Abstract

Why do women’s wages grow more slowly than men’s? Using Swedish administrative data, we answer this question in three steps. First, we analyze men’s and women’s real annual wage growth non-parametrically. We show that men’s and women’s wage growth distributions differ principally in one aspect: women are less likely than men to experience persistent within-firm wage shocks in the right tail of the wage growth distribution. These shocks move workers up their firm’s wage hierarchy, resemble internal promotions, and play a primary role in driving the gender differences in wage growth. Based on this evidence, in the second step we analyze men’s and women’s within-firm mobility. Using a novel wage-based measure of within-firm mobility, we estimate that gender differences in the probability of experiencing large internal promotions account for around 70% of the total difference in men’s and women’s wage growth by age 45. We quantify the contribution of differences in human capital characteristics and occupation, sorting across firms, and hours worked and childbirth to the observed promotion gap. Alongside substantial motherhood penalties, we also document sizable dynamic gender penalties in promotion that are largest early in the lifecycle, reverse after age 40, and are observed both for women who eventually have children and those who remain childless. Lastly, to interpret our findings, we develop a model of promotion dynamics based on Gibbons and Waldman (1999). We conclude that the key empirical facts about promotion and wage growth are not readily explained by behavioral channels, such as gender differences in competitiveness or propensity to negotiate, but are consistent with costs to firms associated with employee leave-taking associated with childbirth, and employer uncertainty about women’s future childbearing and labor supply reductions.
1 Introduction

Why do women’s wages grow more slowly than men’s? In most countries and within most skill groups, men’s and women’s wage trajectories diverge over time, with parenthood contributing significantly to this lifecycle pattern. However, the underlying drivers of this divergence – and their quantitative importance – are still not well-understood. Do women frequently work part-time at firms with otherwise steep trajectories, leading to flatter wage growth? Do they commonly take wage cuts to work at lower-paying but, for example, more flexible firms? Are women promoted less, even when working the same hours as their male colleagues? If so, why would employers make such promotion decisions? Answers to these questions have important implications for labor market policies and for models of wage dynamics.

In this paper, we provide extensive evidence on these and related questions about wage growth within and across firms, and develop a theoretical framework that helps interpret this evidence. We use Swedish administrative data from 1985 to 2013, following men and women over their lifecycle until age 45. Our data is firm-worker linked and includes not only annual labor income, but also an administratively recorded wage variable, even for salaried workers, allowing us to study wage growth with minimal measurement error. Sweden also 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 sample selection or changes in composition of workers with age.

Our paper has three parts: two descriptive and one theoretical. Part one provides a comprehensive, non-parametric characterization of the key features of real annual wage growth. As our first contribution, we establish the main similarities and differences in men’s and women’s annual wage growth distributions. Our findings point to one predominant gender disparity: women have a far lower incidence of large, persistent wage increases in the right tail of the wage growth distribution. These increases are typically on the order of 15% or more, and occur primarily during years in which individuals do not switch firms. As both men and women experience real wage growth below 3% in the majority of years, these right-tail wage increases – which resemble sizable internal “promotions” – generate a disproportionate share of lifecycle wage growth. Our findings indicate that the differential incidence of such large within-firm wage increases account for the large majority of the gender gap in wage growth.

Based on this evidence, in the second part of the paper we characterize differences in men’s and women’s within-firm mobility. Exploiting the employer-employee linked

---

1 See, for example, Bertrand et al. (2010), Goldin (2014), Adda et al. (2017), Kleven et al. (2017).

2 Gender segregation across firms has been documented in several studies. See Hellerstein et al. (2008), Card et al. (2016), Albrecht at al. (2017), and Barth et al. (2017).
nature of our data, we develop a novel measure that identifies promotions based on large (e.g., 10+ log point) wage increases relative to other co-workers at the same firm, which we define formally in Section 3. We estimate that differences in promotion probability account for around 70% of the total gender gap in lifecycle wage growth, even under fairly conservative assumptions about what constitutes a promotion. As our second main contribution, we establish a novel set of stylized facts characterizing gender differences in lifecycle promotion patterns and their drivers, including differences in field of study, hours worked, occupational choice, dynamics around childbearing, and sorting across firms with different promotion opportunities.

Among the drivers we analyze, several that one might expect ex ante to be important – such as differences in choice of major (Bronson, 2015) or in sorting across firms (Hotz et al., 2018) – turn out to play only a small role in driving the promotion gap in the high skill population we study. However, two factors stand out, jointly accounting for around three-quarters of gender differences in promotion probability. The first is a dramatic disparity in promotions – between observationally similar men and women working at the same firm – occurring in the year of and year immediately following birth, when Swedish women are typically on parental leave. Foregone promotions in those isolated years alone account for about 40% of the gender difference in cumulative promotions by age 45, for individuals who ever have children, and generate a large wage differential relative to men within a short time-span. The second factor is a quantitatively large within-firm residual gender penalty in promotion probability that is incurred predominantly prior to first birth, is not associated with what the literature calls a “motherhood penalty” (e.g., Kleven et al. (2017), Angelov et al. (2016)), and is observed also for women who never have children. This gender penalty declines with age and reverses after 40 to favor women. Lastly, while part-time work also contributes to the promotion gap, it is quantitatively less important than either of the previous two factors. The reason for this is that most promotions occur early in the lifecycle, before Swedish women begin to work part-time at significant rates.

In the third part of our analysis, we construct a theoretical model of gender differences in careers within firms. The objective of this section is to interpret the empirical facts documented in parts 1 and 2 of the paper, and to provide a unified framework for them. To do this, we build on a classic model of promotions by Gibbons and Waldman (1999b), in which firms promote individuals of heterogeneous ability as they accumulate human capital and become more productive over time. To this model, we add two features to study gender differences. First, we add a lifecycle component for the worker, with a distinct childbearing phase that is associated with labor supply reductions, especially for women, as observed in the data. Second, we introduce a cost to employers when workers reduce their labor supply and take extended (e.g., parental)
leave, with the cost increasing in job responsibility. The idea behind this assumption is that employers commonly need to find or train temporary replacements for workers on leave, pay other workers overtime, etc.; moreover, such costs are likely to be higher when an employee is a manager who affects the productivity of many workers, vs. a rank-and-file employee.

The model generates rich predictions in line with the data. It produces a persistent “motherhood penalty” in promotion rates associated with childbirth. It also generates a “gender penalty” that initially favors men in the pre-childbearing years, as employers face uncertainty about which women will take a costly leave in the future, and thus promote a lower share of them. The penalty subsequently reverses after the end of the childbearing period, as uncertainty about future labor supply has been resolved, and women passed up for promotion early on advance up the career ladder. Next, we derive and empirically test a direct implication of the model: gender differences in promotion probability prior to first birth should be higher at firms where leaves are expected to be more costly. We provide evidence strongly in line with this prediction. Finally, we use the model to help interpret additional wage dynamics documented in the empirical analysis. We conclude by considering alternative theories about gender differences in wage growth. We argue that these theories have difficulty rationalizing the facts we establish about promotion and wage dynamics.

Our paper builds on the existing literature on gender differences in labor market outcomes (Altonji and Blank (1999), Bertrand (2011)) by providing a novel and detailed anatomy of gender wage growth differentials, and a theoretical framework to help organize this evidence. To the best of our knowledge, our study is the first to investigate wage dynamics by analyzing wage growth distributions non-parametrically, a feature made possible by our data. As such, our paper is closely related to Guvenen et al. (2014), who take a similar approach to document new stylized facts about earnings dynamics. However, they focus not on wage changes as we do, but on annual earnings changes, which exhibit different empirical patterns as they are driven in large part by changes in annual hours worked. Additionally, their paper is not concerned with gender differences or within- and across-firm growth. Our study is also closely related to Card et al. (2016), who study gender differences in sorting across firms and in returns to firm changes. They estimate, using a two-way firm and individual fixed effects model, that these two factors jointly account for about 20% of gender wage differentials in Portugal, a similar number to the one we obtain for Sweden using a less parametric approach. While they study the effect of firm changes on gender wage differentials, we focus primarily on analyzing within-firm growth, which drives the remaining 70-80% of gender wage differences by age 45.

Our paper also builds on the literature on careers within firms (Gibbons and Wald-
man (1999a), Prendergast (1999)). On the empirical side (e.g., McCue (1996), Blau and DeVaro (2007)), we extend this literature by providing comprehensive new evidence about differences in men’s and women’s within-firm mobility: precisely when these differences emerge over the lifecycle, and to what degree they are accounted for by differences in sorting across firms, human capital characteristics, part-time work, and dynamics around childbirth. One reason for this gap in the literature – which we discuss in greater detail in Section 3 – is that there are formidable challenges associated with measuring and therefore studying promotions in a representative population. As a consequence, theoretical research on gender differences in careers within firms has also been surprisingly limited. Key exceptions include Milgrom and Oster (1987), Lazear and Rosen (1990), and a more recent paper by Thomas (2018). The model we develop extends this theoretical literature by focusing on time-to-birth dynamics, gender vs. motherhood penalties in promotion, and wage growth differences also in non-promotion periods.

In the remainder of the paper, we first describe the data and Swedish institutional context (Section 2). In Section 3 we present our measures of internal mobility and discuss existing approaches used in the literature. Sections 4 and 5 present our empirical findings. Section 6 presents the theoretical model. Section 7 analyzes sensitivity of our results. Section 8 concludes.

2 Data and Institutional Context

Our data comes from several administrative registers from Statistics Sweden, covering the years 1985-2013. Our baseline dataset is the LOUISE register, which covers the entire population of Sweden aged 16-75. It includes demographic variables, such as 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. The first is Wage Structure Statistics register, which provides information on a worker’s contracted hours and contracted wages, measured as full-time equivalent monthly earnings. The second is the multi-generational register, which provides details about the dates of all births and hence, the number of children. The third is the employer register, which provides personal identifiers of all employees, as well as certain firm characteristics, such as size, industry, and whether the establishment is in the public or private sector. Using these linked registers, we can analyze firm-specific wage distributions and other firm characteristics by gender or education.

Gender differences in career dynamics are virtually not addressed in the key review pieces and handbook chapters on careers in organizations and personnel economics over the last two decades: e.g., Gibbons and Waldman (1999a), Prendergast (1999), Lazear and Oyer (2007), Gibbons and Roberts (2013).
An advantage of our administrative data is that it records wage — full-time equivalent monthly earnings — and not just the more standard annual labor earnings variable, which is comprised of the product of wage and hours, in addition to possible overtime pay and bonus compensation. Importantly, this allows us to separate annual changes in wages from annual changes in hours in our analysis of growth. This is especially 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. used by 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, and sample weights are provided.\footnote{Results are not dependent on weighting. Individuals who work at firms for which wage data is collected, but are on parental leave in the survey month, do not appear in the Wage Structure Statistics. For these individuals, we interpolate wages during the parental leave year in three ways. The first method, which we use when we report results, averages the wage from the prior and subsequent year. The second and third methods assign just the prior year’s wage, or just the subsequent year’s wage, respectively. None of our results are sensitive to the choice of interpolation.}

In order to analyze lifecycle dynamics, we follow college-educated individuals from cohorts born between 1960 and 1970. These years correspond to the youngest set of cohorts that we are able 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 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 work in the public sector. In Section 7, we provide results when all majors are included.
“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. 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.

3 Measuring Internal Mobility

Over the course of their firm tenure, workers commonly experience years of exceptionally high wage growth, often associated with some form of upward mobility, and periods of low interim growth. In this section, we describe how we use employer-employee matched wage data to measure within-firm mobility and specifically promotion, an important variable for our analysis in Section 5. As our approach is non-standard, we discuss at some length why we select it over more traditional methods. We first describe existing empirical approaches to studying internal mobility, and the challenges to obtaining reliable and transparent measures using these approaches. Next, we present our wage-based approach. Finally, we discuss the relative advantages and drawbacks of our alternative method.

---

5 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.
3.1 Assignment-Based Measures of Promotion

Existing empirical methods for identifying internal promotions rely either implicitly or explicitly on the idea that a change in job assignment is a key characteristic of an upward move within a firm. They fall into three broad categories.

The first method characterizes a career ladder in detail at one specific organization, as in the classic papers by Baker, Gibbs and Holmstrom (1994a,b). This “single-firm” approach allows researchers to develop deep familiarity with a firm, access organizational charts, and create reliable criteria to identify vertical vs. horizontal moves for all employees, as Baker et al. do for a large firm over a two-year period. The drawback of the method is that the firm (or time period) under study may not be representative, and therefore the approach is not well-suited to provide an empirical characterization of promotion dynamics for the general population or over the lifecycle.

The other two methods rely either on self-reported information in survey data to characterize a promotion, as in McCue (1996), or on administratively recorded job titles, as in van der Klaauw and da Silva (2008). A key advantage of these measures is their availability in a number of nationally representative datasets, allowing for population-wide inference. However, they also have substantial disadvantages for researchers. The main drawback of the self-reported measure is that it requires respondents to make subjective assessments about what constitutes a promotion. As Pergamit and Voym (1999) show using the NLSY, the criteria used to make this assessment vary significantly across workers: more than 50% of events that workers call promotions involve no change in position or duties, and a significant share are not accompanied by any wage increase. As a result, it is not clear that the events workers report as promotions are comparable across firms or demographic groups. A second, related consequence is that the criteria used by respondents generally differ from the set of criteria desired by the researcher, such as a change in job assignment.6

The main challenge for researchers when characterizing promotions based on a change in job titles is distinguishing between a lateral and vertical move. This is a general problem when identifying promotions either based on job title or occupational code changes. For example, is a change from “financial specialist” to “business specialist” an upward move? Additionally, job title changes commonly occur purely for administrative reasons, e.g. when a firm alters part of its organizational structure, or changes how it reports some job titles in administrative data. For a researcher who does not wish to classify any job title change as a “promotion,” using this method requires subjective assessments about possibly tens of thousands of job title changes.

---

6 For example, McCue (1995) assumes a change in job assignment characterizes the self-reported measure, writing that the measure’s advantage is to “identify[ ] promotions independently of...wage changes,” as a “change in tasks is what...distinguishes a promotion” (p. 176).
Studies that use data with more than one promotion measure are particularly instructive about these issues. For example, van der Klaauw and da Silva (2008) use administrative, population-wide data from Portugal. Their data has two measures of promotion: administratively recorded job titles, with detailed job descriptions, and employer-reported promotions, also administratively reported. An additional advantage of their data is that job titles are organized into hierarchical levels that are in principle associated with job complexity or responsibility. When comparing the two measures, the authors find that only 30% of hierarchical-level job title changes are also considered promotions by the employer. Additionally, 40% of employer-reported promotions are not associated with any change in detailed job description. These results indicate either that the measures capture very different types of internal mobility events; or, alternatively, that at least one, and possibly both measures capture internal mobility only with significant error. Because the criteria used to define a promotion are not transparent, it is difficult to assess what characterizes these errors, and how they may affect empirical results based on these measures.

3.2 A Wage-Based Measure of Promotion

In contrast to an assignment-based approach to measuring upward moves within firms, we develop a wage-based measure of internal mobility, which exploits information about individuals’ wage increases, relative to those of their co-workers. Rather than focusing on measuring moves up a traditionally defined career ladder, we study how individuals move through a firm’s wage hierarchy. There are a number of ways in which one could construct variables to study such moves. We rely on two variables that we believe are transparent and easily interpretable.

The first measure we construct is a continuous variable that we call “relative wage growth.” It 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. Specifically, for each individual in our population, 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 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 a large wage increase relative to other co-workers within the firm, i.e. “upward mobility.” The relative wage growth measure isolates this second factor, after accounting for systematic differences in average wage growth.

---

7Results are similar when the measure is constructed instead using median wage growth at the firm.
growth across firms in a given year.

The second measure we construct identifies years in which individuals experience especially large upward moves within their firm’s wage hierarchy, and corresponds to our measure of internal “promotion.” Specifically, the measure transforms 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. For reasons that will become obvious in Section 4, we rely primarily on the promotion measure throughout our analysis of internal mobility. Using this variable, we are also able to relate our findings to an existing empirical and theoretical literature that studies promotions primarily as binary events.

Throughout most of the paper, we set $n$ equal to 10 log points when constructing the binary measure, a fairly conservative threshold that focuses the analysis on large promotions. In Section 4 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. While occupational codes are available in our data, we do not use this information to construct our promotion measure. Using occupational codes to define promotions is problematic for two reasons. On the one hand, it is difficult to distinguish between lateral and vertical moves for many changes in occupational code. On the other hand, even detailed, four-digit codes are not sufficiently fine-grained to capture many promotions, e.g., from analyst to project manager to division leader, CFO to CEO, or assistant to associate professor. For this reason, a change in occupation code is neither necessary nor sufficient to indicate a year of upward mobility within the firm. In our data, promotions are commonly associated with changes in occupational code: individuals are about 1.8 times more likely to experience a change in occupational code during a promotion, as compared to years in which they were not promoted. However, because of the aforementioned problems with occupation codes, we do not explicitly use this information to construct our measure.

As our measures are unconventional in focusing on wage changes alone, we conclude by discussing the advantages (and disadvantages) relative to traditional measures for the questions of interest in this study, which are about characterizing internal mobility patterns and their drivers for a representative population.

### 3.3 Comparison of Approaches

As with the other measures, there are drawbacks of our wage-based measure of internal mobility. First, our measures specifically analyze workers’ movements through a firm’s *wage hierarchy*, rather than a traditionally defined career ladder. This may or may not be a disadvantage, depending on the research question. In practice, large relative wage increases and moves up a career ladder are highly correlated, as the literature on
careers in organizations documents (e.g., Gibbons and Waldman (1999a), Prendergast (1999)). However, we cannot rule out that some large wage increases captured by our binary measure are related, for example, to the arrival of favorable outside offers. This is not a significant issue for our empirical results, but affects the interpretation of our findings. We discuss this issue at greater length in Section 7.

Another drawback, specific to the binary version of our measure, is that its definition requires researcher discretion in specifying the relevant threshold, since the underlying variable, relative wage growth, is continuous. This concern can be mitigated by reporting results for a range of thresholds, as we do in this paper. We emphasize, however, that job complexity or responsibility is also a non-binary variable. Thus, when a promotion variable is based on a change in job duties or complexity, a threshold must similarly be applied by the researcher to identify a relevant change in this measure. This is often done implicitly, since an explicit threshold is difficult to specify for job responsibility, a multi-dimensional object. Specifying an explicit threshold (or set of thresholds) for a promotion measure, as we do, also has advantages, namely that it is possible to narrow down the types promotions under study. For example, one can focus the analysis on large upward moves with significant increases in compensation, and not conflate them with changes only in title, accompanied by marginal wage growth.

Relative to the other two methods using nationally representative data, the key advantage of our measure is that it is formally defined, precise about what it measures, and transparent. Our measure cannot be mapped with certainty to a change in job assignment. However, there is no uncertainty about what our measure identifies in the data: large wage increases relative to one’s co-workers at the same firm, where large is explicitly defined. By contrast, it is difficult to assess what types of events are captured by the self-reported or title-based measures, as discussed above. Of course, our measure will not be useful for some questions related to careers within firms. However, its transparency allows researchers to readily assess the set of questions it can address, as well the limitations of empirical analysis using the measure, as we do in this paper. A final advantage is that our approach is readily replicable across all countries in which employer-employee matched wage data is available. For questions such as ours – around gender differences in the labor market – this is particularly useful, as these are likely to be affected by policies that differ significantly across countries.

We conclude with a discussion of related prior studies. Numerous studies have analyzed, using assignment-based measures, whether promotion rates differ for men and women, as well as whether wage growth conditional on promotion differs by gender. However, there is substantial disagreement across studies even about the sign of these differences. For example, studies that find that women have lower promotion rates than men, as we do, include the well-known paper by McCue (1996), as well as Cabral
et al. (1981), and Olson and Becker (1983). On the other hand, Giuliano, Levine, and Leonard (2005) find that men and women have identical promotion rates, while Booth et al. (2005) find that women have higher promotion rates than men. The same is true for average growth associated with promotions. Cabral et al. (1981), Olson and Becker (1983), Giuliano, Levine, and Leonard (2005), and McCue (1996) document increases that are similar for men and women. On the other hand, Blau and Devaro (2007) find that women have higher returns to promotion, while Booth et al. (2005) find the opposite. The difficulties with measuring promotions, discussed above, provide one possible explanation for the range of findings documented in the literature.

Few studies quantify the contribution of promotions to wage growth over the lifecycle. The only available benchmark we know of comes from the study by McCue (1996). McCue finds that self-reported promotions account for around 15% of lifecycle wage growth for individuals in the PSID, for both genders and both college- and non-college educated individuals. However, McCue indicates that limitations to the PSID measure likely lead her to underestimate the contribution of promotions to lifecycle wage growth. Our own results indicate that promotions play a larger role, accounting for around 40 to 45% of lifecycle wage growth. Similarly, we document substantially larger gender differences in promotion rates than McCue, who finds that women’s annual promotion rates are around 0.5 percentage points lower than men’s. The differences between our results and McCue’s are consistent with the idea that traditional approaches are likely to be characterized by significant measurement error. However, our measure could also capture additional events – such as arrival of outside offers that bid up a worker’s wage – that are not associated with traditional moves up career ladders, an issue we discuss further in Section 7.

4 Empirical Analysis I: Gender Differences in Wage Growth

In this section, we establish four stylized facts about wage growth that, we will show, are crucial for understanding gender differences in lifecycle wage trajectories.

4.1 Non-Parametric Analysis

We begin by characterizing non-parametrically men’s and women’s real annual wage growth in Figure 2. The figure plots the distribution of real annual wage growth for individuals in our cohort for all years we observe them, between ages 25 and 45. Thus,

---

8Specifically, PSID interviewers first ask individuals whether they changed jobs in the past year. Only after responding affirmatively, individuals are asked whether they were fired, laid off, switched employers, promoted, etc. If individuals commonly interpret a change in jobs to mean a change in employer, as McCue indicates is likely, many promotions in this population will go unreported.
each point in the histogram is a person-year observation. Women’s wage growth dis-
tribution (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.\textsuperscript{9} As we are plotting real, rather than
nominal wage growth, the graphs feature a relatively high incidence of small negative
wage growth. These observations correspond to nominal wage growth 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
tough 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 changing to
lower-paying firms around time of birth (e.g., Hotz et al. (2017)), these changes are
not associated with a substantially greater incidence of years of high negative wage
growth over the lifecycle, 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 observa-
tions, 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 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).\textsuperscript{10}

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% times 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 differ-

\textsuperscript{9}For a more complete characterization of the tails, see Figure A.1 in the appendix.
\textsuperscript{10}By contrast, Guvenen, Ozkan, and Song (2014) document high reversion after sizable shocks to total
annual earnings. This difference between our results and theirs is to be expected. First, sizable earnings
shocks are driven in large part by fluctuations in hours, e.g. following job loss/recovery. Second, annual
earnings also include one-time bonus payments, not included in our wage measure. Additionally, Guvenen
et al. document that the distribution of total annual earnings growth is left-skewed. This is true also in our
data, even though the wage growth distribution is right-skewed.
ent percentiles, separately for men and women. The higher-order moments, presented at the bottom of Table 2, 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 for men, compared to 0.58 and 18.2 for women. The table also documents that men and women are equally likely to experience negative wage growth, with approximately zero real wage 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 (fact 1). The primary gender difference in wage growth patterns is that women are significantly less likely to experience these years of right-tail growth (fact 2).

### 4.2 Characteristics of Right-Tail Growth

Years of high wage growth are commonly associated in the literature either with firm changes (e.g., Burdett (1978), Topel and Ward (1992), Abowd et al. (1999)), or within-firm promotions (e.g. Baker, Gibbs and Holmstrom (1994), Gibbons and Waldman (1999a)). Therefore, we next analyze the importance of these factors in driving gender differences in right tail and, correspondingly, overall wage growth.

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

We begin by studying the incidence of firm changes in the right tail, for men and women. Column 1 of Table 3 shows the share of wage increases that occur during firm changes, in different wage growth intervals. For the same intervals, columns 2 and 3 display how likely men are to experience a wage increase of a given magnitude, as compared to women. Column 2 focuses on periods when individuals change firms, while column 3 focuses on periods when they do not.
The table features two main findings. First, the majority of right-tail wage shocks occur within-firm. This is true even though firm changes are overall more common during high-growth years. Firm changes constitute 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 Table 3 shows, men are about 26% more likely than women to experience a wage increase of 25+ log points when it is associated with a firm change. By comparison, men are 47% more likely than women to experience a 25+ log point wage increase when it occurs during an individual’s 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 when it is associated with a firm change, compared to 23% more likely for within-firm growth.

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 again proceed with minimal assumptions, starting from the following identity governing average cumulative wage growth in log points by age $a$:

$$\sum_{t=26}^{a} \Delta W_{t}^{k} = \sum_{t=26}^{a} \sum_{j} P_{t}^{k}(j) \Delta w_{t}^{k}(j), \quad k = m, f. \quad (1)$$

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 purposely 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 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

$$g_{j}^{a} = \sum_{t=26}^{a} P_{t}^{m}(j) \Delta w_{t}^{m}(j) - \sum_{t=26}^{a} P_{t}^{f}(j) \Delta w_{t}^{f}(j), \quad (2)$$

where $\sum j g_{j}^{a} = G^{a}$. The outcome of this decomposition indicates that 28% of differences in cumulative wage growth by age 45 are driven by differences in the wage growth associated with firm changes.$^{11}$ To summarize, gender differences in both right-

---

$^{11}$We explicitly avoid a more parametric decomposition, tied to a specific wage setting model in the labor market (see the discussion in Postel-Vinay and Robin (2006)). Card et al. (2016) use a two-way individual and firm fixed effect wage specification to estimate gender wage differentials due to differences in returns to firm changes. Their estimates using this fully parametric model indicate that differential returns to firm changes account for about 20% of gender differences in lifecycle wage growth, which is similar to our estimate.
tail and overall wage growth are driven primarily by differences in within-firm wage growth (fact 3).

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

Based on the evidence above, we lastly investigate whether right-tail, within-firm observations are commonly attributable simply to working at a firm with high average wage growth, or whether they represent years of upward mobility through a firm’s wage hierarchy. For this purpose, we employ the two measures described in Section 3. 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 measure of “promotion,” 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 at different levels of real annual wage growth, 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 given level of wage growth, compared to women. The next two columns summarize the share of total wage growth, in each category, due to average wage growth at the firm (column 2) and relative wage growth (column 3). Column 2 indicates that average firm wage growth is fairly similar across categories, varying from 2.2 to 4.0%. Instead, the majority of the variation in annual wage growth is attributable to relative wage growth (column 3). This indicates that the observed gender differences in the incidence of right-tail, primarily within-firm wage growth correspond to gender differences in the incidence of years of high internal 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 “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.

We conclude with a decomposition similar to the previous one, again relying on equations (1) and (2). This time, instead of conducting the decomposition for two categories – firm changes vs. within-firm growth – we additionally separate within-firm growth into promotion and non-promotion growth, again for \( n = 10 \). Thus, we use three categories, which are exhaustive and mutually exclusive: firm changes, promotions, and non-promotion growth.

Figure 4 graphs the outcome of this decomposition, at each age \( a \). The main finding in Figure 4 is that of the three sources of wage growth, cumulative differences in promotion-related growth are the dominant 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 78% of the differences in wage growth. Differences in growth associated with switching firms account for about 28% of the increase in the wage gap, as quantified previously. Non-promotion growth contributes negatively (-6%), especially after age 39. This decomposition indicates that gender differences in lifecycle wage growth are driven primarily by differences in within-firm mobility (fact 4).

To summarize, this section provided a “bird’s eye” view of men’s and women’s wage growth. We documented that the principal difference between men’s and women’s wage growth distributions is that men have a higher incidence of right tail wage growth. Further analysis shows that this disparity is driven by differences in the probability with which men and women experience years of within-firm growth characterized by high internal mobility, as measured by wage growth relative to other co-workers.

We conclude with a comment. For clarity, we reported results above only for one threshold value used to characterize the binary promotion variable. This focuses the analysis on large internal mobility events. We continue to report results only for this baseline threshold, \( n = 10 \), throughout the remainder of the paper. While the choice of threshold affects quantitative estimates, it does not affect the qualitative results and dynamics that we document, for a large range of values for \( n \). Additional details about this choice of threshold are provided in Appendix A.6. The appendix also documents key results from our analysis for alternative thresholds.

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

Having documented that the gender wage growth gap is driven primarily by differential incidence of internal “promotions;” in this section we focus on accounting for these differences. First, we show how promotion incidence evolves over the lifecycle, and why differences in internal mobility contribute importantly to wage growth differentials. We then analyze the role of human capital, firm characteristics, occupation, and hours worked in driving these patterns. Lastly, we study effects of parenthood.

5.1 Within-Firm Mobility Over the Lifecycle

Table 5 provides summary statistics, by age group, about the wage growth associated with firm switches, promotions, and interim (non-promotion) periods as well as the probabilities with which they occur. The table contains three main findings about lifecycle wage growth, and the contribution of promotions to growth differentials.
First, for both men and women, promotions occur far less frequently than the other two categories, but are also associated with much higher wage growth. For this reason, promotions account for a significant share of lifecycle wage growth, about 40-45% by age 45 (Figure 5). The wage gains following promotion exceed 18 log points (19.7%) at all ages, and are 1-1.5% higher for men. Thus, just one missed promotion corresponds to significant wage losses. Men and women experience low wage growth in interim, non-promotion years, between 1.1-3.3%, depending on age. 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. Importantly, the fact that wage gains exceed 18 log points for both men and women upon promotion indicates that most of the gender difference in promotion-related growth in Figure 4 is driven by difference in the probability of experiencing a large internal mobility event, i.e. by the extensive margin. Equalizing wage gains conditional on promotion only reduces the share of the wage growth gap attributable to differences in promotion to 69.5%.

The 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 17.3% probability of experiencing a promotion in a given year, compared to 21.5% for men, about a 20% 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 return to this finding in our theoretical analysis in Section 6.

The patterns in Table 5 and Figure 6 are in line with the non-parametric results. In Section 4, 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 differ-
ence is the rate with which men and women experience promotions, in line with the differences in right-tail growth documented in Section 4. Missing just one such internal mobility event implies a wage loss of more than 15%, contributing significantly to lower cumulative wage growth over the lifecycle.

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

What accounts for differences in promotion probability, especially early in the lifecycle? While our population of high-skilled men and women is fairly homogeneous, they nevertheless 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. To describe succinctly the role of these factors in accounting for differences in promotion probability we first consider the following regression:

\[ y_{ift} = \mu + \alpha \cdot female_i + X_{it}' \beta + \pi_t + \gamma_f + \epsilon_{ift} \]  \hspace{1cm} (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 \( \pi_t \) and \( \gamma_f \) are year and firm fixed effects. As we add covariates of interest, we primarily describe in this subsection the change in the coefficient \( \alpha \), which corresponds to the gender difference in promotion rates. We focus on individuals under age 35, since these are the ages at which promotions are most common. 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. Finally, in the next subsection we analyze key drivers of the promotion gap dynamically.

In column 1 in Table 6, the baseline gender difference in promotion rates, \( \alpha \), is 4.0 percentage points, controlling only for time fixed effects. Relative to the promotion rate of men under 35, equal to 20.1%, this gap corresponds to a difference of about twenty percent. Next, we add indicators for years of post-graduate education and field of study, and a fourth-order polynomial in years of experience, which increases the estimated gender difference in promotion rates to 4.4 in column 2. 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.\(^{12}\) Therefore, once this difference in choice of major is controlled for, the gap further increases. Next, Column 3 shows that adding a control for years of tenure at the firm has little effect,
since men and women switch firms at similar rates, as we showed earlier. Average tenure in this population is 4.2 years for men and 4.3 years for women.

In Column 4 we add firm fixed effects to the regression, which control for systematic differences across firms in the probability of promotion. Specifically, the addition of this control accounts for gender differences in sorting across firms with different promotion opportunities. Column 4 shows that adding firm fixed effects reduces the coefficient on female by about 10%, from 4.4 p.p. to 4.0 p.p. The small decline in the 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. 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 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. This is in line with our regression results above, that differences in promotion rates of men and women within-firm are most important. 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.

In the next three columns of Table 6, we consider occupational choice. Throughout most of our analysis, we do not add controls for occupation, for the natural reason that many differences in occupation – e.g., manager vs. analyst – are outcomes of promotions. Adding controls for occupation – on top of the controls we already have for field of study – would therefore complicate the interpretation of our results, and potentially control away part of the effect we are seeking to capture. On the other hand, it is possible that women, even when they have the same major as men, are more likely to choose a low-paying occupation not related to their major, possibly one without substantial opportunities for upward movement but, for example, with more flexible hours arrangements. In this case, controls for occupation would be desirable, since they would capture possible differences in chosen career tracks.

In Figure A.3 in the appendix, we document differences in the distribution of men and women across occupations in our population. Women are somewhat more likely than men to work in clerical or administrative support positions, which are typically
associated with fewer opportunities for upward movement. However, the effect of other gender differences in occupation on probability of promotion are not clear-cut. Overall, women are relatively more likely than men to work in occupations related to business and law, in line with their major choices, and less likely to work in occupations related the physical sciences and engineering.

In Table 6, we present results when adding three different types of occupational controls to equation (3). Column 5 adds controls for whether or not an individual works in a clerical or administrative support occupation. The addition of this control has almost no effect, which is not surprising since only a small share of individuals in our population work in such occupations. In column 6 and 7, we add two-digit and three-digit occupation codes, respectively. As expected, this reduces the coefficient on female, but only slightly to 3.9 p.p., which is not statistically different from the baseline coefficient of 4.0 without occupation controls.

In Table 7 we continue the analysis from Table 6, now focusing on gender differences in hours worked. We maintain controls for human capital and firm fixed effects, to concentrate on within-firm gender differences in promotion probability. For this analysis, we use two measures of hours. First, we rely on an administratively recorded variable, contracted hours, which provides detailed information about part-time work and history, indicating hours on a scale from 1% to 100% full-time equivalent (FTE). To complement this variable, which does not capture hours worked above 100% FTE, we 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 labor income. We provide further details about the construction of the hours proxy 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. The average gender difference in weekly hours worked over all age groups is about 5.4 hours, using the proxy measure (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.

Table 7 shows that relative to the baseline (column 1), controlling additionally for hours worked and for part-time history in column 2 reduces the promotion gap to 3.2 p.p., about a 20% reduction in the coefficient. Appendix A.4 details why this reduction is relatively modest. As Table A.5 shows, women are less likely than men to work in categories with the highest weekly hours; thus, controlling for hours worked

---

13For example, an individual contracted at 80% FTE works 32 hours per week.
reduces the coefficient on female, in line with findings by Gicheva (2013). However, the table also shows that women have lower promotion rates for any level of hours worked, even among part-time workers, complicating the simple interpretation that women’s lower promotion rates are entirely a consequence of lower weekly hours. 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 withing category. For all ranges of hours worked, a substantial gender difference in current period promotion rates remains, of about 3 percentage points. This result accounts for the relatively limited decline in the coefficient in column 2 of Table 7.

Finally, Column 3 in Table 7 adds an indicator for whether or not the worker had a child in the current or prior year, and the same indicator interacted with an indicator for being female. We add this control as neither the contracted part-time measure nor the proxy hours measure, which is lagged, fully capture reductions in labor supply associated with parental leave in those years. The addition of this control reduces the coefficient substantially, to about 2.1 percentage points.

To summarize, about half of the within-firm difference in promotion probability for men and women under 35 can be accounted for by gender differences in hours worked and by immediate effects of childbirth events. By contrast, gender differences in field of major, years of experience, occupation and even sorting across firms play a limited role in accounting for differences in promotion probability. We conclude our descriptive analysis by examining these effects of hours worked and childbirth over the lifecycle in the next subsection. We also consider corresponding dynamics in the residual promotion gap.

5.3 Dynamic Effects of Parenthood

The population of high-skill 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, similar to men. 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 most 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 such related labor supply reductions affect gender differences in promotion probability over the lifecycle.

14The reason for this specifically has to do with how individuals who are currently on parental leave are treated in our data. As individuals are still employed when on leave, the contracted hours measure reports their usual hours and does not reflect that their hours are in fact reduced while on parental leave.
A. Total Difference in Promotion Probability, by Time to First Birth

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 promotion probability, i.e. the “promotion gap” 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. (2017)).

We begin by describing the first set of estimates, which measure how the probability of receiving a promotion changes with year relative to first birth \( k \), where \( k = -5, -4, ..., 10 \). The year of first birth corresponds to \( k = 0 \). These estimates are obtained by running the following regression:

\[
y_{it} = \mu + \sum_{k \neq -1} \theta_k D^k_{it} + \sum_k \alpha_k D^k_{it} \cdot female + X_{it}' \beta + \pi_t + \epsilon_{it}. \tag{4}
\]

where \( D^k \) are a set of time-to-birth dummies, equal to one if an individual is \( k \) years from first birth.\(^{15}\) 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, \( \alpha_k \) measures the total male-female difference in promotion probability in every year relative to first birth. The sample for this regression includes only men and women who have ever had children, or about 75% of our original sample. In addition to this baseline specification we also add specifications that control for part-time history and detailed hours worked using the hours proxy variable.

Figure 7 graphs the coefficients \( \alpha_k \). It shows that the time-to-birth dynamics are comprised of three distinct patterns. First, 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 documented in Table 8. These years coincide with parental leave for the majority of women in Sweden.

The second distinct pattern is that 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 two or more years after birth is due to another large 50% reduction.

\(^{15}\)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.
in promotion probability for women at time of second birth, pictured in Figure A.5.\footnote{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.} 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 third key pattern 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.\footnote{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. In particular, individuals are typically compensated for approximately a year of leave 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. We discuss this issue further in Appendix B.5. For institutional details about the parental leave system, see Section 2.}

For this reason, part-time work (the “mommy-track”) only accounts 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. The pre-birth gap is sizable, corresponding to about 29.7\% of the total promotion gap that develops from 5 years prior to first birth to 10 years after first birth. These patterns clarify why detailed controls for hours, for part-time history, and for years of and following births were not enough to account for the promotion gap in the analysis at the end of Section 5.2 (Table 7). As Figure 7 shows, the pre-birth gap is not accounted for by any of these variables.

**B. Motherhood Penalty in Promotion Probability, By Time to First Birth**

The coefficients presented in Figure 7 combine two effects: what the literature calls the motherhood penalty, associated specifically with the effect of childbirth, as well as a gender penalty, i.e. any other factors associated with being female that may affect promotions. To isolate the motherhood penalty, we follow Kleven et al. (2017). We use a specification similar to equation (4), and estimate it separately for men and women:

\[
y_{it} = \mu_{g} + \sum_{k \neq -1} \theta_{k}^{g} D_{it}^{k} + X_{it}' \beta_{g} + \pi_{i}^{g} + \varepsilon_{it},
\]

where \( g \) corresponds to gender. The coefficients \( \theta_{k}^{f} \) and \( \theta_{k}^{m} \) 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 \( \theta_k^f - \theta_k^m \), which is by construction zero in the year before first birth.

To provide an intuition for the difference between these estimates and the ones using specification (4), it is easier to rewrite the equation above, which is estimated separately by gender, instead as one regression in which all variables are interacted fully with female:

\[
y_{it} = \mu + \eta \cdot female_i + \sum_{k \neq -1} \theta_k D_{it} + \sum_{k \neq -1} \alpha_k D_{it} \cdot female + 
\sum_{k \neq -1} \tilde{\alpha}_k D_{it} \cdot female + \tilde{\pi}_t \cdot female + \epsilon_{it}.
\]

(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 promotions, or if employers discriminate against all women of childbearing age, such systematic gender differences would be captured in the estimates of \( \eta, \tilde{\beta}_f \), and \( \tilde{\pi}_f \).

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 \( \alpha_k \) from equation (4) does.

In Figure 8A, we graph the dynamic parenthood effects for men and women separately, i.e. the estimates of \( \theta_k^m \) and \( a_k^f \) from equation (5). As in Figure 7, Figure 8A documents a large drop for women in the probability of promotion in the year of and following childbirth. For men the probability of promotion drops as well, but this decline 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 begin theirs. The effect is smaller for men, as fathers in this cohort on average take less than one-fourth of the total parental leave allocated per child. Several years after first birth, a parenthood effect persists for both women and men, at about three percentage points for mothers, and one percentage point for fathers.

Next, in Figure 8B we plot the motherhood penalty, which is the difference between the motherhood and fatherhood effects graphed in Panel A, and corresponds to the coefficients \( \tilde{\alpha}_k \) from equation (6). Alongside this penalty, we also graph the total male-female promotion gap, from Figure 7. The motherhood penalty accounts for most of the total promotion gap after first birth, with a similarly dramatic drop at first birth. However, 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 8B.\textsuperscript{18} This is closely in line with 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 a little over half of the gender promotion gap.

C. Dynamic Gender Penalty, by Time to First Birth

As the dynamic motherhood penalty accounts for different shares of the total promotion gap in different years, our results imply that residual gender differences in promotion 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 promotions by time to first birth, net of the motherhood penalty and after controlling for observables. The figure effectively illustrates the lifecycle incidence of the unaccounted-for promotion gap in Table 7. 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.

An inspection of individuals who never have children shows a similar dynamic pattern. Panel B plots the set of coefficients when female is interacted with age in our standard regression, with all controls for human capital, firm fixed effects, and hours 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, for both women who ever children and those who remain childless, sizable dynamic gender penalties are observed at younger ages, but reverse around the end of women’s childbearing years.

We conclude with a summary of the key empirical facts about internal mobility. First, a sizable promotion gap between men and women emerges early in the lifecycle, when promotions are most common. Second, it is not readily accounted for by observable gender differences in human capital, tenure, or occupational choice. Third, gender differences in sorting across firms account only for 10% of the promotion gap.

What generates the within-firm promotion gap between observationally similar men and women? Our estimates point to the following decomposition. “Missed” promotions in the year of and year immediately following birth events, when women are typically on parental leave, are the biggest factor, accounting for 40% of the cumula-\textsuperscript{18} 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.
tive promotion gap by age 45, for women who ever have children. There is no evidence of an uptick in women’s promotion rates shortly after first or second birth, suggesting that these promotions are not readily recovered. By comparison, part-time work (the “mommy-track”) accounts for only 21% of the cumulative gap by 45. The gender penalty that women incur prior to first birth accounts for about 30% of the promotion gap by 45. It contributes more than part-time work to observed promotion differences, in large part because it is incurred early on, when promotion rates are highest. This penalty, observed also for women who remain childless, reverses at the end of women’s childbearing years, around age 40. Notably, this reversal mirrors also our earlier findings that the gender gap in wage growth in non-promotion periods and conditional on firm changes is initially higher for women, and reverses after 40. We bring these findings together in the theoretical framework we develop in the next section.

6 A Model of Gender Differences in Careers Within Firms

In this section, we develop a simple theoretical model to interpret the descriptive facts we establish about wage growth and internal mobility of men and women. Because we provide extensive evidence that differences in within-firm wage growth are the main driver of gender differences in wage trajectories, we build a model focused specifically on men’s and women’s careers within firms.

Several classic models of careers can potentially generate a key fact documented in Section 4: that men’s and women’s wage dynamics are characterized by frequent low annual growth, with large periodic wage increases which occur primarily within-firm. 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). We build on one of these workhorse models, by Gibbons and Waldman (1999b), which generates many insights with minimal math and rich intuition. To analyze gender differences in wage growth and internal mobility, we extend their model of promotion dynamics by incorporating childbirth and associated labor supply reductions, as well as the firm’s expectations about these events. We also derive a unique testable implication of our model, and show that it is not empirically rejected.

Before beginning our exposition, we note that one drawback of the model by Gibbons and Waldman is that the classic version of the model generates promotions, but not necessarily exceptionally large wage growth at time of promotion. This can be rectified through one of two possible extensions: either by incorporating private information by the employer about the worker, as is developed in the same paper by Gibbons and Waldman; or by incorporating compensation for effort associated with
different jobs, as we do. However, the addition of either feature complicates the model we develop below, without adding insight or providing further intuition for our findings. For this reason, we focus on the simplest version of the model in the main exposition. We provide proofs in Appendix B for the augmented model, which generates large wage increases associated with promotions, in line with the data. All of the results that we document in this section hold identically for the augmented model.

6.1 General Environment and Benchmark Model

We begin by describing the general environment, which is identical to the full information setting in Gibbons and Waldman (1999b). Specifically, firms are homogeneous and there is free entry into production, with labor as the only input. Workers and firms are risk-neutral and have a time discount rate of zero. A measure one of workers, indexed by $i$, can change firms costlessly. These workers are characterized by innate ability $\theta_i \in [0,1]$, which has a uniform distribution, and which does not change over the $T+1$ periods that they work. However, workers accumulate labor market experience $x_{it}$. All firms observe innate ability and experience, which together generate the worker’s effective ability $\eta_{it}$ in any given period $t$:

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

with $f'(\cdot) > 0$, and $f(0) = 0$. Thus, for any given worker, effective ability $\eta_{it}$ grows over time, even though innate ability $\theta_i$ is constant. Firms consist of $J$ jobs. The output of a worker assigned to job $j$ is linear $\eta_{it}$ and equals

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

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

Depending on their effective ability, workers will be more productive at some jobs than others. A worker with no experience (who therefore has $\eta_{it}=0$) is most productive in job 1, since $d_1 > d_2 > ... > d_J$. 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 1 and 2 solves $d_1 + c_1 \eta_{it} = d_2 + c_2 \eta_{it}$. We denote this solution as $\eta^2$, since it defines the threshold value above which a worker starts to be more productive at job 2. Similarly, $\eta^3$ solves $d_2 + c_2 \eta_{it} = d_3 + c_3 \eta_{it}$, and so on, for the remaining jobs. It is assumed that parameter values for $c_j$ and $d_j$ are such that $\eta^2 < \eta^3 < ... < \eta^J$. This inequality implies that increasingly higher level jobs also require increasingly higher effective ability. Efficient assignment implies that workers are assigned to job 1 if $\eta_{it} < \eta^2$, to job 2 if $\eta^2 < \eta_{it} < \eta^3$, and so on.
In period 0, workers enter the labor market. Promotions are possible in period 1 and in all subsequent periods. Firms make a decision at the start of the period about whether or not to promote a worker. Workers have no experience in period 0. Therefore, they are all hired optimally into job 1. As their experience increases, so does their effective ability. Since $\eta_{it} = \theta_i f(x_{it})$, effective ability increases most rapidly over time for the individuals with the highest innate ability, $\theta_i$. Therefore, these individuals will also be the first to cross the threshold values described above as they accumulate experience, while lower innate ability individuals will be promoted more slowly (see Figure 11A). More generally, for any given (strictly positive) years of experience, the threshold values for promotion can be restated in terms of innate rather effective ability. Let $\tilde{\theta}_j^\tau$ refer to the threshold value of innate ability required to be promoted to job $j$, for someone with $\tau \geq 1$ years of experience, where $\tilde{\theta}_j^\tau = \frac{\eta_j}{f(\tau)}$. Two characteristics define these innate ability thresholds. First, the innate ability required for promotion to a given job $j$ is higher for someone with few years of experience, i.e. $\tilde{\theta}_1^\tau > \tilde{\theta}_2^\tau > ... > \tilde{\theta}_T^\tau$. Second, holding years of experience $x_{it} = \tau$ fixed, the innate ability threshold is higher for higher-level jobs: $\tilde{\theta}_T^\tau > ... > \tilde{\theta}_2^\tau > \tilde{\theta}_1^\tau$ (see Figure 11B).

In this setting, firms make zero profit but optimize the efficiency of output. As Gibbons and Waldman show, it follows immediately that 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 additionally 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. In this benchmark setting, it is never optimal to demote a worker, since effective ability increases monotonically as a worker gains experience.\textsuperscript{19}

Finally, we follow Gibbons and Waldman in considering a restricted subset of parameters for job technologies and human capital accumulation, $c_j$, $d_j$ and $f(x)$, to reduce the number of cases that need to be considered and to focus the analysis to the most insightful ones. In particular, we assume that the parameterizations satisfy the following conditions. First, the highest ability men and women are promoted exactly once in each period, if they work every period. Second, the lowest ability men and women are never promoted.

\textsuperscript{19}Gibbons and Waldman prove that this model also generates a number of desirable results that describe promotion and wage dynamics, such as serial correlation in promotions. As we will not revisit their additional results below, but instead focus on a new set of findings, we refer the interested reader to their original study.
6.2 Extended Model: Gender Differences in Promotion Dynamics

We now extend this benchmark model to study the observed differences in promotion dynamics for men and women, who we assume have identical distributions of innate ability $\theta_i$ and identical human capital accumulation functions $f(x)$. We begin by describing its main features.

A. Lifecycle Structure and Employer Costs

We introduce two main departures from the benchmark environment. As a first departure, we introduce a worker lifecycle with three distinct fertility phases. This feature allows us analyze time-to-birth dynamics, which we showed earlier to be empirically important. The first phase is a pre-birth phase in which most individuals do not yet have children. For simplicity, we assume that the probability of birth in this first phase is zero, though this assumption can be readily relaxed, as we later show. The second phase is the prime childbearing phase. During this phase, individuals have a strictly positive per-period probability of birth. Finally, the third phase characterizes the end of individuals' fertile lifespan, and births in this phase occur with a probability of zero.

In the exposition below, we work with the minimum number of periods, $T = 3$, so that each phase lasts for one period, as summarized in Figure 12. Including additional periods in each of the three phases complicates the model without changing any of its main insights. In period 0, i.e. when individuals first enter the labor market and before promotions 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 that period, with $p_f > p_m$. We normalize $p_m = 0$ without loss of generality.\(^{20}\) The assumption that women reduce their labor supply to zero after 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. In all periods other than $t = 2$, both men and women supply a unit of labor; additionally, the probability of having a child and taking a leave is uncorrelated with ability, two simplifications that we later relax. The timing in each period is described graphically in Figure 12. At the start of the period, firms observe whether a worker gives birth and takes 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.

\(^{20}\)One can think of $p_f$ and $p_m$ a composite probability: the probability of having a child times the probability of taking time off, conditional on having a child, where we normalize the latter to zero for men, and to one for women.
Our second key departure from the benchmark environment is to introduce a cost for the firm when an employee is on leave. It takes the following form:

**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 $k_{j+1} > k_j > 0$.

To describe the intuition behind this assumption, we note that the production technology in the benchmark model abstracts from any spillovers or other complementarities between workers, knowledge hierarchies, slot constraints, or hiring or training costs for temporary replacement workers. Assumption 1 is therefore introduced to capture two commonly held ideas: (1) that employers incur some costs when an employee assigned to a job is on leave, even when employers do not pay for the leave, and (2) that an employee’s absence may be more costly when, for example, the employee is a manager who affects the productivity of many workers, vs. a rank-and-file worker. In line with Swedish labor laws, we assume that firms cannot demote or fire workers based on their current or anticipated childbearing or leave-taking decisions, and cannot write long-term contracts contingent on labor supply behavior after childbirth.

**B. Main Result**

In this setting, we obtain the first set of key results, summarized in Proposition 1:

**Proposition 1** Promotion dynamics for men and women are characterized by the following features:

i. In period 1, women – both future mothers and those who remain childless – are promoted less frequently than men.

ii. In period 2, women who give birth have significantly lower promotion rates than men or childless women.

iii. After all childbearing decisions have been revealed, women who never have children have higher promotion rates than men (a positive “gender effect”); women with children experience both a “motherhood penalty,” and a positive “gender effect” in promotion rates in period 3.

**Proof.** The full proof is provided in Appendix B.2. We present a detailed intuition below, in two steps. First, we discuss the values each period for effective ability, $\eta_j$, required for assignment to job $j$, by gender. Then, we consider the corresponding shares of men and women promoted.

Since men supply the same labor as in the benchmark model, the problem for men is straightforward and identical to the benchmark environment. In all periods, firms
simply apply the standard cut-off values for effective ability to work in job $j$:

$$
\eta_j^l = \frac{d_{j-1} - d_j}{c_j - c_{j-1}}. 
$$

(7)

For women, the problem is more complicated, as women’s labor supply varies over the lifecycle. Since women cannot be fired for taking leave, employers take into account the probability of incurring cost $k_j$ in a current or future period. Therefore, the cut-off values for effective ability for women vary from period to period. The third and final period is the simplest: there is no further possibility of childbearing and leave-taking, and therefore no possibility that a firm will incur a higher cost $k_j$ after promoting a woman. In period 3, the firm’s decision process for women is therefore identical to that for men, with cut-offs for effective ability again described by equation (7).

In period 2, decisions about childbirth are revealed at the start of the period. Share $p_f$ of women have a child and reduce their labor supply to zero, producing $-k_j$ in output. These mothers on leave are promoted with probability zero (their cut-off values are undefined), since employers would only incur a higher cost $k_{j+1} > k_j$ that period upon promoting them. Women who remain childless that period, on the other hand, work and have well-defined cut-off values for promotion. Since all uncertainty about current and future childbearing has now been resolved for childless women, they are promoted in period 2 based on cut-offs that are again identical to those for men and described by equation (7).

Finally in period 1, employers anticipate that female workers may go on a costly leave in period 2. Recall that only promotions from job 1 to job 2 are possible for men and women this period. When deciding whether to promote a female employee to job 2, the firm considers both the additional output she will generate in job 2 in the current period, as well as the higher cost $k_2 > k_1$ that the firm will incur with probability $p_f$ in period 2, since we assume that firms cannot fire or demote workers based on their childbearing decisions. As we show in Appendix B.2, the firm maximizes expected profits by choosing a cut-off value in period 1 for women corresponding to

$$
\eta^* = \frac{d_1 - d_2}{c_2 - c_1} + p_f \frac{k_2 - k_1}{c_2 - c_1} = \eta^2 + p_f \frac{k_2 - k_1}{c_2 - c_1} > \eta^2, 
$$

(8)

where the final inequality follows from the fact that $p_f > 0$, $k_2 > k_1$ and $c_2 > c_1$. Thus, in period 1 the ability threshold for promotion applied to women, $\eta^*$, is higher than that for men, $\eta^2$.

To summarize, effective ability cut-offs for promotions are strictly higher for women than for men in period 1. In period 2, women who give birth are promoted with probability zero (their effective ability cut-offs are undefined). However, women who remain childless have identical $\eta^l_j$ cut-offs as men. In period 3, cut-offs $\eta^l_j$ for all men
and women are identical. We can now derive the share of men and women promoted each period, and describe the intuition for each part of Proposition 1.

Proposition 1(i): In period 1, all men and women enter the period with $\tau = 1$ year of experience, and only one type of promotion is possible (from job 1 to job 2). For men, the threshold value for innate ability to be promoted to job 2 with one year of experience is $\theta^2_1 = \frac{\eta^2}{f(1)}$. Thus, share $(1 - \theta^2_1)$ of men are promoted. For women, $\eta^* > \eta^2$ determines a correspondingly higher threshold value of innate ability required for promotion, which we denote as $\theta^* = \frac{\eta^*}{f(1)} > \frac{\eta^2}{f(1)} = \theta^2_1$. Thus, the share of women who are promoted, $(1 - \theta^*_1)$, is strictly lower than the share of men who are promoted, $(1 - \theta^2_1)$. The difference $\theta^2_1 - \theta^*_1$ is the gender penalty in promotion rates in period 1. Since employers do not know which women will eventually have children, this penalty is incurred by all women, regardless of future childbearing outcomes.

Proposition 1(ii): In period 2, individuals enter the period with two years of experience, and promotions to jobs 2 and 3 are possible. Employers promote share $(1 - \theta^3_2)$ of men to job 3. Additionally, they promote share $(\theta^2_1 - \theta^3_2)$ of men to job 2 for the first time, since $\theta^3_2 = \frac{\eta^3}{f(2)}$ determines the new ability threshold for promotion to job 2 for men with $\tau = 2$ years of experience. By contrast, share $p_f$ of women have a child and go on leave, and are promoted with probability zero, as discussed above. Thus, the difference in promotion rates between mothers and fathers (as well as mothers and childless men) in period 2 is $(1 - \theta^3_2) + (\theta^2_1 - \theta^3_2)$. Additionally, mothers also have lower promotion rates than childless women period 2, since childless women work and have strictly positive probabilities of being promoted.

Proposition 1(iii): We begin by considering the first part of Proposition 1(iii): namely, that childless women’s promotion rates are higher than men’s after all childbearing decisions have been revealed, i.e. after the start of period 2. Recall that effective ability thresholds $\eta^j$ for men and for childless women are identical in period 2. Since both enter with two years of experience, their innate ability thresholds for promotion are also identical. This implies that an identical share $(1 - \theta^3_2)$ of men and childless women are promoted to job 3. However, the shares of men and women promoted for the first time to job 2 differ, since these shares depend also on promotions to job 2 in the prior period, which were determined by the cut-off $\theta^2_1$ for men, and $\theta^*_1$ for women (see discussion of Proposition 1(i)). Thus, share $(\theta^3_1 - \theta^3_2)$ of men are promoted to job 2 for the first time this period, while the share for childless women is $(\theta^*_1 - \theta^3_2)$. In particular, $\theta^3_1 - \theta^3_2 = (\theta^3_2 - \theta^*_2) + (\theta^*_1 - \theta^3_1)$, where the first term on the right-hand side corresponds to the share of men promoted to job 2, and the second term corresponds to the (strictly positive) share of women who were previously “passed up” for promotion.
in period 1, and who are now promoted.\footnote{This assumes that $\bar{y}_1 < \bar{y}_2$. For the case $\bar{y}_1 > \bar{y}_2$, the additional share of childless women promoted is $\bar{y}_2 - \bar{y}_1$ which is also strictly positive, under the parameter restrictions previously discussed.} Thus, women who remain childless have higher promotion rates than men and experience a positive “gender effect” in this period.

Finally, we consider the last part of Proposition 1(iii): that women who had children experience both a “motherhood penalty” in promotion rates in period 3, as well as a positive “gender effect.” Both fathers and childless men enter the period with $\tau = 3$ years of experience, and share $(1 - \bar{y}_3) + (\bar{y}_3 - \bar{y}_3) + (\bar{y}_2 - \bar{y}_3)$ of men are promoted to jobs 4, 3 and 2 for the first time. By contrast, women who had children enter with only $\tau = 2$ years of experience, since they did not work in period 2. The share of mothers promoted is equal to $(1 - \bar{y}_2) + (\bar{y}_1 - \bar{y}_2)$, or equivalently, $(1 - \bar{y}_2) + (\bar{y}_1 - \bar{y}_2) + (\bar{y}_1 - \bar{y}_1)$. Whether men’s or women’s total promotion rate is larger in absolute value in period 3 will depend on the specific parameterizations adopted. Nevertheless, it is clear that there are two separate effects acting on women’s promotion rates in this period. The first is the same gender effect $(\bar{y}_1 - \bar{y}_1)$ that is observed for childless women in period 2, generated by the fact that women who were initially “passed up” for promotion in period 1 now advance to a higher position, as uncertainty about childbearing has been resolved.\footnote{This supposes that $\bar{y}_1 \leq \bar{y}_2$. If $\bar{y}_1 > \bar{y}_2$, the additional share promoted is $\bar{y}_2 - \bar{y}_1$. See previous footnote.} The second effect is a persistent motherhood penalty affecting women’s probability of promotion to different jobs, stemming from their foregone experience. In particular, no women with children are promoted to job 4 this period. Figure 13A illustrates the motherhood and gender penalties generated by the model in each period, for one sample parameterization.

To summarize, the model generates the key empirically observed promotion dynamics. As in the data, motherhood penalties are large and negative in the year of birth, and persist in the post-childbearing period. Additionally, gender effects are initially negative, and then reverse after the probability of women’s future childbearing declines to zero. As we will discuss below, Assumption 1 is critical to these results. Our next proposition relates to this assumption and central mechanism of the model.

C. A Testable Implication

We derive a direct and unique testable implication of the model, which constitutes our second key result. It is summarized in Proposition 2.

**Proposition 2** *The larger the difference $k_2 - k_1$, the larger the gender difference in promotion probabilities in period 1.*

**Proof.** See Appendix B.3. The proposition says that the more costly it is to have a
worker on leave in a higher ranked (e.g. managerial job) relative to the entry level job, the more negative the gender penalty \((\overline{\theta}_2^* - \overline{\theta}_1^*)\) in period 1. This follows directly from equation (8), since \(\overline{\eta}^*\), and therefore \(\overline{\theta}_1^*\), are increasing in \(k_2 - k_1\).

Testing this implication empirically requires direct measures of costs \(k_j\) to employers, which we do not have. However, there are ways to proxy for these costs, to allow us to test the prediction above, which we do in two ways. Our first approach is to look at establishment size. The idea behind this proxy is that an establishment of 1000 employees, with dozens of managers and formalized, on-site 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. 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 very few men in a given workplace use their allotted "daddy months," this may indicate that the firm finds it particularly costly for workers to take time off, and penalizes them. Thus, our second proxy for how costly leaves are to firms is 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.\(^{23}\) At firms where few men take the full leave, we expect a larger pre-birth promotion gap, i.e. a larger gap between men and women who are currently childless but of childbearing age.

Figure 10 graphs the gender promotion gap at establishments of different sizes, for employees who are under 35 and without children. This includes all individuals who have not yet had a first birth and individuals who never have children, to approximate the population of individuals who are working during period 1 of our model. For this population, the figure graphs the coefficients on the interaction between establishment size quintile and female, controlling for firm size, age and time fixed effects, and detailed human capital characteristics. In line with our prediction, Figure 10 documents that the promotion gap is higher at smaller establishments, where one would expect leaves by employees to be costlier to the firm. The pre-birth gender promotion gap decreases monotonically with firm size. At the smallest establishments (under 32

\[^{23}\text{An implicit assumption is that the cost of an entry-level employee on leave varies less across establishments or firms than the cost associated with a manager on leave, the idea being that there are many workers that can potentially fill a lower-level position or be trained for it. It is possible that our second proxy, which uses information about use of daddy months, could capture higher costs not just for } k_2, \text{ but also } k_1. \text{ However, men in Sweden tend to have their first birth in their mid-thirties, when many promotions have already occurred. Therefore, the daddy month proxy should generally capture behaviors of older men who are already in some promoted position.}\]
employees), the pre-birth promotion gap is 4.0 percentage points, while at the largest establishments (more than 717 employees), the pre-birth gap is approximately zero. As men’s promotion rates are very similar across establishment sizes, the pre-birth gender promotion gap is approximately 18% in the smallest establishments, decreases steadily with establishment size, and is approximately zero at the largest establishments.

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. In our population, approximately 25% of individuals work at such a “low uptake” firm. Table 9 shows the regression coefficient on the interaction between female and the binary variable for low daddy month uptake, with promotion again as the outcome variable. At firms where men tend not to take their allotted daddy months within two years after a birth, the promotion gap between childless men and women under 35 was greater, as predicted, by about 1.4 percentage points. Specifically, at firms where daddy month uptake is high, the gender promotion gap for individuals who either are prior to their first birth or remain childless is 2.7 percentage points (i.e., the coefficient on female), or approximately 13.5%. At firms where uptake is low, the gap is 4.1 percentage points. As the overall promotion rates at low uptake firms are slightly higher, this corresponds to a gap of about 18.6%. Thus, the findings using both proxy variables are consistent with Proposition 2: gender penalties for childless women who are of childbearing age are higher at firms where leaves are likely to be costlier to employers.

Proposition 2 relates to a crucial assumption in the model. Without costs $k_j$, which are increasing in job rank, employers would have no reason to promote fewer women in period 1, or even in period 2, when births occur: while output of a promoted worker on leave in period 2 would be zero, the cost to the employer of would be zero as well, since the employer does not pay a wage to the worker. Importantly, foregone output alone is not enough to rationalize the gender or even the motherhood penalty. To generate the key empirical results, a specific cost to the employer is needed – either as we have modeled it, or potentially as a cost of upfront investment in the worker prior to promotion. Unfortunately, little is known about employer costs related to worker leave-taking to date, as other researchers have noted (e.g. Rossin-Slater (2017)), constituting an important area for future research.

24We construct a binary variable, rather than quintiles, as there is relatively little variance in this variable across some quintiles. For example, for approximately half of all observations, the uptake rate of daddy months is between 60% and 70%.
D. Additional Wage Dynamics

We have shown above that the model can generate the key empirical facts around promotion dynamics documented in the paper. Additionally, the model is not rejected based on the testable implication we develop. Lastly, we discuss additional wage dynamics generated by the model. Recall that Section 5.1 described wage growth conditional on promotion, firm changes, or staying in the same job without a promotion. Early in the lifecycle, women experience lower wage growth than men during non-promotions and firm changes; however, this gender difference reverses after age 40 (Table 5). In the model, individuals’ wages grow also in periods when they are not promoted, as effective ability $\eta_i$ increases with experience. Therefore, we can use our framework to help interpret and rationalize this ex ante non-obvious reversal.

To analyze gender differences in wage growth in periods when individuals are not promoted, we derive in Appendix B.4 the wage functions for men and women in each period.

25 Men’s wage function is straightforward and identical to the benchmark model:

$$w_{ij0}^m = d_j + c_j \eta_i$$

Women’s wage function varies by period and takes the following form:

- **Period 0:**
  $$w_{ij0}^f = d_1 + c_1 \eta_i$$

- **Period 1:**
  $$w_{ij1}^f = d_j + c_j \eta_i - p_f k_j$$

- **Period 2:**
  $$w_{ij2}^f = d_j + c_j \eta_2$$

  if childless, on govt.-paid leave otherwise

- **Period 3:**
  $$w_{ij3}^f = d_j + c_j \eta_3$$

Thus, all men and women are all initially hired into job 1 at 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 based on leave in period 2. If firms did not pass on this cost, they would expect to make negative profits. Among those workers who stay in job 1 from period 0 to period 1 (as well as those promoted), wage growth for women is therefore strictly lower than for men of the same ability.

Once women’s childbearing years end, employers no longer anticipate incurring cost $k_j$. Correspondingly, a female worker’s value to the firm rises. In period 2 for childless women, and in period 3 for women with children, wages are therefore bid back up in the market. In those respective periods, women’s wages grow not only with the increase in their effective ability, but also by an additional $p_f k_{j-1}$, where $j-1$ refers to the job held in the prior period. Depending on the parameterization of the technologies for different

---

25 For the wage results presented in this subsection, we require an additional assumption that there is a probability $\epsilon > 0$ of changing employer from period 0 to period 1. As job switching rates are very high at young ages, we do not view this assumption as restrictive. The assumption resolves an indeterminacy issue, as discussed in Appendix B.4. Alternatively, one can impose the stronger assumption that there is a probability $\epsilon > 0$ of changing employer in every period, with the same result.
jobs, human capital and costs \( k_j \), 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. Thus, the model generates a gap in non-promotion wage growth favoring men early on in the lifecycle, and embeds a clear mechanism that can produce a corresponding gap favoring women later in the lifecycle.

The model’s predictions for wage growth conditional on switching firms are similar. This is illustrated in Figure 13B. Under the assumption of an exogenous separation rate \( \epsilon > 0 \) at the end of each period, leading to a firm change, the model predicts strictly lower wage growth for women during firm changes in period 1, as we prove in Appendix B.4, and a possible reversal of this wage growth gap in period 3. This result, in line with the data, is driven by the same mechanism: women’s market value is lower than men’s in period 1, but rises again once there is no more uncertainty about future childbearing.

E. Relaxing Assumptions and Extensions

Throughout this section we made several simplifications to keep the analysis tractable. These included the following assumptions: that individuals have a child in period 1 with zero probability; that ability and probability of childbirth are uncorrelated; and that all men and women supply a unit of labor after the childbearing period, including women who had children. In Appendix B.5, we show that relaxing these assumptions does not affect any of the results in Propositions 1 and 2.

Our model generates one notable counterfactual prediction: after the end of the childbearing period, women who never have children experience not only higher promotion rates, but in fact converge completely with men in wages and job assignments. A likely reason for this counterfactual prediction 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 (e.g., Lise (2016)). For this reason, early lifecycle promotion differences do not have sufficiently persistent effects in our model. We leave such important extensions for future research.\(^{26}\)

6.3 Alternative Explanations

Our model can generate the key empirical facts of the paper about promotion, time-to-birth dynamics, and conditional wage growth; additionally, we develop a test of the model and do not reject it empirically. However, it is possible that other explanations can also account for the empirical patterns we document. We therefore conclude with

\(^{26}\)Gibbons and Waldman (2006) develop an extension along these lines to study cohort effects.
a discussion of plausible alternative explanations. As motherhood penalties are more straightforward to rationalize, we focus on whether alternative explanations can fit remaining features of the data, in particular the change in gender penalties before vs. after women’s childbearing years.

We begin with theories about systematic gender differences in productivity (e.g. Azmat et al. (2018)), propensity to negotiate (Babcock et al. (2006)) or propensity to behave competitively (e.g., Niederle and Vesterlund (2007)). Each of these three theories can generate gender gaps in promotions favoring men. However, as a primary explanation, their drawback is that they can only rationalize early lifecycle differences, but not those after age 40. If lower propensity to compete or negotiate, for example, is a systematic characteristic of women, one should observe corresponding negative gender effects over the entire lifecycle. A possible caveat is that the selection of working women could change with age – for example, 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 largely ruled out by our setting, as 95% of women in our high skill population are in the labor force at all ages.

A second, related explanation 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 patterns 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.

A third explanation is that young women, anticipating future labor supply reductions, have less incentive to accumulate human capital and to work high hours, 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. This explanation is promising, but has the following drawbacks. First, it is not clear that women’s anticipated leave-taking behavior decreases their incentive to accumulate human capital or work high hours prior to birth. Paid leave in Sweden compensates workers according to their prior earnings, at an 80% replacement rate. Thus, for an individual that plans to take leave the next year, each additional hour worked generates 1.8 times the wage earnings, an order of magnitude more than the typical estimated returns to human capital (e.g., Imai and Keane (2004), Keane (2011)).  

We discuss this issue at greater length in Appendix B.5. A second limitation is that the explanation

---

27 For the same reason, women have significant incentive to accumulate as much human capital as possible even earlier on in the lifecycle, to maximize also their hourly wage prior to taking leave.
cannot readily account for gender promotion penalties for those who never have children. Finally, men and women work similar hours prior to first birth, and the pre-birth promotion gaps we document are present even after flexible controls for hours worked. We cannot rule out that we miss some hours worked above the full-time threshold using our measures; however, women experience lower promotion rates than men at all levels of hours worked, even among women and men who work part-time (see Section 5.2). We leave a more thorough investigation of these alternative theories to future research.

7 Limitations and Sensitivity

7.1 Limitations and Discussion

Our wage-based measure of promotion has a number of advantages, as discussed in Section 3, but also the limitation that promotions cannot be mapped to moves along a clearly defined career ladder, as in studies focused on a single firm (e.g., Baker, Gibbs and Holmstrom (1994a,b)). One potential concern using our measure is that the large within-firm wage increases we observe reflect matches of outside job offers (e.g., Burdett and Mortensen (1998), Postel-Vinay and Robin (2002)), rather than changes in job assignment. While this would not affect our empirical findings, it does affect their interpretation. Three patterns in the data suggest that the promotions we identify are most likely associated with the more traditional definition, related to job assignment. First, as discussed in Section 3, individuals who experience a promotion under our definition are also likely to experience a change in occupational code. Second, a gender promotion gap of about 3.7 percentage points is observed also at public sector establishments, which have regulated wage schedules and limited flexibility to match wage offers. Third, we find that most promotions occur early on in the lifecycle, whereas the literature estimates that external offers affect wage growth primarily later on in the lifecycle (e.g., Bagger et al. (2014)).

Second, as noted in Section 3, our measure requires us to specify a threshold. In our analysis, we use the same threshold for men and women. This is, we believe, the clearest and most transparent approach. However, it could be that women are paid substantially less for identical moves up a career ladder (e.g., as measured by change in job complexity or duties), 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 somewhat gender differences in the probability of experiencing large internal mobility events, and understate differences in wage growth conditional on such events. Combining our wage-based measure with data on a subset of firms with detailed information on organizational structure can shed additional light on this issue.

Finally, we note two caveats to our findings. First, the variable of interest in our
study is wage, not annual labor earnings. Existing studies find that moves up the career ladder not only increase base pay but also bonus compensation (e.g., Ekinci, Kauhanen, and Waldman (2015)). This may further amplify gender differences in total earnings trajectories, as well as the contribution of promotions to these differences. Second, the findings we document are for a specific institutional and policy environment. If costs to employers associated with lengthy parental leaves differ significantly from those associated with high quit rates after birth, observed gender and motherhood penalties may also differ across countries with different institutional environments, such as Sweden and the U.S. We leave an investigation of this important question for future work.

7.2 Sensitivity

Throughout the paper, we set the threshold \( n \) for the relative wage growth that defines a promotion equal to 10 log points, focusing the analysis on relatively “large” movements within the firm. In Appendix A.6, we analyze how our results change with alternative choices for \( n \). Additionally, we consider a promotion measure based on median (instead of mean) wage growth of college-educated co-workers at the firm. We present the results in Table A.6. By construction, as the threshold for \( n \) increases, the share promoted declines. 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 promotion rate to just 0.11 in the years when promotions are most common, ages 26 to 35. Nevertheless, the share of lifecycle gender wage differentials accounted for is still quite high, at 64%. The results are very similar to baseline results when one constructs the promotion measure based on median wage growth of co-workers at the firm, rather than average wage growth. 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 characterized both by lower promotion rates than in the baseline population, as well as by substantially lower growth in non-promotion years. As a result, the share of gender differences in lifecycle wage growth accounted for by promotions in the full population declines, to 62%. Instead, male-female differences in non-promotion growth now play a larger role, 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.
8 Conclusion

Why do women's wages grow more slowly than men's? In this paper, we provide detailed descriptive evidence on this question, focusing on high-skilled men and women, and develop a theoretical framework to help interpret this evidence. We first show, nonparametrically, that men's and women's wage growth distributions are similar in most respects; quantitatively, the most important difference is that women experience right-tail, within-firm wage growth with substantially lower frequency than men. These events are characterized by high growth relative to other workers at the same firm, resembling large internal “promotions;” their differential incidence by gender accounts for around 70% of the total gap in male-female wage growth.

Next, we document in detail promotion dynamics over the lifecycle, using a novel measure of promotion that exploits the employer-employee matched nature of our data. An analysis of gender differences in human capital, occupation, and sorting across firms with different promotion opportunities indicates that most of the promotion gap is a difference between observationally similar men and women at the same firm. Its biggest driver is a substantial penalty incurred by mothers specifically in the year of and immediately following childbirth events, when women are typically on leave, accounting for 40% of the cumulative promotion gap by age 45. Part-time work by mothers accounts for a much smaller share by comparison, 21%, primarily because part-time work is more common later in the lifecycle, when promotions occur less frequently. Finally, a dynamic gender penalty, not accounted for by motherhood, also plays a quantitatively significant role. For women who ever have children, the gender penalty in promotions incurred prior to birth corresponds on average to 30% of the cumulative gap by 45. This gender penalty reverses around age 40. To interpret these findings, we develop a model of gender differences in promotion dynamics based on Gibbons and Waldman (1999). The facts we document are consistent with costs to firms associated with employee leave-taking, and employer uncertainty about women's future childbearing and labor supply reductions. By contrast, the empirical facts are not readily rationalized by behavioral explanations, such as gender differences in competitiveness or propensity to negotiate.

Our findings have implications for future research and policy. First, they indicate that within-firm wage and promotion dynamics are the primary driver of gender wage growth differentials, at least within our population of college-educated individuals. By contrast, sorting to lower-paying and lower-growth firms and differential returns to firm-to-firm changes play a decidedly secondary role in driving gender differences in wage trajectories, a finding also in line with Card et al. (2016). Calls to incorporate more theoretical insights from personnel economics and the literature on careers within firms into analyses of gender differences in wage trajectories (e.g. Goldin (2014)) are
strongly supported by our results. Second, our findings indicate that a large share of the lifetime promotion gap is incurred over a strikingly short period, corresponding to the years when Swedish women are typically on parental leave. Additionally, the dynamic gender penalties we document, their larger magnitude at smaller establishments and at firms where few men take leave, and the reversal of gender penalties after the end of women’s childbearing years, are all consistent with employer costs associated with such leaves. Obtaining data on employer costs related to work interruptions and leave-taking is an important area for future research, with significant implications for optimal policy design.
References


### Table 1: Summary Statistics, 1960-1970 Cohort

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

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

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

### Table 3: Wage Growth During Firm Changes vs. Within-Firm Wage Increase

<table>
<thead>
<tr>
<th>Wage Growth</th>
<th>% Wage Increases Occurring During Firm Changes</th>
<th>Likelihood of Experiencing Wage Growth Shock: Men to Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>During Firm Change</td>
</tr>
<tr>
<td>&lt;0</td>
<td>12.1%</td>
<td>0.99</td>
</tr>
<tr>
<td>0 - 0.05</td>
<td>8.1%</td>
<td>0.94</td>
</tr>
<tr>
<td>0.05 - 0.10</td>
<td>11.2%</td>
<td>0.98</td>
</tr>
<tr>
<td>0.10 - 0.15</td>
<td>15.5%</td>
<td>1.00</td>
</tr>
<tr>
<td>0.15 - 0.25</td>
<td>21.3%</td>
<td>1.05</td>
</tr>
<tr>
<td>0.25+</td>
<td>29.7%</td>
<td>1.26</td>
</tr>
</tbody>
</table>
### Table 4: Within-Firm Growth: Average Growth at Firm, Relative Wage Growth, and Promotions

<table>
<thead>
<tr>
<th>Likelihood, Average Wage Growth at Firm, Relative Wage Growth, and Promotions</th>
<th>(1) Men to Women</th>
<th>(2) Average Wage Growth at Firm</th>
<th>(3) Relative Wage Growth</th>
<th>(4) Share Obs. That Are Promotions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 0.075</td>
<td>0.92</td>
<td>0.022</td>
<td>0.009</td>
<td>0.02</td>
</tr>
<tr>
<td>0.075 to 0.15</td>
<td>1.09</td>
<td>0.035</td>
<td>0.071</td>
<td>0.24</td>
</tr>
<tr>
<td>0.15 to 0.225</td>
<td>1.21</td>
<td>0.040</td>
<td>0.141</td>
<td>0.81</td>
</tr>
<tr>
<td>0.225 to 0.25+</td>
<td>1.44</td>
<td>0.025</td>
<td>0.263</td>
<td>0.98</td>
</tr>
</tbody>
</table>

### Table 5: Probability and Size of Wage Growth, By Type

| Firm Switch Promotion Non-Promotion |
|---|---|---|---|
| Men | Women | Men | Women | Men | Women |

| Annual Probability |
|---|---|---|---|---|---|
| Ages 26-30 | 0.295 | 0.291 | 0.215 | 0.173 | 0.490 | 0.535 |
| Ages 31-35 | 0.247 | 0.234 | 0.160 | 0.131 | 0.593 | 0.635 |
| Ages 36-40 | 0.192 | 0.186 | 0.108 | 0.093 | 0.701 | 0.721 |
| Ages 41-45 | 0.145 | 0.147 | 0.075 | 0.069 | 0.780 | 0.785 |

| 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: A promotion is defined as a large, discrete wage jump relative to one’s co-workers at the same firm, in a year when an individual did not switch employers. See text for details.

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

<table>
<thead>
<tr>
<th>(1) Female</th>
<th>(2) Human Capital</th>
<th>(3) Tenure at Firm</th>
<th>(4) Firm FE</th>
<th>(5) Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.039***</td>
<td>-0.044***</td>
<td>-0.044***</td>
<td>-0.040***</td>
<td>-0.041***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

Notes: All columns control for time fixed effects. Human capital controls include indicators for age, years of higher education, field of major, as well as a quadratic in years of experience. N = 190,404. Sample includes all individuals ages 26 to 35 in years when a firm switch did not occur. Column 4 controls for being in a clerical or administrative support occupation. Columns 5 and 6 control for two- and three-digit occupation codes, respectively. In this and all subsequent tables, a promotion is defined as a large, discrete wage increase relative to one’s co-workers. See text for details.
Table 7: Gender Difference in Probability of Promotion Ages 26 to 35: Controls for Hours Worked

<table>
<thead>
<tr>
<th>Dep. Variable: Probability of Promotion</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-0.040***</td>
<td>-0.032***</td>
<td>-0.021***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Baseline Controls &amp; Firm Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Part-Time History &amp; Hours Worked</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Controls for Year of &amp; Year Following Birth</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*** Significant at 1% level. Notes: All regressions include controls for year, age, years of higher education, field of major, a quadratic in years of experience, as well as controls for part-time history and specific hours worked within the ranges provided. \( N = 190,404 \). Sample includes all individuals ages 26 to 35 in years when a firm switch did not occur.

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

<table>
<thead>
<tr>
<th>Promotion Gap</th>
<th>Promotion Rate</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>5 to 1 years before first birth</td>
<td>-0.032</td>
<td>0.198</td>
</tr>
<tr>
<td>Year of and first year after</td>
<td>-0.087</td>
<td>0.163</td>
</tr>
<tr>
<td>2 to 10 years after first birth</td>
<td>-0.015</td>
<td>0.124</td>
</tr>
</tbody>
</table>

Notes: Calculations above are for individuals who ever had children, in years when a firm switch did not occur.

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

<table>
<thead>
<tr>
<th>Low Daddy Month Take-Up Firm ( \times ) Female</th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Daddy Month Take-Up Firm</td>
<td>-0.014**</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Female</td>
<td>0.021***</td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>-0.027***</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

** Significant at 5% level. *** Significant at 1% level. Notes: Coefficients are from a regression that includes indicators for year, age, years of higher education, and field of major, as well as a quadratic in years of experience. The outcome variable is whether an individual was promoted in a given year. A firm is considered to have low uptake of daddy months if fewer than half of male employees at the firm take their 60 days of allotted parental leave within the first two years after the birth of their child.

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

Notes: Graph follows college-educated individuals from the 1960-1970 birth cohorts in Sweden. For sample details, see Section 2. Source: Statistics Sweden.
**Figure 2:** 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 -0.25 and 0.25.

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

Notes: The histograms tabulate individual-year level observations of real annual wage growth. The tails of the distribution are collapsed to mass points, at -0.25 and 0.25.
Figure 4: Decomposition of Gender Difference in Cumulative Wage Growth Since Age 25

Figure 5: Decomposition of Lifecycle Wage Growth

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

A. Share Switching Firms

B. Share Promoted, If Stayed At Firm

Notes: Panel A graphs the share switching firms out of all workers. Panel B graphs the conditional share promoted, out of those who did not switch firms in the current period.
Figure 7: Gender Gap in Promotions, by Years Relative to First Birth

![Graph showing the gender gap in promotions relative to first birth.](image)

Notes: Confidence intervals omitted for clarity. See Table A.8 for standard errors for all coefficients. All regressions include baseline controls for year, age, years of higher education, field of major, and a quadratic in years of experience. The promotion gap represents the coefficient on “female,” or the gender difference in promotion probability.

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

A. Parenthood Effects

![Graph showing parenthood effects.](image)

B. Motherhood Penalty & Total Promotion Gap

![Graph showing motherhood penalty and total promotion gap.](image)

Notes: The “motherhood penalty” in Panel B is equal to the difference between the motherhood and fatherhood effects, graphed in Panel A. The “total promotion gap” plots the estimates for the full promotion gap from Table 7. Confidence intervals omitted for clarity. See Table A.8 for standard errors. All regressions include baseline controls for year, age, years of higher education, field of major, and a quadratic in years of experience.
Figure 9: Gender Effects by Time to First Birth (Parents) and Gender Effects by Age (Childless)

A. Gender Effect, Ever Have Children

B. Gender Effect, Childless

Notes: The “gender effect” in Panel A is equal to the difference between the total promotion gap and the motherhood penalty, graphed in the same figure. The corresponding gender effect in Panel B is the set of coefficients on the interactions between female and age, in a regression with promotion probability as the outcome variable, and including controls for years of education, field of study, years of experience, tenure, hours worked, and firm fixed effects. Shaded area corresponds to the 95% confidence intervals.

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

Notes: The shaded area represents the 95% confidence interval. Regressions include controls for year, age, years of higher education, field of major, 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 $\theta_H$, and one with low innate ability, $\theta_L$. Cut-offs $\eta_2, \eta_3, \eta_4$ 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: Sample simulation is for individuals with children only. The technology parameters are $(d_1, d_2, d_3, d_4) = (1, 0.8, 0.6, 0.4)$, $(c_1, c_2, c_3, c_4) = (0.5, 1.3, 1.5)$. The human capital accumulation function corresponds to $(f(1), f(2), f(3)) = (0.95, 1.1, 1.3)$; $p_f = 0.8$; and $k_1 = 0.1, k_2 = 0.15$. In panel B, conditional wage growth is calculated for all individuals who were either not promoted or switched jobs, across all jobs.
Appendix A  Empirical Analysis

This appendix provides additional evidence related to the empirical analysis of wage growth and internal mobility.

A.1 Wage Growth Analysis

This section provides supplementary figures and graphs that complement the analysis in Section 4. Figure A.1 documents the same patterns as Figure 2, but with the tails collapsed instead at 0.5 and -0.5 log points, showing the majority of the tail behavior for men and women.

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 levels of wage shocks, mean reversion is low. 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.

**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

<table>
<thead>
<tr>
<th>$\ln w_{t+1} - \ln w_t$</th>
<th>-0.2</th>
<th>-0.1</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln w_{t+5} - \ln w_t$</td>
<td>-0.1</td>
<td>0.51</td>
<td>0.17</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>0.0</td>
<td>0.28</td>
<td>0.49</td>
<td>0.26</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>0.1</td>
<td>0.08</td>
<td>0.15</td>
<td>0.32</td>
<td>0.16</td>
<td>0.06</td>
</tr>
<tr>
<td>0.2</td>
<td>0.06</td>
<td>0.09</td>
<td>0.19</td>
<td>0.26</td>
<td>0.13</td>
</tr>
<tr>
<td>0.3</td>
<td>0.03</td>
<td>0.05</td>
<td>0.10</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>0.4</td>
<td>0.02</td>
<td>0.03</td>
<td>0.06</td>
<td>0.14</td>
<td>0.19</td>
</tr>
<tr>
<td>0.5</td>
<td>0.02</td>
<td>0.03</td>
<td>0.06</td>
<td>0.19</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table A.2: Concentration of Lifecycle Wage Growth

<table>
<thead>
<tr>
<th>Share of individuals who achieved, during three yrs. of greatest wage growth:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>50% of lifecycle wage growth</td>
<td>0.80</td>
</tr>
<tr>
<td>60% of lifecycle wage growth</td>
<td>0.55</td>
</tr>
<tr>
<td>70% of lifecycle wage growth</td>
<td>0.34</td>
</tr>
<tr>
<td>50% of lifecycle wage growth (excluding first three years after graduation)</td>
<td>0.81</td>
</tr>
</tbody>
</table>

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 5.2. In particular, we provide 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 in Section 3. 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 upward movement, both in our cohort and among high-skill workers overall.

In firms with fewer opportunities for upward movement, women and men represent a roughly equal share of workers. However, at firms that are in the upper half of the distribution for the yearly share of workers promoted, women represent a lower share of the firm’s high-skilled employees, around 35-43%. 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 5 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.

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

<table>
<thead>
<tr>
<th>Gap</th>
<th>Share of Gap Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Gap In Promotion Rates</td>
<td>3.88</td>
</tr>
<tr>
<td>Counterfactual: Same Distribution Across Firms</td>
<td>3.05</td>
</tr>
<tr>
<td>Counterfactual: Same Promotion Rate Within Firms</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Notes: Calculations are for all individuals ages 26 to 35. 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 promotion is defined as a large, discrete wage jump relative to one’s co-workers. See text for details. N = 190,404.
A.3 Classification of Occupations

The two-digit occupation codes used in the paper 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.

Physical, mathematical and engineering science 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.

Physical and engineering science 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; Busi-
ness 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. The 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.

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

---

**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 exceed 100% FTE (40 hours per week). For women, the proxy combines the effect of hours worked above full-time, as well as time away for parental leave, and therefore is lower in most periods. Since the proxy measure captures additional variation in hours, we rely on it whenever we analyze specific hours worked.

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**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages 26-30</td>
<td>0.04</td>
<td>0.11</td>
<td>39.1</td>
<td>38.0</td>
<td>41.9</td>
<td>38.6</td>
</tr>
<tr>
<td>Ages 31-35</td>
<td>0.04</td>
<td>0.19</td>
<td>39.2</td>
<td>37.1</td>
<td>41.9</td>
<td>34.9</td>
</tr>
<tr>
<td>Ages 36-40</td>
<td>0.05</td>
<td>0.23</td>
<td>39.1</td>
<td>36.7</td>
<td>40.5</td>
<td>33.7</td>
</tr>
<tr>
<td>Ages 41-45</td>
<td>0.04</td>
<td>0.18</td>
<td>39.2</td>
<td>37.2</td>
<td>40.6</td>
<td>36.0</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Weekly hours:</th>
<th>Promotion Rate</th>
<th>Share of Men</th>
<th>Share of Women</th>
<th>Promotion Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 to 32</td>
<td>0.147</td>
<td>0.051</td>
<td>0.140</td>
<td>-0.027***</td>
</tr>
<tr>
<td>32 to 39</td>
<td>0.164</td>
<td>0.123</td>
<td>0.139</td>
<td>-0.036***</td>
</tr>
<tr>
<td>39 to 41</td>
<td>0.159</td>
<td>0.341</td>
<td>0.339</td>
<td>-0.032***</td>
</tr>
<tr>
<td>41 to 48</td>
<td>0.199</td>
<td>0.321</td>
<td>0.212</td>
<td>-0.031***</td>
</tr>
<tr>
<td>48 or more</td>
<td>0.244</td>
<td>0.145</td>
<td>0.053</td>
<td>-0.031***</td>
</tr>
</tbody>
</table>

The promotion rate is calculated for men, and is conditional on not having switched firms in the observation year. All hours calculations use the proxy hours measure. See text for details.

### A.5 Normalization of the Motherhood Penalty

In this section, we briefly discuss the required normalization for the motherhood penalty, and how the choice of normalization affects any qualitative results in Section 5.3.

In the existing literature (e.g., Kleven et al. (2018), 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.

Note that any change in 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). Specifically, motherhood penalties still dominate after first birth, while gender penalties are more prominent.
early in the lifecycle. 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. Under the 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 3, 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 A.6 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.
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

<table>
<thead>
<tr>
<th>Promotion Gap</th>
<th>Promotion Rate (M)</th>
<th>Wage Gain F</th>
<th>Wage Gain M</th>
<th>% Explained by Promotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 7.5</td>
<td>-0.047*** (0.002)</td>
<td>0.28</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>n = 10</td>
<td>-0.040*** (0.002)</td>
<td>0.20</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>n = 12.5</td>
<td>-0.032*** (0.002)</td>
<td>0.15</td>
<td>0.24</td>
<td>0.22</td>
</tr>
<tr>
<td>n = 15</td>
<td>-0.026*** (0.002)</td>
<td>0.11</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>Median-based,</td>
<td>-0.042*** (0.002)</td>
<td>0.22</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>n = 10</td>
<td>Absolute wage growth,</td>
<td>0.29</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td>10+ percent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 major 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 ages 40-45 is quite large, at 0.24, compared to 0.25 in 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 0.26. 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 switches is approximately similar in the two groups. However, the importance of gender differences in non-promotion growth increases in the full population. A combination of two factors accounts for this pattern. First, wage growth in non-promotion periods is lower in the omitted majors than in the baseline population. Second, women make up a larger share of the omitted majors. Correspondingly, the share of gender differences in wage growth explained by differences promotion-related growth decreases moderately, to 62%.
Table A.7: Summary Results: Baseline vs. Full Population

Panel A: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Baseline Population</th>
<th>Omitted Majors</th>
<th>All Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promotion Gap</td>
<td>0.040***</td>
<td>-0.033***</td>
<td>-0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Mean Wage (Men), Ages 40 to 45</td>
<td>10.54</td>
<td>10.35</td>
<td>10.46</td>
</tr>
<tr>
<td>Mean Wage Gap, Ages 40 to 45</td>
<td>0.25</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>Men</td>
<td>60,353</td>
<td>12,398</td>
<td>72,751</td>
</tr>
<tr>
<td>Women</td>
<td>42,602</td>
<td>33,839</td>
<td>76,441</td>
</tr>
</tbody>
</table>

Panel B: Decomposition Results

<table>
<thead>
<tr>
<th>% of Gender Difference in Wage Growth Explained by:</th>
<th>Firm Switches</th>
<th>Promotions</th>
<th>Non-Promotion Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Population</td>
<td>0.28</td>
<td>0.78</td>
<td>-0.06</td>
</tr>
<tr>
<td>Full Population</td>
<td>0.31</td>
<td>0.62</td>
<td>0.07</td>
</tr>
</tbody>
</table>

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.

A.8 Other Supplementary Tables and Figures

Figure A.5: 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

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

* 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 6, we noted that a drawback the classic model by Gibbons and Waldman (1999b) that we build on is that it generates promotions, but not necessarily exceptionally large wage growth at time of promotion. This can be rectified through one of two possible extensions: either by incorporating private information by the employer about the worker, as is developed in the same paper by Gibbons and Waldman; or by incorporating compensation for effort associated with different jobs, the approach we take. For simplicity, we left this out of the description of the model in Section 6, but add this feature to the exposition and proofs below.

To do this, we assume that an (observable) effort cost $e_j$ for the worker is associated with each particular job, with $0 \leq e_1 < e_2 < ... < e_J$. This assumption is meant to capture that jobs higher up the career ladder tend to be associated not only with more complexity (requiring higher human capital), but also 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 added effort costs. A worker is initially hired to job 1. The worker will be assigned efficiently to job 2 when his ability exceeds the $\eta$ that solves

$$d_1 + c_1 \eta_{it} - e_1 = d_2 + c_2 \eta_{it} - e_2.$$  

We denote this solution, as before, as $\eta_2$, where

$$\eta^2 = \frac{d_1 - d_2 + (e_2 - e_1)}{c_2 - c_1}.$$  

As in the benchmark model without effort costs, the worker is paid the competitive wage $w_{itj} = d_j + c_j \eta_{it}$. However, wages now increase discontinuously at the time of the job change. To see why, consider Figure A.6. In the model without effort costs, recall that individuals are promoted as soon as their effective ability is such that their output in job 2 exceeds their output in job 1. This point is marked by $\eta''$ in the figure. However, as the figure shows, wage does not jump discontinuously at this point. By contrast, in the model with effort costs, a worker will strictly prefer working at job 1 if his ability is equal to $\eta''$, or to any value below $\eta^2$, since in that case $u_{it1} > u_{it2}$. To be induced to work in job 2, a worker with $\eta_{it} < \eta^2$ would have be paid a wage that is higher than his output. Therefore, the worker is promoted only at the higher threshold, and the wage jumps by $e_2 - e_1$ at time of promotion.
In equilibrium, free entry, costless switching of workers between firms, and competition for workers will yield the efficient job assignment, i.e. \( d_j + c_j \eta_{ij} - e_j \) is maximized for each worker. This implies that workers in the benchmark environment with effort costs are promoted according to the following cut-offs:

\[
\eta_j^j = \frac{d_{j-1} - d_j + (e_j - e_{j-1})}{c_j - c_{j-1}}.
\]

(9)

Figure A.6: Promotion from Job 1 to Job 2

B.2 Proof of Proposition 1

To prove proposition 1 for the augmented model, we begin with the cut-off values for men and women, respectively. As already discussed, for men 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 \( \eta \) in period 3 that determine job assignment for women are identical to those for men: \( \eta_3^2, \eta_3^3, \eta_4^4 \), 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: $\bar{\eta}^2$ and $\bar{\eta}^3$, 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. In period 1, there is only one possible type of promotion for men and women: from job 1 to job 2. Let $V_1(j=1)$ 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 $V_1(j=2)$. In this case, we have:

$$V_1(j=1) = d_1 + c_1\eta_t - e_1 + p_f \cdot (-k_1) + (1 - p_f) \cdot V_2^*$$
$$V_1(j=2) = d_2 + c_2\eta_t - e_2 + p_f \cdot (-k_2) + (1 - p_f) \cdot V_2^*$$

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 $\bar{\eta}^*$ that equalizes expected values of female employees in jobs 1 and 2 solves

$$d_1 + c_1\bar{\eta}^* - e_1 + p_f \cdot (-k_1) = d_2 + c_2\bar{\eta}^* - e_2 + p_f \cdot (-k_2).$$

The solution to this equation is

$$\bar{\eta}^* = \frac{d_1 - d_2 + (e_2 - e_1)}{c_2 - c_1} + p_f \frac{k_2 - k_1}{c_2 - c_1} = \bar{\eta}^2 + p_f \frac{k_2 - k_1}{c_2 - c_1} > \bar{\eta}^2. \quad (10)$$

where the final inequality follows from the fact that $p_f > 0$, $k_2 > k_1$ and $c_2 > c_1$. Thus, in period 1 the ability threshold for promotion applied to women is higher than for men. Next, we derive the cut-off values $\theta^j_1$ 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 1 to job 2), 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 $\bar{\eta}^2$, 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 2, which is $\bar{\theta}^2_1 = \frac{\bar{\eta}^2}{f(1)}$. Thus, share $(1 - \bar{\theta}^2_1)$ of men are promoted. For women, the higher threshold value $\bar{\eta}^*$ determines a correspondingly higher threshold value of innate ability, which we denote as $\bar{\theta}^*_1 = \frac{\bar{\eta}^*}{f(1)} > \frac{\bar{\eta}^2}{f(1)} = \bar{\theta}^2_1$. Thus, under our
assumption that \( \theta \) is continuous and uniformly distributed between 0 and 1, the share of all women who are promoted, regardless of future childbearing status, \((1 - \theta_1^*)\), is strictly lower than the share of men who are promoted, \((1 - \theta_2^*)\). The difference \(\theta_2^* - \theta_1^*\) is the gender penalty in promotion rates in period 1. QED.

**Proof of Proposition 1(ii):** In period 2, fathers have strictly positive promotion probabilities to jobs 2 and 3, under the restricted set of parameterizations and under the assumption that \( f'(<) > 0 \). 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 \((1 - \theta_2^3)\) of both fathers and childless men to job 3, and share \((\theta_2^2 - \theta_2^3)\) to job 2 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 have substantially lower promotion rates than 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 significantly 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 \((1 - \theta_3^3)\) are promoted to job 3, same as for men. For promotion rates to job 2 in period 2, there are two possible cases for childless women. For the case that \(\theta_1^* < \theta_2^3\), childless women’s promotion rate to job 2 will be \((\theta_1^* - \theta_2^3) + (\theta_2^3 - \theta_2^3)\). For the case that \(\theta_1^* > \theta_2^3\), childless women’s promotion rate to job 2 will be \((\theta_1^* - \theta_2^3) + (\theta_2^3 - \theta_2^3)\).

Note that in both cases, the second term – \((\theta_1^* - \theta_2^3)\) (case 1) or \((\theta_2^3 - \theta_1^*)\) (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 \(\theta_2^3\) in period 3 for men. Both fathers and childless men have \(\tau = 3\) years of experience, and share \((1 - \theta_3^3) + (\theta_2^3 - \theta_3^3) + (\theta_2^3 - \theta_2^3)\) of men are promoted to jobs 4, 3 and 2 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 4 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 $(1 - \theta_2^3) + (\theta_1^3 - \theta_2^3) + (\theta_1^2 - \theta_2^2)$ are promoted for the case that $\theta_1^* \leq \theta_2^3$, and share $(1 - \theta_2^3) + (\theta_1^3 - \theta_2^3) + (\theta_2^2 - \theta_3^3)$ are promoted for the case that $\theta_1^* > \theta_2^3$. 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 positive “gender effect” in period 3.

**B.3 Proof of Proposition 2**

The proof of proposition 2 is a direct implication of equation (10). Since in equation (10), $\bar{\eta}^i$ is increasing in $k_2 - k_1$, this implies that $\bar{\eta}^i f(\bar{\eta}^i)$ is also increasing in $k_2 - k_1$. Therefore, the negative gender penalty $\overline{\theta}^1 - \overline{\theta}^* \theta$ becomes larger (more negative) with $k_2 - k_1$. 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, $w_{mijt} = d_j + c_j \eta_{it}$. 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 $w_{mijt} = d_j + c_j \eta_{it}$, 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 $-k_j$, and cannot be fired in this period. We begin with period 3, which is the most straightforward as there is no uncertainty about childbearing or labor supply. Correspondingly, in period 3 women are also compensated according to their production, $w_{fit} = d_j + c_j \eta_{it}$, 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 \( \Pi_{j1} \) 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 \( \pi^*_2 \) 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 \( \pi^*_2 \) 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 \( \epsilon \), 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.\(^{28}\)

To summarize, the wage function for men is identical to that in the benchmark environment: \( w_{ijt}^m = d_j + c_j \eta_{it} \). Women’s wage function varies by period and takes the

\(^{28}\)In the latter case, the wage for women in period 1 would be \( w_{ij1}^f = d_j + c_j \eta_{i1} - p_f k_j(1 - \epsilon) \), since employers take into account the now lower probability of incurring cost \( k_j \) in the following period.
following form:

Period 0: \( w_{ij0}^f = d_1 + c_1 \eta_{i0} \)

Period 1: \( w_{ij1}^f = d_j + c_j \eta_{i1} - p_{fj}k_j \)

Period 2: \( w_{ij2}^f = d_j + c_j \eta_{i2} \) if childless, on govt.-paid leave otherwise

Period 3: \( w_{ij3}^f = d_j + c_j \eta_{i3} \).

With these wage functions in hand, it is possible to analyze wage growth both during firm changes and in non-promotion periods. In general, this model is not designed to study firm changes in depth. However, by introducing an assumption that there is a positive separation probability \( \epsilon > 0 \) at the end of every period, we can analyze average wage growth for those who must switch firms following exogenous separation. Under a positive separation probability, the wages look identical to those derived above, except that for women in period 1, the wage would be \( w_{ij1}^f = d_j + c_j \eta_{i1} - p_{fj}k_j(1 - \epsilon) \), as discussed in footnote 28. After a firm separation, individuals move to (an identical) new firm, that hires them either into the same job as before, or promotes them into 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 \( p_{fj}k_j \) 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 \( \eta_{i1} \). Additionally, some high-ability women who would have been promoted if they were men do not advance to job 2 in 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.

Later in the lifecycle, the opposite occurs. Consider a woman who was on leave in period 2, with a most recent wage of \( w_{ij1}^f \). Since there is no more uncertainty about her childbearing, her wage is now bid back up to the 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. Therefore, 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. For this reason, we also do not present this result as a formal proposition, but only demonstrate that it is true for
some parameterizations, such as the one in Figure 11.29

A similar pattern holds for periods in which no promotion occurs. According to the wage functions derived above, 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 true for some 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 period 1, a result that we indicated does not depend on the uncertainty about future childbearing and leave-taking has been resolved.

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

\[
\Delta w_1^m = \left( c_1 \cdot \frac{1}{2} (0 + \eta^2) \right) \cdot \eta^2 + \left( d_2 - d_1 + c_2 \cdot \frac{1}{2} (\eta^2 + 1) \right) \cdot (1 - \eta^2)
\]

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

\[
\Delta w_1^f = \left( c_1 \cdot \frac{1}{2} (0 + \eta^2) - p_f k_1 (1 - \epsilon) \right) \cdot \eta^2 + \left( d_2 - d_1 + c_2 \cdot \frac{1}{2} (\eta^2 + 1) - p_f k_2 (1 - \epsilon) \right) \cdot (1 - \eta^2)
\]

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

\[
\Delta w_1^m = \left( c_1 \cdot \frac{1}{2} (0 + \eta^2) \right) \cdot \eta^2 + \left( d_2 - d_1 + c_2 \cdot \frac{1}{2} (\eta^2 + 1) \right) \cdot (1 - \eta^2)
\]

\[
\Delta w_1^f = \left( c_1 \cdot \frac{1}{2} (0 + \eta^2) - p_f k_1 (1 - \epsilon) \right) \cdot \eta^2 + \left( c_1 \cdot \frac{1}{2} (\eta^2 + 1) - p_f k_1 (1 - \epsilon) \right) \cdot (1 - \eta^2)
\]

Note that the first and third term are strictly lower for women than men, since \( p_f k_1 (1 - \epsilon) > 0 \). The second term is strictly lower for women than men, since \( p_f k_2 (1 - \epsilon) > 0 \), and \( d_2 - d_1 + c_2 \cdot \frac{1}{2} (\eta^2 + \eta^2) > 0 \), or alternatively \( \frac{1}{2} (\eta^2 + \eta^2) > \frac{d_2 - d_1}{c_2 - c_1} \). The final inequality follows from: \( \frac{1}{2} (\eta^2 + \eta^2) > \eta^2 = \frac{d_2 - d_1}{c_2 - c_1} \).
B.5 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 and 1, and therefore accumulate human capital at the same rate in those periods.

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 $\theta_i$ 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 $\theta$, 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 $\theta$, 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. This is a potentially strong assumption which we now examine.

We first note that, in our model, 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 $k_j$ in our model 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, $(d_j + c_j \eta_j)h$. 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 5.3. 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 \( \gamma_i \) 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

\[
\eta_{it} = \theta_i f(h_0, \ldots, h_{t-1}),
\]

for \( t > 0 \). Wage functions take a similar form as before. A man in period \( t \) is paid

\[
w_{i};t = d_j + c_j \eta_{it} = d_j + c_j \theta_i f(h_0, \ldots, h_{t-1}).
\]

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

\[
w_{it} = \omega(h_0, \ldots, h_{t-1}; g_i, \theta_i, t),
\]

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:

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

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

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:

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

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

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

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

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

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