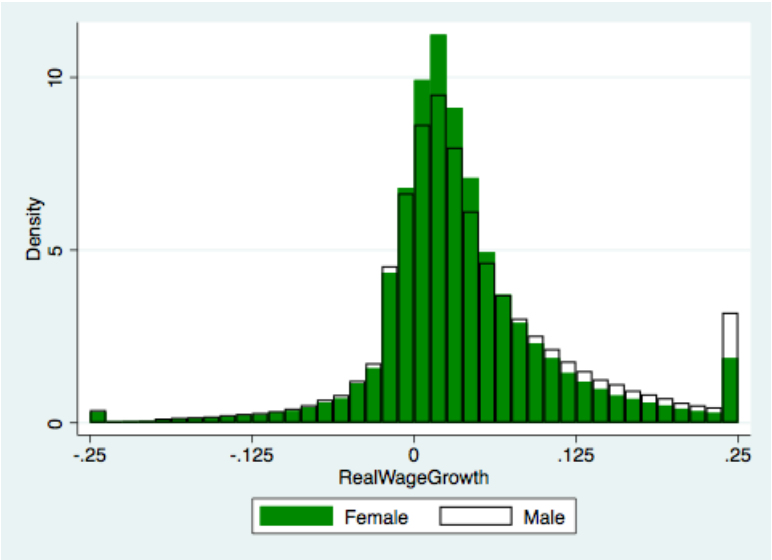
	Baseline Population	Omitted Majors	All Graduates
Promotion Gap	-0.040***	-0.033***	-0.037***
	(0.002)	(0.002)	(0.001)
Mean Wage (Men), Ages 40 to 45	10.59	10.36	10.54
Mean Wage Gap, Ages 40 to 45	0.19	0.19	0.25
Men	60,353	12,398	72,751
Women	42,602	33,839	76,441

Panel B: Decomposition Results

	% of Gender Difference in Wage Growth Explained by:		
	Firm Switches	Promotions	Non-Promotion Growth
Baseline Population	0.26	0.75	-0.01
Full Population	0.39	0.58	0.03

In Panel A, the promotion gap is estimated for individuals ages 26 to 35 in years when a firm switch did not occur. Controls include indicators for year, age, years of higher education, field of major, and firm, as well as a quadratic in years of experience. All other estimates are for the full population of individuals ages 25-45, unless otherwise indicated.

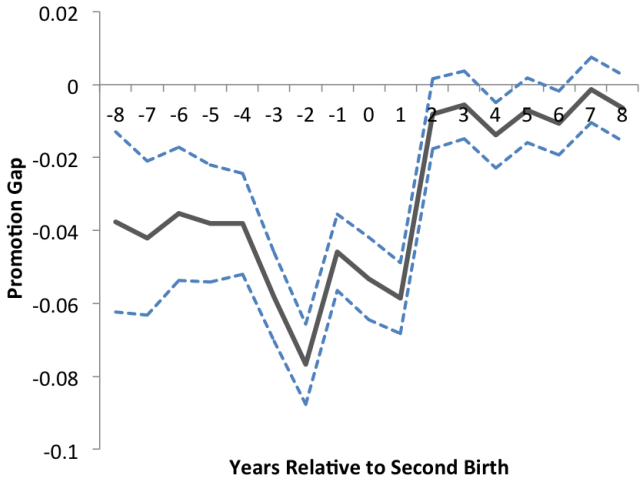
Figure A.5: Distribution of Real Annual Wage Growth, Ages 25 to 45 (Full Population)



Finally, Figure A.5 shows that non-parametric patterns in wage growth we documented hold similarly for the full population.

A.8 Other Supplementary Tables and Figures

Figure A.6: Gender Gap in Promotions, by Years Relative to Second Birth



Notes: The dashed series graph the 95% confidence interval. The regression is for individuals who ever have two or more children, and include indicators for year, age, years of higher education, field of major, and firm; a quadratic in years of experience; and controls for part-time work and part-time history. The promotion gap represents the coefficient on "female," or the gender difference in promotion probability.

Table A.8: Probability of Promotion, by Time to Birth

Years to First Birth:	(1)	(2)	(3)	(4)	(5)
-5	-0.038	-0.033	-0.020	0.018	-0.001
	(0.009)***	(0.010)***	(0.011)*	(0.004)***	(0.004)
-4	-0.022	-0.023	0.000	0.018	0.002
	(0.008)***	(0.008)***	(0.010)	(0.006)***	(0.005)
-3	-0.038	-0.037	-0.029	0.009	0.007
	(0.007)***	(0.008)***	(0.009)***	(0.005)*	(0.005)
-2	-0.036	-0.037	-0.027	0.007	0.008
	(0.006)***	(0.007)***	(0.008)***	(0.005)	(0.005)***
-1	-0.033	-0.031	-0.032	0.000	0.000
	(0.006)***	(0.006)***	(0.007)***	(0.000)	(0.000)
0	-0.089	-0.090	-0.086	-0.065	0.000
	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.004)
1	-0.086	-0.084	-0.077	-0.076	-0.010
	(0.004)***	(0.005)***	(0.005)***	(0.005)***	(0.004)***
2	-0.033	-0.018	-0.006	-0.035	-0.019
	(0.004)***	(0.005)***	(0.007)	(0.004)***	(0.004)***
3	-0.039	-0.029	-0.007	-0.038	-0.011
	(0.004)***	(0.005)***	(0.006)	(0.004)***	(0.004)***
4	-0.035	-0.024	-0.020	-0.038	-0.012
	(0.004)***	(0.005)***	(0.005)***	(0.004)***	(0.004)***
5	-0.026	-0.017	-0.014	-0.030	-0.011
	(0.004)***	(0.004)***	(0.005)***	(0.004)***	(0.004)***
6	-0.021	-0.009	-0.001	-0.030	-0.012
	(0.004)***	(0.004)***	(0.005)	(0.004)***	(0.004)***
7	-0.020	-0.008	-0.003	-0.032	-0.012
	(0.004)***	(0.004)***	(0.005)	(0.004)***	(0.004)***
8	-0.024	-0.014	-0.005	-0.026	-0.003
	(0.004)***	(0.004)***	(0.005)	(0.004)***	(0.004)
9	-0.021	-0.013	-0.008	-0.028	-0.005
	(0.004)***	(0.004)***	(0.005)*	(0.005)***	(0.004)
10	-0.010	0.000	0.003	-0.032	-0.006
	(0.004)***	(0.004)	(0.005)*	(0.004)***	(0.004)*

* Significant at 10% level. *** Significant at 1% level. Notes: Columns (1) to (3) provide point estimates and standard errors for the three series in Figure 7. Column (1) refers to the baseline results. Column (2) adds controls for part-time work and part-time history. Column (3) adds controls for hours worked using the proxy variable. Columns (4) and (5) provide point estimates and standard errors for the three series in Figure 8. Column (4) refers to the motherhood gap and column (5) refers to the fatherhood gap.

Appendix B Augmented Model and Proofs

B.1 Augmented Model

In the beginning of Section 5, we noted that the full information benchmark model by Gibbons and Waldman (1999b) that we build on generates only moderate wage growth at time of promotion. Several extensions can amplify promotion-related growth. First, it is possible to incorporate symmetric learning by firms about unobserved ability of the worker, as is developed in the same paper by Gibbons and Waldman. These wage increases can be even further augmented by allowing learning about ability to be private information to the employer, with publicly observed promotions acting as a signal of worker ability, as in Waldman (1984). Another alternative is to incorporate compensation for higher effort associated with more complex jobs. We have experimented with all of these approaches, and take the latter one due to its simplicity. Since incorporating this feature does not substantively change any of the main findings, we omit it from the model description in Section 5, but add it to the exposition and proofs below for completeness.

To do this, we assume that an (observable) effort cost e_j for the worker is associated with each particular job, with $0 \leq e_0 < e_1 < ... < e_J$. This assumption is a reduced-form way to capture that jobs higher up the career ladder not only require more human capital, but are also associated with more responsibility and correspondingly greater required levels of effort or disutility (e.g., stress) on the part of the worker, for which they must be compensated. A worker's utility is denoted by $u_{itj} = w_{itj} - e_j$.

Consider, for the moment, the benchmark environment with effort costs. A worker is initially hired to job 0. The worker will be assigned efficiently to job 1 when his ability exceeds the η that solves

$$d_0 + c_0\eta_{it} - e_0 = d_1 + c_1\eta_{it} - e_1.$$

We denote this solution, as before, as $\overline{\eta}^1$, where

$$\overline{\eta}^1 = \frac{d_0 - d_1 + (e_1 - e_0)}{c_1 - c_0}.$$

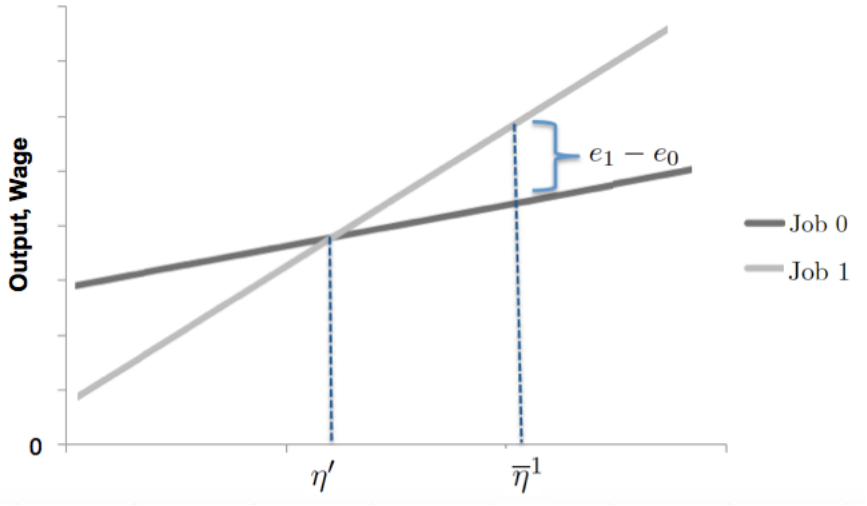
As in the benchmark model without effort costs, the worker is paid the competitive wage in this frictionless environment, $w_{itj} = d_j + c_j\eta_{it}$. However, wages now increase discontinuously at the time of the job change. To see why, consider Figure B.1. In the model without effort costs, recall that individuals are promoted as soon as their effective ability is such that their output in job 1 exceeds their output in job 0. This point is marked by η' in the figure. However, as the figure shows, wage does not jump discretely at this point. By contrast, in the model with effort costs, a worker will strictly prefer working at job 0 if his ability is equal to η', or to any value below $\overline{\eta}^1$, since in that case $u_{it0} > u_{it1}$. To be induced to work in job 1, a worker with $\eta_{it} < \overline{\eta}^1$ would have be paid a wage that is higher than his output. Therefore, the worker is promoted only at the higher threshold, when his output in job 1 exceeds his output in job 0 by a discrete amount of $e_1 - e_0$, and the wage similarly jumps by an additional $e_1 - e_0$ at time of promotion.

In equilibrium, free entry, costless switching of workers between firms, and perfect

competition for workers will yield the efficient job assignment: that is, output net of effort costs, $d_j + c_j\eta_{ij} - e_j$, is maximized for each worker. Thus, in the augmented benchmark environment, workers are promoted according to the following cut-offs:

$$\overline{\eta}^j = \frac{d_{j-1} - d_j + (e_j - e_{j-1})}{c_j - c_{j-1}}. \tag{9}$$

Figure B.1: Promotion from Job 0 to Job 1



B.2 Proof of Proposition 1

To prove proposition 1 for the augmented model with gender and motherhood, we begin by considering the cut-off values for effective ability for men and women, respectively. Men supply the same labor as in the benchmark model, since only women are assumed take leave in period 2. Therefore, nothing changes for men relative to the frictionless benchmark environment, and for them the cut-off values are given by equation (9) above.

To describe cut-off values for women, we begin with period 3, when there is no possibility of childbearing. In this period, all firms in the market know that there is no possibility of incurring cost k_j if they promote a woman. Therefore a firm's problem for women that period is identical as for men. Thus, the cut-off values for η in period 3 that determine job assignment for women are identical to those for men: $\overline{\eta}^1, \overline{\eta}^2, \overline{\eta}^3$, as defined by equation (9).

In period 2, childbirth is possible. Women enter the period and share p_f have a child and reduce their labor supply to zero, which is observable prior to the promotion decision. These women are promoted with probability zero, since employers would only incur a higher cost $k_{j+1} > k_j$ if they promote those women in that period, with no benefit. For women who remain childless that period, however, all uncertainty about current and future childbearing has been resolved. These childless women are promoted in period 2 based on cut-offs that are again identical to those for men: $\overline{\eta}^1$ and $\overline{\eta}^2$, as defined by equation (9).

Finally, in period 1 firms that choose to promote/hire a woman to a given job j in the current period will incur a higher cost k_j with probability p_f in the following period. This follows from the assumption that firms cannot fire or demote workers based on leave-taking, which introduces a friction in the model. In period 1, there is only one possible type of promotion for men and women: from job 0 to job 1. Let $V_1(j=0)$ equal the expected output net of effort costs in periods 1 and 2 of a female worker who is not promoted at the start of period 1. Alternatively, if she is promoted at the start of period 1, her expected value corresponds to $V_1(j=1)$. In this case, we have:

$$V_1(j=0) = d_0 + c_0\eta_{it} - e_0 + p_f \cdot (-k_0) + (1-p_f) \cdot V_2^*$$
$$V_1(j=1) = d_1 + c_1\eta_{it} - e_1 + p_f \cdot (-k_1) + (1-p_f) \cdot V_2^*$$

where V_2^* indicates a woman's expected output net of effort costs if she remains childless in period 2, a value that is identical in both equations since V_2^* does not depend on job assignment in period 1. Consequently, the period 1 threshold value $\overline{\eta}^*$ that equalizes expected values of female employees in jobs 0 and 1 solves

$$d_0 + c_0\overline{\eta}^* - e_0 + p_f \cdot (-k_0) = d_1 + c_1\overline{\eta}^* - e_1 + p_f \cdot (-k_1).$$

The solution to this equation is

$$\overline{\eta}^* = \frac{d_0 - d_1 + (e_1 - e_0)}{c_1 - c_0} + p_f \frac{k_1 - k_0}{c_1 - c_0} = \overline{\eta}^1 + p_f \frac{k_1 - k_0}{c_1 - c_0} > \overline{\eta}^1. \tag{10}$$

where the final inequality follows from the fact that $p_f > 0$, $k_1 > k_0$ and $c_1 > c_0$. Thus, in period 1 the ability threshold for promotion applied to women is higher than for men. Next, we derive the cut-off values θ_τ^j for men and women in each period, to prove each part of Proposition 1.

Proof of Proposition 1(i): In period 1, all men and women have exactly $\tau = 1$ year of experience, and only one type of promotion is possible (from job 0 to job 1), under the restricted set of parameterizations we consider, in which the highest ability men and women not on leave are promoted exactly once each period. For men, the threshold value $\overline{\eta}^1$, as defined by equation 9, determines the corresponding threshold value of innate ability required for the promotion of an individual with one year of experience to job 1, which is $\overline{\theta}_1^1 = \frac{\overline{\eta}^1}{f(1)}$. Thus, share $(1 - \overline{\theta}_1^1)$ of men are promoted. For women, the higher threshold value $\overline{\eta}^*$ determines a correspondingly higher threshold value of innate ability, which we denote as $\overline{\theta}_1^* = \frac{\overline{\eta}^*}{f(1)} > \frac{\overline{\eta}^1}{f(1)} = \overline{\theta}_1^1$. Thus, under our assumption that θ is continuous and uniformly distributed between 0 and 1, the share of all women who are promoted, regardless of future childbearing status, $(1 - \overline{\theta}_1^*)$, is strictly lower than the share of men who are promoted, $(1 - \overline{\theta}_1^1)$. The magnitude of the gender penalty in period 1 is $\overline{\theta}_1^* - \overline{\theta}_1^1$. QED.

Proof of Proposition 1(ii): In period 2, fathers have strictly positive promotion probabilities to jobs 1 and 2, under the restricted set of paramaterizations described in Section 5.1, and under the assumption that $f'(\cdot) > 0$. Since the probability of taking leave for fathers is normalized to zero, fathers and childless men are promoted at identical

rates. Specifically, employers promote share $(1-\overline{\theta}_2^2)$ of both fathers and childless men to job 2, and share $(\overline{\theta}_1^1-\overline{\theta}_2^1)$ to job 1 for the first time. By contrast, share p_f women have a child and go on leave, and are promoted with probability zero, as the cut-off value for effective ability determining promotion for these women is infinite. Therefore, mothers experience a dramatic drop in promotion rates relative to fathers in period 2. Since childless women have zero probability of taking leave and, like men, have a strictly positive probability of being promoted, women who give birth in period 2 similarly have lower promotion rates than childless women. QED.

Proof of Proposition 1(iii): We begin by considering the first part of Proposition 1(iii), that women who never have children experience a higher rate of promotion than men after all childbearing decisions have been revealed. This corresponds to period 2 of the model. After the start of period 2, there is no more positive probability of future childbearing, and firms subsequently apply the same effective ability thresholds determining promotion for both men and childless women in periods 2 and 3. The period 2 promotion rates for men have been derived above. Among childless women, share $(1-\overline{\theta}_2^2)$ are promoted to job 2, same as for men. For promotion rates to job 1 in period 2, there are two possible cases for childless women. For the case that $\overline{\theta}_1^* < \overline{\theta}_2^2$, childless women's promotion rate to job 2 will be $(\overline{\theta}_1^1-\overline{\theta}_2^1)+(\overline{\theta}_1^*-\overline{\theta}_1^1)$. For the case that $\overline{\theta}_1^* > \overline{\theta}_2^2$, childless women's promotion rate to job 1 will be $(\overline{\theta}_1^1-\overline{\theta}_2^1)+(\overline{\theta}_2^2-\overline{\theta}_1^1)$. Note that in both cases, the second term – $(\overline{\theta}_1^*-\overline{\theta}_1^1)$ (case 1) or $(\overline{\theta}_2^2-\overline{\theta}_1^1)$ (case 2) – is greater than zero, indicating that women's promotion rate is higher than men's and that they experience a positive "gender effect" in this period.

Next, we consider the second part of Proposition (iii): relative to fathers, women who had children experience both a negative "motherhood penalty" in promotion rates in period 3, as well as a positive "gender effect." We begin by deriving the cut-off values for θ_τ^j in period 3 for men. Both fathers and childless men have $\tau = 3$ years of experience, and share $(1-\overline{\theta}_3^3)+(\overline{\theta}_2^2-\overline{\theta}_3^2)+(\overline{\theta}_2^1-\overline{\theta}_3^1)$ of men are promoted to jobs 3, 2, and 1 for the first time. By contrast, women who had children enter with only $\tau = 2$ years of experience, since they did not accumulate human capital in period 2. This lowers their effective ability relative to men, and therefore their promotion probability, generating a "motherhood penalty." In particular, no women with children are promoted to position 3 this period. However, mothers in period 3 also experience a positive "gender effect" from the fact that uncertainty around their childbearing and associated labor supply has now been resolved, which increases their probability of promotion in the current period. In particular, share $(1-\overline{\theta}_2^2)+(\overline{\theta}_1^1-\overline{\theta}_2^1)+(\overline{\theta}_1^*-\overline{\theta}_1^1)$ are promoted for the case that $\overline{\theta}_1^* \leq \overline{\theta}_2^2$, and share $(1-\overline{\theta}_2^2)+(\overline{\theta}_1^1-\overline{\theta}_2^1)+(\overline{\theta}_2^2-\overline{\theta}_1^1)$ are promoted for the case that $\overline{\theta}_1^* > \overline{\theta}_2^2$. In both cases, the last term corresponds to women who were initially "passed up" for promotion in period 1, but now advance to a higher position. Thus, alongside the motherhood penalty, these women experience a reversal of the gender penalty in period 3. QED.

B.3 Proof of Proposition 2

The proof of proposition 2 is a direct implication of equation (10). In equation (10), $\overline{\eta}^*$ is increasing in k_1-k_0. This implies that $\overline{\theta}_1^* = \frac{\overline{\eta}^*}{f(1)}$ is increasing in k_1-k_0, and therefore

also the magnitude $\overline{\theta}_1^* - \overline{\theta}_1^1$ of the gender penalty. QED.

B.4 Derivation of Wages

Under the assumptions of homogeneous firms, free entry into production, labor as the only input, and costless switching of workers between firms, men are compensated according to their production, $w_{ijt}^m = d_j + c_j\eta_{it}$. Any compensation above this wage would lead to negative profits for the firm, while any compensation below this wage would allow a competing firm to offer $\varepsilon > 0$ higher wages to attract the worker and still make positive profit. Therefore, in equilibrium, men are always paid $w_{ijt}^m = d_j + c_j\eta_{it}$, and firms earn zero profits.

For women, the wage function is more complex, since some women do not work in period 2 – leading to output of $-k_j$ – but cannot be fired in this period. We begin by describing their wages in period 3, which is the most straightforward period as there is no uncertainty for the employer about women's childbearing or labor supply, and therefore about incurring cost k_j. Consequently, in period 3 women are also compensated according to their production in equilibrium, $w_{ijt}^f = d_j + c_j\eta_{it}$, similar to men, following the same reasoning as above. The same is true in period 2 for women who remain childless and supply a unit of labor, as there is no further uncertainty about their current or future labor supply. Women who have a child and go on leave in period 2 are on government-paid leave by assumption, and thus are not offered a wage by the firm that period.

Finally, we consider women's wages in periods 0 and 1. To derive wages for these two periods, we require an additional assumption that there is a strictly positive probability, $\epsilon > 0$, that an individual is separated from their employer at the end of period 0, and works for a new employer in period 1. This assumption is needed to resolve an indeterminacy issue which we discuss shortly, after describing the wage functions. As job switching rates are very high at young ages, we do not view this as a restrictive assumption.

In period 1, employers who hire a woman to work in job j anticipate that they will incur cost k_j in with probability p_f in period 2, under our assumption that demotion or firing based on childbearing is not possible, in line with Swedish labor laws. Expected profit Π_{j1} for the firm of hiring a female worker in job j in period 1 is therefore

$$\Pi_{j1} = d_j + c_j\eta_{it} - w_{ij1}^f + p_f \cdot (-k_j) + (1 - p_f) \cdot \pi_2^*$$

where π_2^* indicates expected profit from the female worker if she remains childless (and thus works) in period 2. Since firms receive zero expected profits in equilibrium, both π_2^* and $\Pi_{j1} = 0$. Consequently, $w_{ij1}^f = d_j + c_j\eta_{it} - p_f k_j$.

Finally, it is important to note that without the additional assumption we introduced above – that there is a strictly positive probability, $\epsilon > 0$, that an individual changes employer from period 0 to period 1 – both employers and female workers would be indifferent between the following contracts: one that reduces women's wage in period 0 by $p_f k_j$, but not in period 1; one that alternatively reduces wages in period 1 by $p_f k_j$, but not in period 0; or one that splits the cost $p_f k_j$ across the two periods. However, when there is a positive separation probability ϵ, the wage reduction of $p_f k_j$ can only occur in period 1, and women are paid the same wage as men in period 0. The reason

for this is that for women, a contract with a wage penalty that is incurred partly or fully in period 0 is strictly inferior: following a separation shock, these women would have to incur the same penalty again in period 1, since no new firm would hire them at a wage above $d_j + c_j\eta_{it} - p_f k_j$. In equilibrium, firms will try to attract female workers by offering them a higher wage in period 0, up to $d_1 + c_1\eta_{i0}$, with incidence of the penalty $p_f k_j$ falling entirely in period 1. Note that this is true whether one assumes a strictly positive separation rate only at the end of period 0, or at the end of all periods.[42].

To summarize, the wage function for men is identical to that in the benchmark environment: $w^m_{ijt} = d_j + c_j\eta_{it}$. Women's wage function varies by period and takes the following form:

Period 0:	$w^f_{ij0} = d_0 + c_0\eta_{i0}$	
Period 1:	$w^f_{ij1} = d_j + c_j\eta_{i1} - p_f k_j$	
Period 2:	$w^f_{ij2} = d_j + c_j\eta_{i2}$	if childless, on govt.-paid leave otherwise
Period 3:	$w^f_{ij3} = d_j + c_j\eta_{i3}$.	

B.5 Wage Growth During Firm Changes and in Non-Promotion Periods

With the above wage functions in hand, it is possible to analyze wage growth both during firm changes and in non-promotion periods. To make such an analysis of firm changes possible, we introduce an assumption that there is a positive separation probability $\epsilon > 0$ at the end of every period, requiring workers to switch firms. A model with homogeneous firms, of course, is not designed to study firm-to-firm moves in depth. Nevertheless, the different predictions for wage growth in equilibrium for men and women exogenously separated from their firm is instructive.

Under a positive separation probability, wage functions are identical to those derived earlier, except that for women in period 1, the wage would be $w^f_{ij1} = d_j + c_j\eta_{i1} - p_f k_j(1-\epsilon)$, for the reasons described in footnote 42. Women who are separated from their firm at the end of period 1 and have a child in period 2 will not be hired, and will simply be on leave that term. All other individuals in all other periods will move to an identical new firm following separation. The firm assigns them either into the same job as before, or promotes them to a higher position, according to the same cut-offs as derived previously.

The main result we can obtain for wage growth conditional on firm changes is that women's wage growth in period 1 is strictly lower than men's, but can exceed men's wage growth later in the lifecycle. Women's wages grow more slowly than men's during a firm change in the pre-birth period (i.e., from period 0 to period 1) for two reasons. First, wages for men and women of the same ability in the same job are identical in period 0, but differ by $p_f k_j$ in period 1, in men's favor, as shown by the equations above. Therefore, there is a corresponding wage growth differential for men and women who are assigned to the same job in period 1 and who have the same η_{i1}. Additionally, some high-ability women who would have been promoted if they were men do not advance to job 2 in

[42]In the latter case, the wage for women in period 1 would be $w^f_{ij1} = d_j + c_j\eta_{i1} - p_f k_j(1-\epsilon)$, since employers take into account the now lower probability of incurring cost k_j in the following period.

period 1, further reducing their wage growth that period relative to men. The total effect is that women's wage growth in period 1 is strictly lower than men's.

To see this formally, consider that average wage growth from period 0 to period 1 can be expressed as average wage growth conditional on staying in job 1, times the share of individuals who stay in job 1, plus average wage growth conditional being promoted, times the share promoted. Recall that individuals start with zero effective ability in period 0, earning d_1, but accumulate one year of experience after the first year. Without loss of generality, we normalize $f(1) = 1$. Since innate ability is assumed to be distributed uniformly, average wage growth for men can therefore be expressed as:

$$\Delta\overline{w}_1^m = \left(c_0 \cdot \frac{1}{2}\left(0 + \overline{\eta}^1\right)\right) \cdot \overline{\theta}^1 + \left(d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^1 + 1\right)\right) \cdot (1 - \overline{\theta}^1)$$

By contrast, women experience both a different cut-off for promotion, and the penalty $p_f k_j$. Therefore,

$$\Delta\overline{w}_1^f = \left(c_0 \cdot \frac{1}{2}\left(0 + \overline{\eta}^*\right) - p_f k_0\right) \cdot \overline{\theta}^* + \left(d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^* + 1\right) - p_f k_1\right) \cdot (1 - \overline{\theta}^*)$$

We can re-write both of these expressions as follows:

$$\begin{aligned}
\Delta\overline{w}_1^m &= \left(c_0 \cdot \frac{1}{2}\left(0 + \overline{\eta}^1\right)\right) \cdot \overline{\theta}^1 + \left(d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right)\right) \cdot (\overline{\theta}^* - \overline{\theta}^1) + \\
&+ \left(d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^* + 1\right)\right) \cdot (1 - \overline{\theta}^*) \\
\Delta\overline{w}_1^f &= \left(c_0 \cdot \frac{1}{2}\left(0 + \overline{\eta}^1\right) - p_f k_0\right) \cdot \overline{\theta}^1 + \left(c_0 \cdot \frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right) - p_f k_0\right) \cdot (\overline{\theta}^* - \overline{\theta}^1) + \\
&+ \left(d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^* + 1\right) - p_f k_1\right) \cdot (1 - \overline{\theta}^*)
\end{aligned}$$

Note that the first and third term are strictly lower for women than men, since $p_f k_j > 0$. The second term is also strictly lower for women than for men, since $p_f k_0 > 0$, and additionally $d_1 - d_0 + c_1 \cdot \frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right) > c_0 \cdot \frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right)$, or alternatively $\frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right) > \frac{d_0 - d_1}{c_1 - c_0}$. The final inequality follows from: $\frac{1}{2}\left(\overline{\eta}^1 + \overline{\eta}^*\right) > \overline{\eta}^1 = \frac{d_0 - d_1}{c_1 - c_0}$.

While women's wage growth in period 1 is *strictly* lower than men's, their wage growth during firm changes later in the lifecycle *can* exceed men's. Consider a woman who was on leave in period 2, with a most recent wage of w_{ij1}^f, which contains the wage penalty $p_f k_j(1 - \epsilon)$. Since there is no more uncertainty about her childbearing, her wage is now bid back up by this amount. She will be paid the same wage paid to men who have the same ability and are assigned to the same job; additionally, some of the women passed up for promotion in period 1 are now promoted. However, note that in period 3, these women also have less accumulated human capital than men, potentially driving down their promotion probabilities relative to men. For this reason, we can only say that wage growth in period 3 *can* be higher for women than men during firm changes, since this will not be true for some parameterizations of the model. Therefore we do not present a formal proposition, but only demonstrate that the result is true for some parameterizations, such as the one in Figure 11.

In the case when workers do not change firms and are also not promoted, a similar pattern holds. The wage functions derived earlier show that men who were not promoted in period 1 experience wage growth that is $p_f k_1$ higher than that of women of the same ability. Additionally, under a large set of parameterizations, the *average* wage growth of non-promoted men from period 0 to period 1 will also be higher than corresponding *average* wage growth of non-promoted women, as in Figure 11B. The reason that this is not necessarily true for all parameterizations is that there are two reasons why men and women who were not promoted have different wage growth in period 1, which work in opposite directions. First, the women who are not promoted in period 1 have on average higher effective ability than men, since some high-ability women are passed up for promotion. This raises the average wages of women relative to men in the population that was not promoted, since wages depend on ability. On the other hand, for any given ability, the women who were not promoted still incur a penalty of $p_f k_1$, lowering their wages relative to men. Finally, later in the lifecycle, women's wage growth in non-promotion periods can exceed that of men, as in as in Figure 11B, since women's wages are bid back up by firms once uncertainty about future childbearing and leave-taking has been resolved.

B.6 Relaxing Assumptions

In the model presented above, we made several assumptions that kept the analysis tractable. We now revisit these assumptions, and what the model would predict when they are relaxed. Specifically, we relax the following assumptions: that individuals have a zero probability of having a child in period 1; that ability and probability of childbirth are uncorrelated; and that men and women both supply a unit of labor inelastically in periods 0, 1, and 3.

It is straightforward to see that the first two assumptions do not drive any of the results. By introducing a positive probability of childbirth and labor supply reduction in period 1, we simply introduce the possibility of incurring a "motherhood penalty" one period earlier, since those women who have children and take leave in period 1 would be promoted with probability zero. All remaining (childless) women would incur the same gender penalty in period 1 as in the present version of the model, since employers continue to expect a positive probability of incurring k_j in period 2 for these women. The second simplifying assumption – that a woman's innate ability θ_i is independent of the probability of having children and taking leave – is also an innocuous assumption. Suppose instead that $p_f(\theta)$ is continuous and decreasing in θ, so that higher ability women have lower probability of having a child and taking a leave. As long as $p_f(\theta) > 0$ for all θ, it follows immediately from the model that the threshold for promotion for women in period 1 must still be higher than the threshold for men.

Our third simplifying assumption concerns men's and women's labor supply each period. As in the benchmark model by Gibbons and Waldman (1999), there is no effective labor supply decision in our model. Specifically, we assume that individuals always supply one unit of labor except in period 2, when women who give birth reduce their labor supply (to zero) with an exogenous probability. As a result, women who ever have children work are assumed to work full-time in periods 0, 1 and 3, same as men, and also to accumulate

a year of human capital in each of these periods. This means we exclude the possibility of working part-time, for example, and any costs to the employer associated with part-time work. These are potentially strong assumptions which we now examine.

We first note that the assumption that men and women supply the same amount of labor in period 3 is in fact innocuous. First, this is the final period, meaning that human capital accumulation in this period does not matter for future promotions. Second, cost k_j is incurred by the employer only if the employee is on leave. Therefore, firms do not have any incentive to penalize women for working $1 > h > 0$ hours in period 3, and women are simply paid for their output, $(d_j + c_j\eta_{ij})h$. All results go through as before.

Of course, the model could be made more realistic by allowing k_j to be a smooth, decreasing function of hours worked, rather than to jump discretely at zero working hours. This introduces a non-linearity also in strictly positive hours worked. For example, suppose a function for $k_j(h)$ such that $k_j(0) > k_j(0.5) > k_j(1) = 0$ for all j, where $h = 0.5$ indicates part-time work. Additionally, $k_{j+1}(h) > k_j(h)$, for $h < 1$. The idea behind such an assumption would be that it is costlier for the employer to have a part-time manager, than a part-time rank and file worker. In this new environment, the model's main conclusions are still not affected. To see why, consider what happens if, with some positive probability, women with children experience a taste shock such that they choose to work part-time ($h = 0.5$) in period 3. The model would predict, in line with the data, that promotions would be lower in period 3 for women who choose to work part-time than for those who choose to work full-time, since employers want to avoid the higher cost $k_{j+1}(0.5) > k_j(0.5)$ and thus raise promotion thresholds for these women. Indeed, this would amplify the total motherhood penalty in period 3. However, a reversal in the gender penalty would still be observed after childbearing decisions have been revealed. This is obviously true for women who do not have children, and for those with children who work full-time in period 3, since employers no longer have any possibility of incurring a positive cost $k_j(0)$ or $k_j(0.5)$. However, a positive gender effect is observed even for women who decide to work part-time in period 3, as long as $k_j(0) > k_j(0.5)$, since employers no longer have a possibility of incurring cost $k_j(0)$.

Finally, our assumption that men and women supply the same amount of labor exogenously in the pre-birth periods (both periods 0 and 1) is the strongest one we make, and not entirely innocuous. Indeed, one possible alternative explanation for the patterns we observe is that women, in anticipation of future labor supply reductions, have substantially less incentive to work high hours than men in the pre-birth periods, even period 0, and therefore accumulate less human capital and experience lower promotion rates already in period 1. In the discussion below we examine the credibility of the assumption that men and women supply the same labor in the pre-birth period, and implications of the model when it is weakened.

First, we note that the assumption is not obviously counterfactual. While part-time work is not uncommon after first birth for women, only 6% of women in our (high skill) population work part-time prior to first birth, compared to 4% of men. For this reason, controlling for part-time work virtually does not affect the pre-birth promotion gap we document in Section 4.3. Nevertheless, our measures of hours may miss hours worked above full-time, which may be an important determinant of promotion. If men

are far more likely to work overtime hours that are not observed in the data, then their human capital accumulation will also be higher early in the lifecycle and prior to first birth, and consequently their promotion rates. One interesting and important question is what predictions our model would generate if individuals instead *chose* their labor endogenously in periods 0 and 1, based on anticipated future labor supply.

To credibly address this question, we must extend the model presented in Section 6.2 to include both disutility from hours worked as well as human capital accumulation that depends on hours worked. For this analysis, we also model parental leave benefits for women, with compensation based on the prior period's income at an 80% replacement rate, as in Sweden.

Consider the following environment, in which individuals choose hours worked h_{it} each period. Individual utility each period is equal to

$$u_{it} = c_{it} - \frac{1}{2}\gamma(z_{it})h_{it}^2,$$

where c_{it} is consumption. Disutility from work, $\gamma(\cdot)$ depends on z_{it}, where z_{it}=1 in period 2 if an individual is a woman and has a child, and is equal to 0 otherwise. Thus, for men, z_{it}=0 in all periods. We assume that $\gamma_i(1) = \infty$, implying that women always work zero hours with probability of one, and thus always take parental leave. For simplicity, we assume that the parameters for γ_i are such that men always work and do not take parental leave. Thus, we do not model a parental leave option for men. This is not an issue for this analysis, since we are specifically interested in whether women choose to supply less labor than men in periods 0 and 1 due to their anticipated labor supply reductions.

As before, effective ability is a function of innate ability and the stock of human capital. The latter depends positively on hours worked each period, so that effective ability is now equal to

$$\eta_{it} = \theta_i f(h_0, ..., h_{t-1}),$$

for $t > 0$. Wage functions take a similar form as before. A man in period t is paid $w_{ijt} = d_j + c_j\eta_{it} = d_j + c_j\theta_i f(h_0, ..., h_{t-1})$. Women who are not on leave are also paid according to this wage function in all periods except the first period, when their wage is additionally reduced by $p_f k_j$, as discussed in Section 6.2. To keep the notation as simple as possible, we represent the per period wage by $w_{it} = \omega(h_0, ..., h_{t-1}; g_i, \theta_i, t)$, where g_i refers to gender.

Mothers on leave in period 2 do not receive a wage, but instead receive an income modeled after the Swedish parental leave system. The income is calculated based on the woman's total earnings in period 1, so that $y_{i2}^{PL} = \phi w_{i1} h_{i1}$ if $z_{i2} = 1$ and $h_{i2} = 0$, and is equal to zero otherwise; $\phi = 0.8$, and corresponds to the Swedish replacement rate of 80%.

Finally, workers (but not employers) know with certainty whether they wish to have a child and whether they will reduce their labor supply in period 2. Workers in period 0

solve the following maximization problem:

$$\max \sum_t \left(c_{it} - \frac{1}{2}\gamma_i(z_{it})h_{it}^2 \right)$$

$$\begin{aligned} \text{s.t.} \quad c_{it} = y^{PL} \quad &= \phi w_{i1} h_{i1} && \text{if } z_{it} = 1, \\ c_{it} = w_{it} h_{it} &= \omega(h_0, ..., h_{t-1}; g_i, \theta_i, t) h_{it} && \text{otherwise.} \end{aligned}$$

To see how labor supply differs in periods 0 and 1 for women who anticipate taking time off in the future, as compared to men who do not, we derive the first order conditions in each period determining labor supply, for interior solutions. In period 2, of course, we have a corner solution for women who give birth, and their labor supply is zero that period.

In the final period, the first-order condition for both men and women simply equates wage to disutility from work, i.e. $w_{i3} = \gamma_i(0)h_{i3}$. However, in the remaining periods, individuals also take into account that their hours today affect future earnings through human capital accumulation, or potentially through their pay while on parental leave. For men (as well as for women who choose to remain childless), the first order conditions imply the following equations determining labor supply in periods 0, 1, and 2:

$$\gamma_i(0)h_{i2} = w_{i2} + h_{i3}\frac{d\omega(h_0, h_1, h_2; g_i, \theta_i, 3)}{dh_{i2}} \tag{M2}$$

$$\gamma_i(0)h_{i1} = w_{i1} + h_{i3}\frac{d\omega(h_0, h_1, h_2; g_i, \theta_i, 3)}{dh_{i1}} + h_{i2}\frac{d\omega(h_0, h_1; g_i, \theta_i, 2)}{dh_{i1}} \tag{M1}$$

$$\begin{aligned} \gamma_i(0)h_{i0} = w_{i0} &+ h_{i3}\frac{d\omega(h_0, h_1, h_2; g_i, \theta_i, 3)}{dh_{i0}} + h_{i2}\frac{d\omega(h_0, h_1; g_i, \theta_i, 2)}{dh_{i0}} + \\ &+ h_{i1}\frac{d\omega(h_0, h_1; g_i, \theta_i, 1)}{dh_{i0}} \end{aligned} \tag{M0}$$

By contrast, for women with children, $h_{i2} = 0$ and the first order conditions for periods 0 and 1 are as follows:

$$\gamma_i(0)h_{i1} = w_{i1} + h_{i3}\frac{d\omega(h_0, h_1, h_2; g_i, \theta_i, 3)}{dh_{i1}} + \phi w_{i1} \tag{F1}$$

$$\gamma_i(0)h_{i0} = w_{i0} + h_{i3}\frac{d\omega(h_0, h_1, h_2; g_i, \theta_i, 3)}{dh_{i0}} + h_{i1}(1+\phi)\frac{d\omega(h_0, h_1; g_i, \theta_i, 1)}{dh_{i0}} \tag{F0}$$

A comparison of (M1) and (F1) provides insight into what drives men's and women's labor supply in period 1, and what a reasonable set of parameterizations would imply for gender differences in hours worked that period. In fact, the two conditions are quite similar, except for the third term on the right-hand-side in (F1) and (M1). For men, the third term represents the additional expected earnings from the human capital accumulated from higher hours today. For women who anticipate having children, the third term represents the additional expected earnings from parental leave benefits. Recall that the value of ϕ in the Swedish system is around 0.8. In other words, women's direct compensation per unit of labor in period 1 is, effectively, not w_{i1}, but $1.8w_{i1}$. Since $0 < h_{i2} < 1$, the slope on the human capital accumulation function would have to be implausibly high,

in order for the third term in (M1) to exceed $0.8w_{i1}$. Similarly, in period 0 women, compared to men, do not anticipate that their human capital accumulation today pays off through period 2 wages, but instead anticipate that it pays off through higher future parental leave benefits.

Whether the first order conditions derived above imply higher hours worked for women or men in the years prior to first birth is an empirical question that requires a structural estimation of the parameters and, in addition to the factors mentioned above, will also depend on differences men's women's hours worked in period 3, h_{i3}, and differences in period 1 wage, w_{i1}. However, the equations above indicate that it is not at all obvious that women have less incentive to work high hours than men early in the lifecycle. Under reasonable parameterizations, women may even have greater incentive to do so under the replacement rates in the Swedish parental leave system. For this reason, and the fact that we observe almost universal rates of full-time work for both men and women in the pre-birth period in the data, the assumption that men and women supply a unit of labor inelastically in periods 0 and 1 is, in our view, well-motivated. We leave a further investigation of these issues for future research.

Appendix C Wages: Sampling and Weighting

For our analysis, we merge data from Wage Statistics Sweden with the main register data, LOUISE. LOUISE covers the entire adult Swedish population. Wage Statistics Sweden collects information about wage (full-time equivalent monthly earnings) and contracted hours once yearly for employees with positive hours in the survey month. All public-sector employees and all workers at firms with at least 500 employees are surveyed, while firms with fewer than 500 employees are only sampled each year. At sampled firms, all workers with positive hours are surveyed. We observe wages for about 60% of workers every year.

In this appendix, we consider issues around sampling and weighting relevant to our analysis. Note that small firms must be sampled for two consecutive years for us to calculate wage growth for their employees. Generally, standard sample weights will not fully account for this lower probability of appearing in the data two years in a row. Similarly, workers who switch either to or from a small firm in a given year will also appear in our data for two consecutive years with a reduced probability. In what follows, we test sensitivity of our results to alternative weighting procedures, including those that take into account the reduced probabilities of observing smaller firms in two consecutive periods.

We then conduct additional robustness tests. In particular, we verify whether the population that we observe in Wage Statistics Sweden – and for whom we construct wage growth statistics – has income profiles that are representative of those of the entire Swedish population. For this purpose, we rely on the variable total annual labor income, which is contained in LOUISE and therefore available for all adults.

C.1 Wage Growth Distributions Under Alternative Sampling Weights

To analyze how alternative sample weighting procedures affect wage growth patterns, we focus on the full distribution of real annual wage growth, as in Figure 2 in the main text. The wage growth distributions of men and women, without using any weights, are pictured in Panel (i) of Figure C.1. Panel (ii) replicates the same histogram, but using the standard weights provided by Wage Statistics Sweden. Notably, using the provided sample weights does not affect the distribution in any meaningful way.

Next, in Panels (iii) and (iv) of Figure C.1, we construct our own set of weights, using information on firm size of workers, which is found in the firm register and therefore available for all working adults. We refer to these weights as FS-1 and FS-2 in the figures. To construct the weights used in Panel (iii), we group firms by number of employed workers into the following 10 categories: 1 to 9, 10 to 34, 35 to 99, 100 to 249, 250 to 499, 500 to 999, 1000 to 1999, 2000 to 4999, 5000 to 9999, and 10,000+ employees. In each firm size category, we calculate the probability that workers at those firms appear in the wage statistics data. Next, we assign a weight to each individual that is the inverse of that probability, giving us the weights corresponding to FS-1 in Panel (iii). As expected, results using these alternative weights correspond closely to those in Panel (ii), since the sample weights provided by Wage Statistics Sweden are also based primarily on firm size.

Finally, we construct weights that take into account the reduced probability that

an individual at a firm of a given size is observed two periods in a row. Specifically, we construct a 10x10 matrix, using the same categories as before, where each cell corresponds to firm size category in period t and firm size category in period $t-1$. For every cell, we then calculate the probability of observing an individual in the wage statistics data in both periods. The final sampling weight assigned to each individual is the inverse of that probability, yielding weights FS-2. Results using those weights are graphed in Panel (iv) of Figure C.1. Note that these weights are the most appropriate ones when analyzing wage growth, since they account for the probability that a worker at a firm of a given size is observed in the wage data for two consecutive periods.[43]

Figure C.1 shows that the wage growth patterns we document are not sensitive to alternative weighting procedures. For easier visual inspection, Figure C.2 jointly graphs the baseline distribution from Panel (i) of Figure C.1, using no weights, with that from Panel (iv) of Figure C.1, which uses the weights that take into account probabilities of being sampled two years in a row based on firm size. As Figure C.2 shows, the two sets of distributions do not differ in any meaningful way. Therefore, we present unweighted results and do not repeat tables and figures for alternative weights.[44] However, as the above figures make clear, the wage growth patterns, and therefore our results, are robust to alternative weighting procedures.

C.2 Distributions of Two- and Three-Year Wage Growth

Next, we analyze two-year and three-year wage growth distributions to verify that we do not miss important wage growth patterns by focusing strictly on year-on-year wage growth. This analysis is especially relevant for the left tail of the distribution. For example, we may miss negative wage growth for individuals who separate from a firm for some period of time and therefore do not appear in two consecutive years in our data. Figure C.3 graphs histograms of two- and three-year individual wage growth, with an adjusted x-axis to reflect the longer periods under consideration. As the figure shows, the qualitative results are identical to the ones in Figure 2. In particular, men and women have similar incidence of negative wage growth, with men slightly more likely to experience significant negative growth. As before, the distributions exhibit high kurtosis and right skewness, with men substantially more likely to experience high growth in the right tail of the distributions.

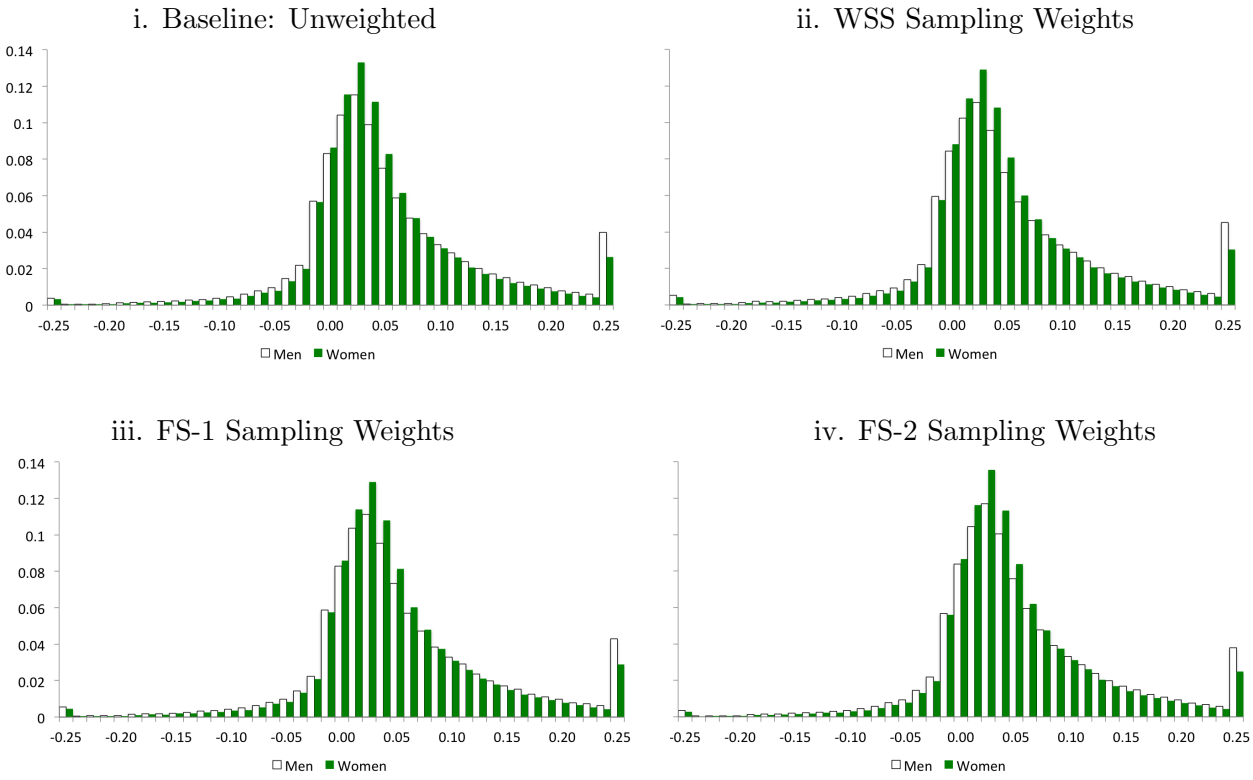
C.3 Total Annual Labor Income: Full Population vs. Individuals in Wage Statistics

Lastly, we verify whether individuals for whom we have wage data are representative of the entire population of Swedish workers by comparing total annual labor earnings for the two groups. Figure C.4 graphs earnings profiles for the two groups. While men's earnings

[43] We have experimented both with varying the number of firm size categories as well as the cut-offs for the firm categories, and in all cases the wage growth distributions look similar. Results available upon request.

[44] The choice to present unweighted results follows party from our conversations with Swedish scholars. There is some disagreement about the application even of sampling weights provided by Wage Statistics Sweden when studying a panel over time, and therefore a common preference for unweighted results. We did not find alternative weighting procedures, including those described above, to affect our results.

Figure C.1: Wage Growth Distributions Under Alternative Weighting Procedures



Notes: See note in Figure 2. WSS sampling weights refer to the weights provided by Wage Statistics Sweden. FS-1 and FS-2 are weights we construct based on information about workers' firm size, where FS-2 weights are adjusted for the fact that firms have to be surveyed for two consecutive periods to construct wage growth statistics for their employees. See text for details.

are marginally lower for those who appear in the Wage Statistics data, as compared to those in the whole population, overall the income profiles are nearly identical. For women, there are no noticeable differences.

Figure C.2: Comparison: Baseline (Unweighted) vs. FS-2 Sampling Weights

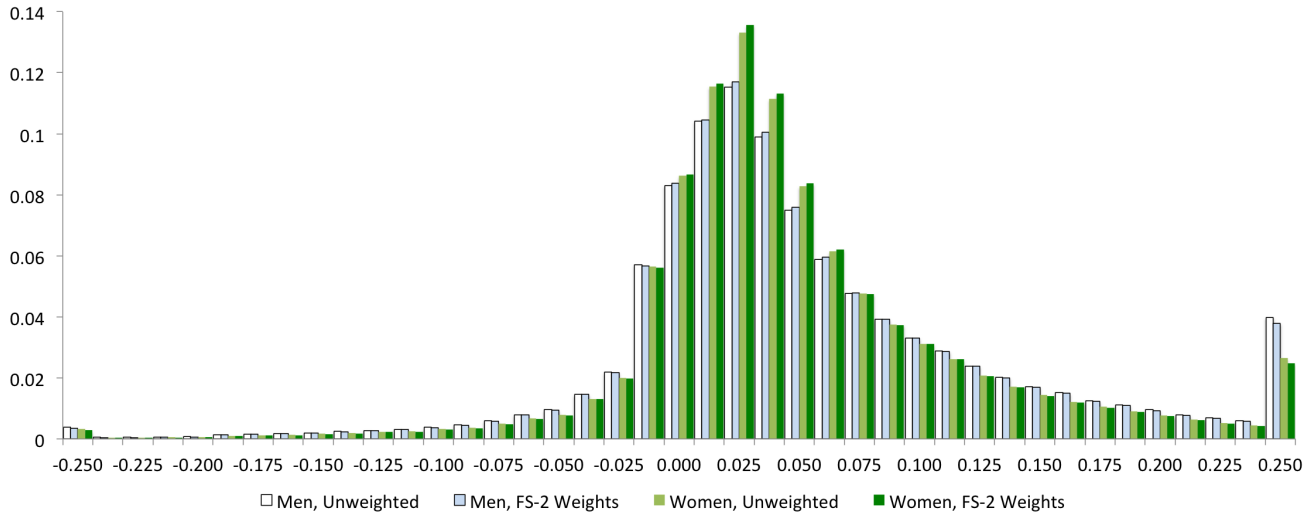
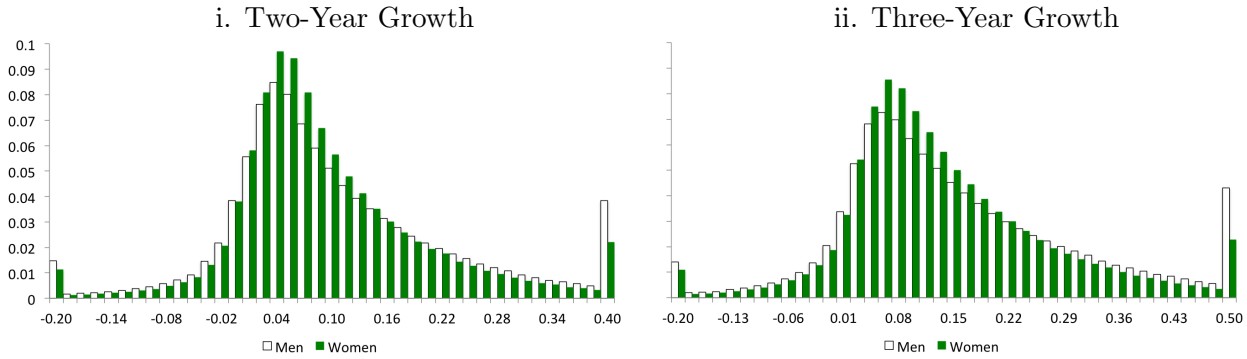
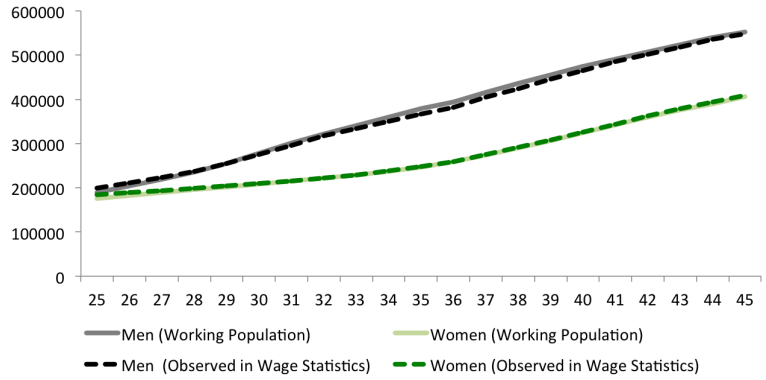


Figure C.3: Two- and Three-Year Wage Growth Distributions



Notes: The histograms tabulate person-year observations of two- and three-year real wage growth, for individuals ages 25 to 45. The tails of the distribution are collapsed to mass points.

Figure C.4: Annual Income of Full Population vs. Individuals in Wage Statistics



Notes: The working population is defined as anyone with annual labor income of at least 5000 Swedish krona. Individuals in Wage Statistics are those for whom we observe wages in two consecutive periods.