Women in the Pipeline: A Dynamic Decomposition of Firm Pay Gaps

Jishnu Das* Clément Joubert†

February 16, 2022

Abstract

This paper proposes a new decomposition method to understand how gender pay gaps arise within firms. The method accounts for pipeline effects, non-stationary environments, and dynamic interactions between pay gap components. The decomposition is applied to a new data set covering all employees at the World Bank Group between 1987 and 2015 and shows that historical differences in the positions for which men and women were hired account for 77 percent of today's average salary difference, dwarfing the roles of entry salaries, salary growth, or retention differences. Forward simulations show that 20 percent of the total gap can be assigned to pipeline effects that would resolve mechanically with time.

JEL Classification: J16, J31, J7, M5

Keywords: Gender Pay Gap, Dynamic Decomposition, Pipeline effects

*McCourt School of Public Policy and Walsh School of Foreign Service. E-mail: jishnu.das@georgetown.edu.

†World Bank. E-mail: cjoubert@worldbank.org (corresponding author).

1 Introduction

Aggregate pay gaps (also known as unconditional or raw pay gaps) can serve as holistic measures of gender disparities in pay, and are often quoted in the press.[1] Assuming innate differences in productivity and preferences are small across genders, large aggregate pay gaps must reflect discrimination and inefficiency in compensation, in job assignment (hires, promotions, retention) or, upstream, in the production of human capital.[2]

In contrast, legal recourse against sexual discrimination in compensation historically relies on the narrower concept of "Equal Pay for Equal Work" (EPEW), which is concerned with pay gaps between equally productive workers in identical occupations but silent on occupational segregation or inequality in promotions.[3]

Consequently, new transparency regulations are expanding firms' responsibility beyond EPEW. In the United Kingdom, the "Equality Act 2010 (Gender Pay Gap Information) Regulations 2017" now requires all firms with 250-plus employees to report the difference between the mean hourly rate of pay of all male employees and that of all female employees.[4]

To understand how firms will (or should) respond to these new transparency

[1]For example, gender equality advocates calculate the calendar date at which women start "working for free", i.e. when the remaining fraction of the year equals the percentage gender pay gap (G.V. (2017)).

[2]The assumptions underlying this argument would include small inherent gender differences in productivity outside the labor market, such as in child rearing or home production.

[3]A large literature has evaluated violations of the EPEW principle by statistically controlling for differences in job characteristics and/or measures of a worker's productivity: See Blau and Kahn (2017) for a survey. As Flabbi (2010) discusses, residual gender pay gaps could be attributed either to discrimination or to unobserved productivity differences, unless economic and distributional assumptions are made to separate those two factors.

[4]Obligations to report pay gaps have also emerged in France, Denmark, Belgium, Germany, Italy, the Netherlands, Ireland, Switzerland, Australia, and Canada (see, for example, *Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee* 2017). A similar bill (AB-1209) was vetoed by California governor Brown in October 2017.

requirements, we must first understand how aggregate pay gaps arise within a firm. This paper addresses three key considerations that have received little attention in the literature. First, historical patterns in hiring can generate lasting pay gaps long after the original imbalances have disappeared. Consider a firm that satisfies EPEW and applies identical promotion rates to both genders, but hired fewer women than men in the past. Because of "pipeline" effects, it will exhibit an aggregate gender pay gap and an apparent glass ceiling.[5] If the role of pipeline effects is mistakenly assigned to other factors, firm policy responses aimed at fixing the pay gap could create further imbalances. Second, existing gender pay gap decompositions assume a constant environment, in the sense that the returns to endowments (aka. the "wage structure") are time-invariant.[6] In fact, we show in our appplication that firm processes, and the gender pay gap itself, can exhibit large fluctuations over time. Third, existing studies tend to examine each firm process separately, even though gender disparities in hires, salary growth and exits interact dynamically.[7] As an example, high turnover tends to decrease the quantitative importance of salary growth differences relative to initial salaries differences.

This paper proposes a new dynamic decomposition method that can be applied to aggregate pay gaps to quantify the relative contributions of the gender compo-

[5]Suppose a firm employs two-period-lived male and female workers. In year two the firm hires 50 men and 50 women, but in the previous year it hired more men, say 90 men against 10 women. Each period, half of the workers are promoted to higher paying managerial jobs, with no discrimination against women. Therefore, in year two the firm will employ 45 male managers, 95 male workers, five female managers and 55 female workers. This firm exhibits an aggregate gender wage gap in year two without any discrimination in promotions or within-job salaries.

[6]See Filmer et al. (2005) for an example using a 1997 cross-section of our panel.

[7]See, for example, Bielby and Baron (1984) and Petersen and Morgan (1995) on occupational segregation, Albrecht et al. (2003), Arulampalam et al. (2007), Guvenen et al. (2014), Fernandez and Campero (2017), Angelov et al. (2016), Kleven et al. (2018) or Albrecht et al. (2018) on salary growth and glass ceilings, and Blau and DeVaro (2007) on the mixed evidence on gender promotion gaps. Recent work has shown how occupational segregation feeds back into within-job wage differences as female managers are associated with better female compensation: e.g. Cohen and Huffman (2007), Matsa and Miller (2011), Flabbi et al. (2016), Kunze and Miller (2017).

sition of hires, entry salaries, salary growth, and retention, while accounting for pipeline effects and a changing firm environment.[8] The method estimates auxiliary, reduced-form models for each of the decomposition factors in each year, before aggregating them through microsimulations. We then obtain the decomposition through counterfactual simulations of male and female salaries in which each source of pay differences is shut down at a time. The simulations can also be extended forward to quantify pipeline effects. Rather than assume that data are drawn from a steady state, as in traditional wage decomposition methods, our approach accounts for the constant changes in hiring, firing, growth, and shrinkage that are typical in most firms.

To apply our methodology, we assemble a new panel containing 27 years of all personnel records from the World Bank Group (WBG), a multilateral finance organization with more than 16,000 employees in 2015. The panel length, data size, recency, and measurement precision afforded by these personnel records offer the perfect setting for the specific purpose of this paper, which is to analyze the within-firm dynamics of gender pay gaps. The firm is a policy-relevant unit of analysis because firms increasingly bear the legal responsibility of eliminating gender pay gaps. In addition, we show that the size and trends exhibited by gender pay gaps and employment differences in our sample are similar to those found among U.S. employees as a whole.

Our application of the decomposition method to the WBG offers a striking example of how past hiring stocks can affect pay gaps today. Our data show a decline in the aggregate gender gap from 50 cents on the dollar in 1987 to 23 cents on the dollar in 2015. For the mean salaries in our data, this amounts to an annual difference of US$27,400. However, we find that three-quarters of the 2015 gap was due to

[8] The Fortran and Stata programs that implement the methodology are available at https://github.com/clemjoub/Firm-Pay-Gaps-Decomposition

historical differences in the types of jobs at which women and men were hired. By contrast, less than 10 percent of the pay gap is due to differences in entry salaries; 5 percent is due to differences in salary growth (including promotions and raises) and less than 1 percent is due to differences in retention.[9]

We are also able to quantify how much of the gap will tend to dissipate over time in the absence of any policy or environmental change. When we simulate what would happen to the pay gap if 2015 hiring patterns were kept in place and no changes were made to compensation, we find that over the next 10 years, the aggregate pay gap continues to decline by one fifth (or five percentage points) but then stabilizes. The reason why we see these patterns is that, historically, most hires in technical positions were men; because these hires formed the pipeline for management jobs today, they continue to exert a negative influence on the gender pay gap. By 2015, the bias in hiring among the technical staff had virtually disappeared, which explains our simulated reduction in the pay gap. The persistent residual gap reflects the continued over-representation of women in support staff positions; reducing the gap further requires that men are hired into supporting staff positions as well.

Lastly, we show that the patterns we uncover are not specific to the gender pay gap but also apply to pay gaps associated with employee nationality (broadly categorized in developed versus developing countries). Our methodology could similarly be applied to pay gaps between employees of different races, or disability status, among other examples.

Our main contribution is therefore to propose a simulations-based dynamic decomposition method to understand how gender pay gaps are produced within firms.

[9]The residual 9% correspond to gender pay gaps among incumbent staff in the first year in our data: this portion of the gap cannot be decomposed into hiring, growth and exits because these processes took place before we start observing those employees.

Studies performing dynamic decompositions based on a fully specified economic model can be found in the literature, applied to different contexts (e.g. Keane and Wolpin 2010, Joubert 2015). In contrast to that approach –but in line with classic decomposition methodologies– we do not account for behavioral responses or permanent unobserved heterogeneity. Therefore, our counterfactual simulations should not be interpreted as policy experiments but rather as accounting exercises.[10]

Two non-structural studies share similarities with our approach. Bourguignon et al. (2008) decompose country-level income distribution differences. Although their exercise is static, they also use simulations to aggregate semi-parametric models of each determinant of income dispersion, and generate counterfactual distributions in which the differences in the parameters governing each individual factor are shut down one at a time. Gayle et al. (2012) implement a decomposition that explicitly captures individual career dynamics but the object of their decomposition is the gap in the cumulative earnings of U.S. male and female executives, rather than a cross-sectional firm-level pay gap and hiring pipelines. Also, their approach relies on stationary analytical formulae rather than micro-simulations.

A second contribution of the study is to produce new evidence on the career stages at which pay gaps arise in organizations. The evidence comes from an original panel data set of uncommon length and width, extracted from recent personnel records of the World Bank Group. We show this organization exhibits the classic internal labor market features identified in other large US companies and discussed in Baker and Holmstrom (1995). Existing gender pay gap studies using single-firm data separately document the gender differences in job assignment (Malkiel and Malkiel,

[10] In fact, we lack the data to examine such behavioral responses where the gender gap is most salient – historical hiring patterns. For instance, when there is a change in the president of the WBG, churn increases sharply. It may be that the applicant pool changes in anticipation of this churn or immediately after. Unfortunately, the institution does not retain data on the historical applicant pool, and therefore incorporating behavioral responses is currently outside the scope of our work.

1973; Ransom and Oaxaca, 2005), entry salaries (Gerhart, 1990), salary growth and promotion (Jones and Makepeace, 1996; Hersch and Viscusi, 1996; Barnet-Verzat and Wolff, 2008) and/or retention (Petersen and Saporta, 2004; Giuliano et al., 2005; Gobillon et al., 2014), but they do not provide, as we do, a quantitative comparison of these gender pay gap sources.

The originality of our approach is to quantify the career stages at which pay gaps arise, rather than to disentangle the contribution of observable worker endowments (such as experience or education) from that of prices as classic regression-based decompositions do. It allows follow-up investigations and remedial efforts to be focused on those organizational processes where the gaps are largest. If, for instance, most of the gender gap arose in entry salaries, a classic Oaxaca–Blinder decomposition could then be applied to entry salaries.[11] If instead, salary growth was the main pay gap source, follow-up efforts could investigate gender differences in performance rating, promotions, project allocations, or the impact of career interruptions due to children.

2 Decomposition Framework

2.1 A heuristic model of careers

The dynamic decomposition we propose is most relevant when employment is governed by an internal labor market (Doeringer and Piore (1985)), characterized by long-term worker-firm employment relationships ("careers"), which are not cap-

[11] The Oaxaca–Blinder decomposition evaluates the extent to which differences in an outcome variables (wages) can be explained by differences in characteristics (experience or education) versus differences in returns to these characteristic (see the seminal papers Oaxaca 1973 and Blinder 1973; Fortin et al. 2011 for a recent survey of the literature; and Filmer et al. 2005 for an application to a 1997 cross-section of World Bank personnel files).

tured by existing static decomposition methods. In the traditional economic theory of the firm, where salaries are instead determined and regularly updated in competitive spot markets, there is less scope for the dynamic linkages captured by our methodology.[12] Section 3 discusses the extent to which employment relationships in our application exhibit the characteristics of an internal labor market. To facilitate the interpretation of our decomposition results, we embed the method in a heuristic model of careers within an organization. Along the way, we point at the economic mechanisms that could be endogenized in a structural exercise before contrasting our approach with existing regression-based decompositions and with decompositions performed within a structural model in section 2.3.

We consider an organization in which new employees are hired into one of K ports of entry. Employees belong to different groups such as gender, disability status, or race. Each port of entry gives access to a separate career track k constituted of a sequence of positions that accrue increasing salaries. Group and track are permanent and immutable characteristics, so we can stratify the organization –and the sample in the empirical application– according to these two dimensions. To ease the notation we dispense with the track and group indices for the rest of this section.

In the labor economics literature, hiring is typically modeled as a matching process (possibly) with search frictions and asymmetric information on the match quality.[13] This type of model can yield endogenous distributions of offered and accepted salaries. We instead take as exogenous the number of hires and the distribution of salaries among new entrants in period t and denote them as: n_t^h and f_t^h. The dispersion in salaries among entrants to a track can be interpreted as the fact that new employees enter slightly higher or lower on the track. The subscript t accommodates

[12]The process of skill accumulation would still generate persistence in pay gaps in the context of spot markets if some groups have better access to tasks that increase human capital.

[13]For a review of this literature see Oyer and Schaefer (2011)

changes in labor market conditions or in internal hiring policies over time.

Every period, employees advance through their track and increase their salary. Salary increases could reward human capital accumulation, or reflect the firm's learning about the match quality or compensating employees as part of an incentive schemes. In our framework, an employee's percentage salary increase r_{it} depends on their current wage w_{it} and an unobserved idiosyncratic component ϵ_i^r, that could reflect effort or unobserved ability. The resulting conditional distribution of raises is denoted $f_t^r(./w)$. Larger increases in salary may be associated with promotions and smaller increases would correspond to regular raises in the absence of one. The salary increase function is group- and track-specific, so some tracks and some groups may exhibit faster salary progressions than others on average, possibly reflecting the higher average ability of their members. However, we do not allow for persistent sources of unobserved heterogeneity within employees in the same track and group. These assumptions allow us to reduce the set of "state variables" to only the current wage, which greatly simplifies the empirical application while still allowing for the dispersion of wages among incumbents to expand over time. We discuss how the decomposition results can be interpreted under these assumptions in section 2.3.

Once per period, employees may receive outside opportunities in the form of external job offers, self-employment or home production and decide whether to exit the organization. The expected present discounted utility offered by the most attractive opportunity is denoted V_{it}^l. Staff compare it with their expected present discounted value of staying in the firm $V_t(w_i, \epsilon_i^l)$, where ϵ_i^l summarizes unobserved determinants of the value of staying not captured by the current wage w_i. The probability that a worker with current salary w_i chooses to exit the firm is denoted by $p_t^l(w_i) = p(V_t(w_i, \epsilon_i^l) < V_{it}^l)$. The resulting number of leavers and distribution of salaries among them are n_t^l and f_t^l.

Every period, the organization inherits the distribution of salaries among incumbents, and in particular the pay gaps it exhibits, but it can update the rules and circumstances governing hires, salary increases and exits. Hiring rules affect the distribution of salaries offered to candidate entrants in each port of entry; rules governing promotions and raises determine the salary increases received by each employee as a function of their current salary; and retention policies affect the probabilities of exit (again as a function of current salary). This timing structure is designed to capture "pipeline effects", i.e. the fact that contracted commitments, internal norms, administrative rules and outside labor market incentive compatibility constraints prevent the organization from equating staffing and salaries across genders instantaneously. It is restricted in practice to adjusting promotions, hiring salaries and retention over time, taking the incumbent salary distribution as given.

2.2 Decomposing Pay gaps

We now consider two groups of employees, say women and men, denoted by the subscripts $g = f$ and $g = m$ and assume for simplicity that the organization has only one career track. For each group, combining the distribution of salaries among incumbents, new hires and leavers as well as the conditional distribution of salary increases (whether associated with a formal promotion or not) yields a law of motion for the overall distribution of salaries f_t^g:

$$f_{g,t+1}(.) = \mathcal{G}_{gt}\Big[n_{gt}, f_{gt}(.), f_{gt}^r(./w), n_{gt}^h, f_{gt}^h, n_{gt}^l, f_{gt}^l\Big] \tag{2.1}$$

$$n_{g,t+1} = n_{gt} + n_{gt}^h - n_{gt}^l \tag{2.2}$$

Iterating equations 2.1 and 2.2 backward to an initial period t_0 (say, the first year of available data) yields a relationship between the distribution of salaries for group

g at time τ as a function of the initial salary distribution of salaries at t_0, and the sequence of raise, hires, and exits from t_0 to τ:

$$f_{g\tau} \quad = \quad \mathcal{H}_{g\tau}\left[n_{gt_0}, f_{gt_0}, \{f^r_{gt}(./w)\}^{\tau}_{t_0}, \{n^h_{gt}\}^{\tau}_{t_0}, \{f^h_{gt}\}^{\tau}_{t_0}, \{n^l_{gt}\}^{\tau}_{t_0}, \{f^l_{gt}\}^{\tau}_{t_0}\right] \tag{2.3}$$

It follows that any difference in corresponding moments of f^f_τ and f^m_τ can be decomposed into four factors:

1. Differences in the salary distributions of incumbents at t_0 ($n^m_{t_0}$ and $f^f_{t_0}$ versus $n^f_{t_0}$ and $f^m_{t_0}$) [Legacy factor],
2. Differences in the distributions of hiring salaries (n^f_{ht} and f^f_{ht} versus n^m_{ht} and f^m_{ht}, $t = t_0, \tau$) [Hiring factor],
3. Differences in salary increases, including promotions and raises (r^f_t versus r^m_t, $t = t_0, \tau$) [Salary growth factor], and
4. Differences in exits (n^f_{lt}, f^f_{lt} versus n^m_{lt}, f^m_{lt}, $t = t_0, \tau$)[Retention factor].

In an organization with several career tracks, we can further separate the hiring factor above into (i) the relative number of men and women hired into different tracks [Composition of hires factor], and (ii) the initial salaries accepted by men and women for each given position upon hiring [Entry salaries factor].

The decomposition factors we just defined cannot be obtained via a tractable analytical formula. Instead, we design a dynamic, stochastic micro-simulator, described in section 4.1, which produces salary distributions from t_0 to t under the baseline and five counterfactual specifications indexed by c:

- baseline (c = 0): all parameters are gender specific,

- c = 1: female employees have the same probability as males of leaving the firm, conditional on their salary;
- c = 2: in addition, female employees draw their salary increases from the same distribution as males;
- c = 3: in addition, the distribution of female entry salaries is the same as that of males in each hiring position;
- c = 4: in addition, the firm hires the same number of female as male employees in each position; and
- c = 5: in addition, the distribution of salaries and grades among women is the same as among men at t_0.

The female salary distribution in year t obtained in counterfactual c is denoted by $f^f_{t,c}$. Denote the gap of interest $m(f^m_t) - m(f^f_t)$, approximated in our baseline scenario ($c = 0$) by: $m(f^m_{t,0}) - m(f^f_{t,0})$, where $m(.)$ is a moment of the distribution such as the mean. The contribution of factor c to the salary gap is defined as the reduction in the gender pay gap as we move from $c - 1$ to c, divided by the total gap: $\frac{[m(f^m_{t,c-1}) - m(f^f_{t,c-1})] - [m(f^m_{t,c}) - m(f^f_{t,c})]}{m(f^m_{t,0}) - m(f^f_{t,0})}$. Because only female salaries change in the counterfactual simulations, this simplifies to: $\frac{m(f^f_{t,c}) - m(f^f_{t,c-1})}{m(f^m_{t,0}) - m(f^f_{t,0})}$. In counterfactual 5, female and male salaries have the same distribution, so the contribution of factor c can also be written as: $\frac{m(f^f_{t,c}) - m(f^f_{t,c-1})}{m(f^f_{t,5}) - m(f^f_{t,0})}$. Other orderings of the decomposition can be performed in analogous ways. We show that our conclusions are robust to changes in the ordering of factors in section 5.5.

2.3 Interpretation of the decomposition

Our decomposition is original in that it measures the contribution to pay gaps of the different stages in a worker's career. By contrast, existing pay gap decompositions typically seek to separate the role of "endowments" (such as schooling or experience) from that of the "wage structure" (i.e. the "prices" of or "returns" to these endowments).[14] In particular, differences in endowments are considered "explained", while differences in the price of these endowments are "unexplained" and often interpreted as discrimination. Our goal is not to determine *why* pay gaps occur (in particular, whether they reflect discrimination), but rather *when* they form within organizations. The relative size of the decomposition factors (hiring, salary increases, retention) can help identify which policy levers are most likely to affect the aggregate gender pay gap in a firm.

While the decomposition is not based on a fully specified behavioral model, it does impose a "timing structure" on the salary data, in which the distribution of salaries evolves only incrementally over time under the influence of entries, exits and salary increases among incumbents. This structure generates inertia or "pipeline effects", and a number of non-behavioral dynamic interactions between decomposition factors. For example, if staff turnover is high, the weight of entry salaries in the overall distribution –and therefore the importance of the hiring factor– will be magnified. Conversely, if turnover is low, the salary growth factor will be more prominent. As another example, if higher salaries command steeper raises, increasing women's entry salaries will also improve their salary progression mechanically. If higher paid staff are more likely to leave, the effect of increasing entry salaries or raises on the gender pay gap will be dampened, etc.

Our decomposition counterfactuals face the same impediments to a causal in-

[14]Fortin et al. (2011)

terpretation as the other non-structural decomposition methods surveyed in Fortin et al. (2011): we propose a "statistical decomposition, of descriptive but not predictive value" (Bourguignon et al. (2003)). For example, the salary growth counterfactual gives women the same salary increases found among men (conditional on a staff's current salary) but assumes that entry salaries and exits remain the same (again, conditional on current salaries). This assumption shuts down interesting and relevant dynamic interactions between the decomposition factors: removing disparities in career progression across genders is likely to impact the relative qualities of male and female job applicants, their reservation wages or the level of effort of new hires, to name only a few examples. Our counterfactuals do not incorporate such endogenous responses in the same way that an Oaxaca-blinder decomposition will not allow women to acquire more schooling if the counterfactual wage structure, estimated from men's wages, rewards it more.

While it would be a valuable extension of our paper, structurally decomposing the aggregate pay gap in an organization would be particularly challenging because it would require jointly modelling the many economic decisions that affect it (human capital acquisition, job search, candidate screening, recruitment, incentives provision, and training to name a few). Indeed, policies aimed at eliminating gender differences in the aggregate pay gap are likely to trigger responses along each of these margins. The personnel economics literature has produced many candidate mechanisms for each these processes, and concluded that several of them are likely at play simultaneously (Oyer and Schaefer (2011)). To our knowledge, no structural model combining endogenous hiring, endogenous salary increases and endogenous exits exists in the literature. A fruitful approach could be to use the non-structural decomposition we propose to identify which career stages are quantitatively most important for the gender pay gap, and perform a structural decomposition that endogenizes that particular career stage to quantify the endogenous responses to candidate remedial policies.

Our methodology could also be extended to incorporate additional observed and unobserved permanent heterogeneity. In the case of observed heterogeneity, this could be done by stratifying the sample, estimations and simulations along, for instance, schooling. Our "composition of hires" factor could then be further separated into the relative numbers of men and women *with each given schooling levels* hired at each grade. The schooling differences could be shut down separately from the rest to assess the role of schooling differences. While this could be done in future applications of our method, our data do not contain good socio-demographic information such as schooling attainment or schooling quality.[15]

Adding heterogeneity to the procedure could have two technical benefits. The first benefit is to obtain a better approximation of the input processes. For example, if there were two very distinct groups of employees within a cell of the stratification that we consider (which already includes gender, nationality group and entry position), entry salaries could have a bimodal distribution that a normal distribution would not be able to approximate. However, as we show in section 4.3, our specifications allow us to fit average salaries by group and the aggregate pay gap very closely.

A second benefit of incorporating permanent heterogeneity would be to capture more complex dynamic linkages between processes in our counterfactual. For example, counterfactually improving the retention of women –to match that of men, say– should have an impact on the distribution of raises if the women who leave tend to have steeper salary progressions. This type of linkage is not captured by

[15]Note, in the case of schooling, that the grade of entry is very highly correlated with schooling: in fact schooling requirements are explicitly associated with each grade of entry. Therefore, the within-entry grade mix of schooling levels is unlikely to play a first order role in our application.

our procedure. Incorporating permanent unobserved heterogeneity would require a different estimation strategy, in which all parameters are estimated jointly by maximum likelihood or method of moments. The procedure would pick parameters that maximize the fit of simulations to the data, including all dynamic and cross-process correlations.

We decided to leave this extension out of the scope of the current paper for the following reasons. First, to implement the joint estimation of input parameters, we would need to reduce the number of such parameters by at least an order of magnitude. This would greatly reduce the flexibility of our model and its ability to fit the data, possibly offsetting the expected benefits described above. Second, this exercise would still fall short of capturing behavioral responses due to the absence of a structural model. Third, this approach would make both the practical implementation and interpretation of this method much more complex. Fourth, as discussed in the next section, we find no gender disparity in how the average leaver compares to the average stayer at the Bank, when we consider their last salary or their accumulated performance ratings.

3 Application: the World Bank Group's headquarters

3.1 Context

We apply the decomposition methodology presented in the previous section to personnel records on employees of the WBG's headquarters in Washington, DC spanning 33,436 individual employees over 28 years. The uncommon length and breadth of the dataset make it ideal to capture how gender pay gaps build up over the course of long careers, the focus of this study.

In 2015, the last year in our data set, the WBG's headquarters exhibited an aggregate pay gap of 22.8%, remarkably close to the gap of 21% found among U.S. employees as a whole.[16] The WBG gap was significantly larger in 1987, the first year in our data (48.6% vs. 31.0%) but the pay gap dynamics are similar to the U.S. labor market's: great reductions in the pay gap in the 80s and 90s (-10.9pp and -7.6pp respectively), followed by slower convergence in the 2000s. While the gap plateaued in the U.S. around 2005, the slow-down took place more gradually at the WBG which saw significant convergence until 2010. One of the main explanatory trends identified in Blau and Kahn (2017) for the U.S. is the increased representation of women at high-paying technical and managerial positions. We describe a very similar evolution in the context of our data over the same period in section 3.4.

The comparable evolution of the gender pay gap at the WBG's headquarters and in the U.S. took place despite some specific features exhibited by the organization. Most prominently, employees of the the Washington, DC office come from 60 countries, and U.S. citizens are a significant minority among hires over our data period (Table 1). Special visa arrangements with the U.S. government allow the WBG to hire and bring in staff to central headquarters from all around the world. Visa considerations hinder the access of non-U.S. citizen staff to the local U.S. labor market and therefore are likely to affect their exit decisions.

Despite some unusual characteristics, the WBG personnel data exhibit the classic features of an internal labor market described by Doeringer and Piore (2020), summarized in Waldman (2012) and documented in studies of other large firms such as Baker et al. (1994*a*) and Baker et al. (1994*b*). These features include: job ladders, "ports of entry", promotions from within, and wages attached to jobs which we discuss in turn below.

[16]OECD (2021), Gender wage gap (indicator). doi: 10.1787/7cee77aa-en (Accessed on 23 September 2021)

Table 1: Nationality of International Hires at the WBG

	1987-1995		1996-2005		2006-2015	
	N	%	N	%	N	%
US	3,179	34.33	3,428	32.56	2,931	24.75
India	633	6.84	603	5.73	745	6.29
Great Britain	429	4.63	419	3.98	427	3.61
France	373	4.03	416	3.95	486	4.1
Other	4,279	46.2	5,663	53.78	6,879	58.08

Table entries correspond to the number and proportion of new staff hired from each country in each period.

The job ladder at the WBG is denoted by grades that run from GA to GL (the president of the WBG), and is represented in figure 1. Grades GA–GD are the grade levels for administrative and client support (ACS) staff. GE corresponds to analyst level. GF and GG contain the bulk of professional technical staff. For expositional purposes, in the rest of the paper we will use "support staff" to refer to grades GA through GD; "technical staff" for grades GE, GF, and GG and non-manager GH; and "managerial staff" for managers in grades GH and above, sometimes separating GI and above grades under the denomination "senior managers".[17] For each gender, figure 1 plots the proportions of external hires and internal promotions among staff entering a given grade. For example, 22% of new female GF staff were promoted internally from GE, and 78% were hired externally. The figure also shows the proportion of all hires who start at each given grade. For instance, external hires at GF correspond to 28% of all female external hires. While new staff can and do enter at any grade, 83% of external hires are made at grades GB, GE, GF and GG, which can

[17] In addition to staff, the WBG has "unassigned or ungraded" employees who are composed of long and short-term consultants as well as a small number of staff outside the salary and promotion structure of the WBG, such as the WBG's executive directors, who are representatives of the WBG's member countries, and their advisers.

be thought of as the "ports of entry" into the organization. These grades also see a larger inflow of external hires compared to other grades, which are instead filled by a majority of promotions from within (see figure 1).[18]

To hire staff externally, a hiring manager must draft terms of reference (TORs) that include the specific hiring grade of the job (grade ranges are not allowed), as well as the minimum experience and education required.[19] The adequacy of the grade and job description must be approved by Human Resources before the position is advertised. Therefore, male or female candidates cannot negotiate a grade increase during the hiring process as described by HR procedures.[20]

After a candidate is selected, the salary is decided through an HR process that proposes an initial salary within a band (depending on education and experience) that is then revised following a negotiation with the applicant. The higher the salary within the band, the more the sign-offs required, with top salaries within the bank requiring vice-presidential approval. We note though that the bands are quite wide and allow salaries to differ by 30% around the band's midpoint. Therefore, male or female candidates have significant room to negotiate their salary.

The WBG is similar to other firms studied in the personnel literature in that salary growth mirrors the progression through grades or positions. As in these other contexts, salaries also increase outside of grade promotions (Baker and Holmstrom (1995)). However, the salary "bands" described above are such that salary increases, for a given performance rating, get smaller in the upper parts of a grade's salary

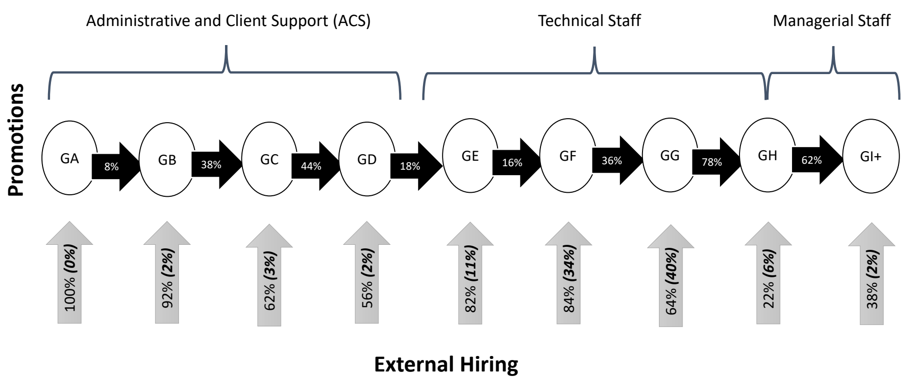
[18]By contrast, exits happen at all grades, which is also a feature shared by the other firms described in the literature, as noted by Baker and Holmstrom (1995) among others.

[19]The correspondence between grades, job descriptions and experience and education requirement is decribed in internal guidelines and the official Staff Rule 04.01.

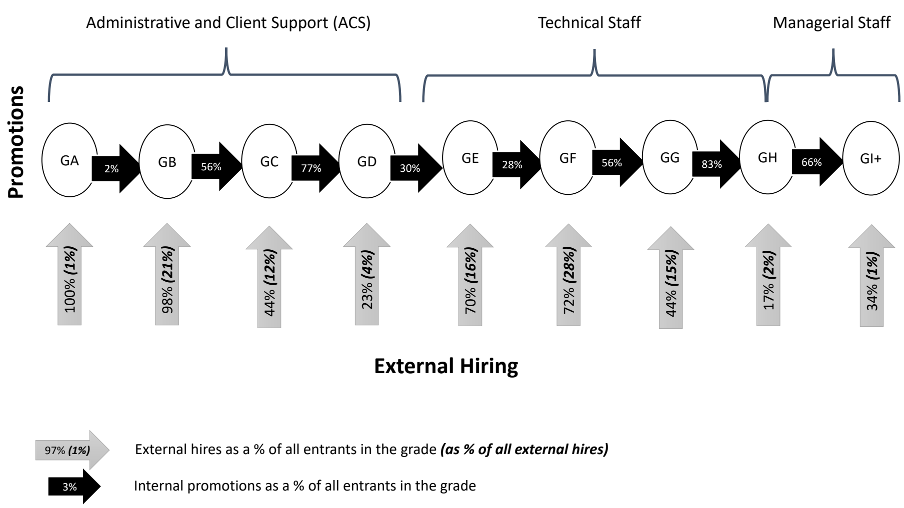
[20]It conceivable that these procedures could be circumvented, if a manager has a person in mind for a job, and negotiates a grade informally before launching the formal hiring procedure. Since this runs contrary to explicit rules of the organization, it is difficult to assess the extent to which this may happen.

Figure 1: The job ladder: external hires and promotions by grade

(a) Male staff



(b) Female staff



The figure plots the proportions of external hires and internal promotions among staff entering a given grade. The bold number in parentheses correspond to the proportion of all hires who start at each given grade. For example 22% of new female GFs were promoted internally from GE (or exceptionally, demoted from a higher grade), and 78% were hired externally. External hires at GF correspond to 28% of all female external hires band. Therefore, wages are attached to grades in the sense that salary progression will be significantly hindered and eventually stalled in the absence of a promotion. As one simple statistic 87% of the variation in salaries in our data are explained just by grade fixed-effects.

Currently, promotions at the World Bank are handled within the department until grade GG, after which a competitive process with multiple applicants is required. While the exact promotion rules have changed over the duration of our data, what has not changed is that a promotion commands an immediate salary increase. In addition, after a promotion, a staff's salary will find itself in a lower segment of the new grade's salary band (relative to the old grade's band), allowing future raises to be larger.

While our application to the WBG remains a case study, our methodology is designed to be portable to other personnel records in the hope that evidence on the role of pipeline effects and careers on gender pay gaps will accumulate from different contexts. The method requires few observable characteristics and is relatively simple conceptually and numerically. In and of itself, the WBG case provides valuable complementary evidence to the overwhelmingly US and Europe-centric body of gender pay gap facts gathered in the literature, allowing us to gauge the differences and similarities in gender pay gap dynamics coming from a regionally and culturally diverse workforce that includes representatives of many nationalities at all levels of skill and responsibility.

3.2 Data and Sample

The WBG's Human Resource Longitudinal Database covers all staff employed by the WBG between 1987 and 2015 and was assembled for the purpose of this study. It is structured in a panel from 1987 to 2015 with staff uniquely identified through a universal, permanent personnel identifier (UPI). Data from each year are a snapshot taken on June 30 that contains information on the staff's universal personnel identifier (UPI), compensation and benefits (for instance, salaries), personal background (gender, age), professional situation (for instance, professional grade), location (HQ or country-office based), role and movements within the organization (promotions and lateral moves), and information on the yearly performance rating (SRIs) from 2000 onward.[21] For the purpose of applying our decomposition, we consider the annual monetary compensation of each employee, the grade at which they were hired, their gender and nationality.[22] The WBG also distinguishes between employees coming from countries that contribute to the World Bank, typically high-income countries ("Part 1") versus those who are borrowers, typically low- and middle-income countries ("Part 2").[23]

The WBG has more than 100 country offices, each of which hires most of their staff in a local labor market where salary opportunities available to each gender, and the WBG's market power, differ. To obtain a homogeneous sample, we restrict our analysis to staff in the Washington headquarters hired on a U.S. dollar salary plan, commonly known as "internationally recruited staff" (62.2 percent of all WBG hires in 2015).[24]

We make two additional sample restrictions. Among the 259,618 records corre-

[21] The Technical Appendix and codebook provide further details.

[22] Non-monetary benefits (such as health insurance) are substantial at the WBG, but homogeneous within the sample we consider, particularly across genders (described below).

[23] This terminology comes from the way the WBG classifies its country members. Part 2 countries joined the WBG as potential loan recipients and thus roughly correspond to low- and middle-income countries.

[24] In addition, substantial country-specific expertise is required to convert local salaries to dollar equivalents. Given the starting date of 1987 in our data, the dissolution of the Soviet Union and the emergence of local currencies, the creation of the euro replacing European currencies, as well as multiple spells of hyperinflation in countries ranging from Turkey to Ecuador, during which country office salaries often switched currency denominations, all need to be addressed on a case-by-case basis.

sponding to the universe of WBG headquarters employees between 1987 and 2015, we exclude the 4,689 records of executive directors and their staff, and of secondments (staff loaned and paid by other institutions). We also exclude a total of 229 records for which missing or anomalous grades were found (6 records), recorded gender changed over time (97 records), recorded salary was 0 (88 records), or recorded salary was clearly outside the grade range in the corresponding year (38 records).

3.3 Three Relevant Historical Trends

Compensation and hiring practices have not remained static over the period of our data. In fact, multiple institutional changes and HR policies could have affected hiring and turnover as well as the salary structure. One advantage of our dynamic decomposition methodology is that it can account for these changes, allowing hiring patterns, salary increases and exit rates to be different in each year of the panel. This is in contrast to a classic Oaxaca-Blinder decomposition (see Filmer et al. (2005), for an application to the WBG) in which, for instance, returns to experience are captured by a time-invariant parameter. Because this feature is likely to characterize most organization over the course of an employee's career, and is a key motivation for our methodology, we present information on three important institutional dynamics.

First, the period covered by our data exhibits large changes in the staff's grade composition. Figure 2 shows that over the period of our data, there was an increase in the fraction of technical and managerial level staff as a fraction of total staff from 64 percent in 1987 to 85 percent in 2015. This increase is consistent with both increasing automation of routine tasks and shifting of routine tasks from support staff to technical and managerial staff. The proportional increase in technical and managerial staff was primarily in the technical grades of GE, GF, and GH; no change was seen in the proportion of managers to staff between the years of 2000 and 2015. The decline in support staff, the majority of whom are women, will automatically reduce the wage gap in the organization as the workforce becomes more homogenized. At the firm level, this is an important force as several studies have shown how firms strategically use contractual workers or outsourcing to address regulatory requirements, recentering the question of the boundaries of the firm (Bertrand et al. (2021)).

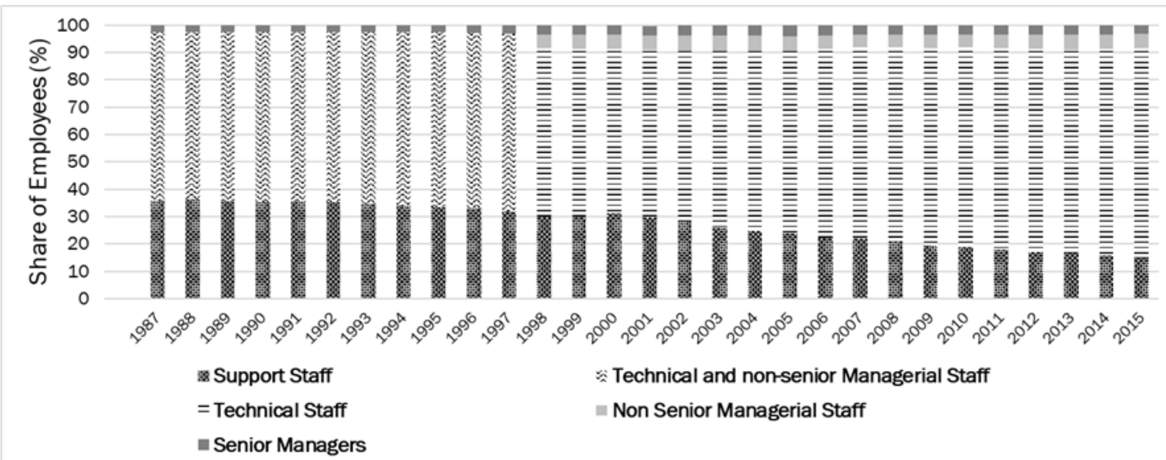
A second feature of the data is that exit rates are not stationary. Figure 3 plots annual exits from the institution as a percentage of regular staff, excluding staff exits due to mandatory retirement. Exits at the WBG are an average of 9 percent a year but have fluctuated in cycles between 6 and 11 percent following large institutional reforms. Exits peak in reform years (which are usually associated with new presidential terms) but then drop because reforms bring exits "forward." It also appears that exits track economic performance in the United States, rising when the economy is strong. Note that the 9 percent exit rate implies that 50 percent of staff leave the institution every 8 years. When we consider long tenures, high rates of attrition leave substantial room for selection effects in the salary gaps of those employees who remain at the institution. Because our methodology only allows exit selection based on the last salary (rather that predicted future career paths), this feature must be kept in mind in the result section when interpreting counterfactuals involving exit patterns.

A third feature of the data is that salaries have not grown at the same pace in all occupations. Salaries in higher grades exhibit much larger progressions over the span of our data, mimicking similar trends in the US economy over the period.

Table 2 shows the mean real salaries of employees at each grade over time. We compare all salaries to a base of 100 for grade GA in 1987.[25] Note the considerable variation in salaries within each grade. Typically, the 10th percentile and the 90th percentile of the within-grade salary distribution differ by 20 to 40 percent. Even

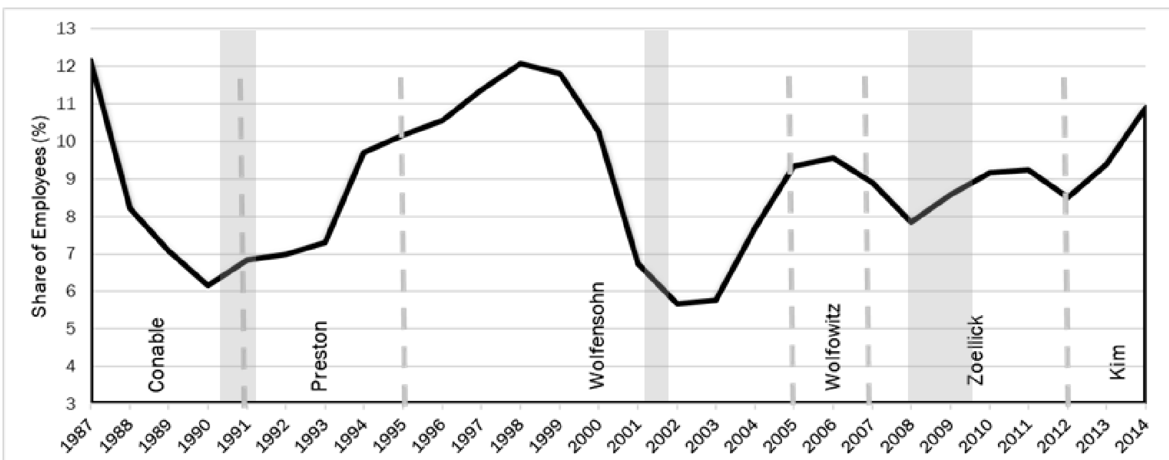
[25] To preserve anonymity, we leave blank the cells where there are too few employees and do not present results for grade GK.

Figure 2: Proportion of Support Staff versus Technical Staff versus Managerial Staff (1987–2015)



The figure shows the fraction of WBG employees that are in Administrative and Client Support positions (grades GA-GD), technical staff (grades GE-GG and some of grade GH), non-senior management (some of grade GH) and senior management (grades GI and above). The managerial status of GH staff is only available in our data after year 1998.

Figure 3: Annual Exit Rates from the WBG (1987–2015) and WBG presidents' tenures



The figure shows the number of WBG employees that exit the institution as a fraction of all staff in each given year who have not reached mandatory retirement age, as they overlap with the tenures of WBG presidents (dotted lines) and economic recessions in the U.S. (shaded areas).

within technical grades, the mean real salaries have increased more for grades GI, GJ and GK compared to GF and GG staff. Annual real salaries increased by three-tenths of 1 percent between 1987 and 2014 among GB–GD and GG staff, seven-tenths of 1 percent for GE and GH level staff, and 1.1 to 1.6 percent for GI–GK level staff.

Each of these trends over the period of our data can affect the salary gaps between subgroups. For instance, the decline in support staff, who are predominantly women in jobs with lower salaries, implies that the average salary of women relative to men will rise in the institution. Similarly, differences in the profiles of staff leaving the WBG will affect the salaries of those who remain. Finally, differential increases in salaries for different grades can affect aggregate gaps. First, as support staff tend to be women, their lower salary growth over time will imply that the aggregate gap will also increase. Second, staff are promoted over time. If men are promoted faster to GH (for instance) relative to women and GH salaries are growing faster, this will again induce an increase in the aggregate gap over time.

3.4 Gender Differences in Hires, Salary Growth, and Exits

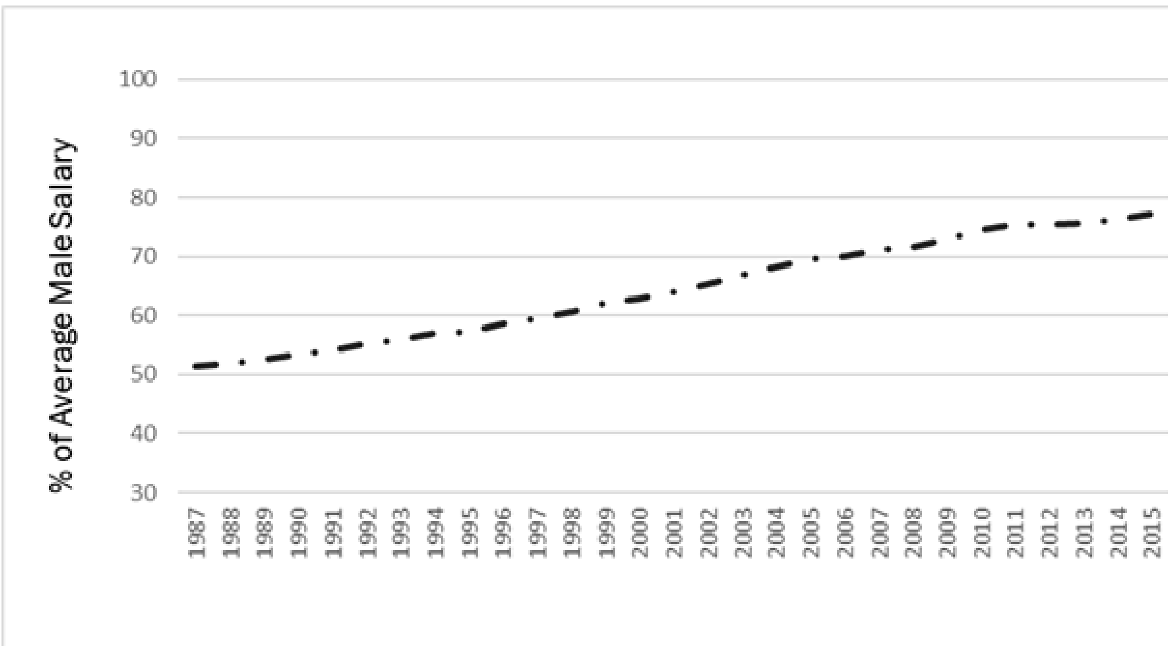
In 1987, the mean salary of a female staff member at the WBG was 52 percent that of a male staffer. By 2015, this had increased to 77 percent (Figure 4). As a point of comparison, pay gaps between gender are about twice the size of pay gaps between employees from high-income countries (denoted as "Part 1") versus middle- and low-income countries (denoted as "Part 2"). Figure 5 interacts gender and country of origin: a clear ordering emerges with the highest salaries for Part 1 males throughout the period of our data. The differences between men and women within country groups are about twice as large as the differences between Part 1 and Part 2 staff, within gender groups. Contrary to the gender pay gap, the nationality pay gap has not significantly declined over the period.

Table 2: Salaries by Grade and over time at the WBG

		Fiscal Year					
Grade		1987-90	1991-95	1996-00	2001-05	2006-10	2011-15
GA	Mean Salary	100	103	96	100	101	99
	Number of Staff	122	62	72	74	34	8
GB	Mean Salary	112	116	116	118	121	119
	Number of Staff	2,600	3,099	1,828	1,068	381	253
GC	Mean Salary	141	150	148	149	154	153
	Number of Staff	4,859	6,212	6,159	5,156	4,155	3,718
GD	Mean Salary	171	185	185	182	188	189
	Number of Staff	1,093	2,088	2,478	3,642	3,506	3,084
GE	Mean Salary	193	218	217	214	219	220
	Number of Staff	2,618	2,719	2,917	4,000	3,917	4,007
GF	Mean Salary	256	275	273	278	282	284
	Number of Staff	1,822	2,944	3,522	6,017	7,084	8,615
GG	Mean Salary	362	378	372	377	389	390
	Number of Staff	7,644	10,746	10,323	9,973	11,813	13,842
GH	Mean Salary	457	492	484	501	534	539
	Number of Staff	2,901	4,428	5,407	6,161	6,741	8,008
GI	Mean Salary	547	610	610	644	698	712
	Number of Staff	548	743	920	1,168	1,160	1,224
GJ	Mean Salary	619	706	704	782	851	877
	Number of Staff	69	100	160	188	152	163

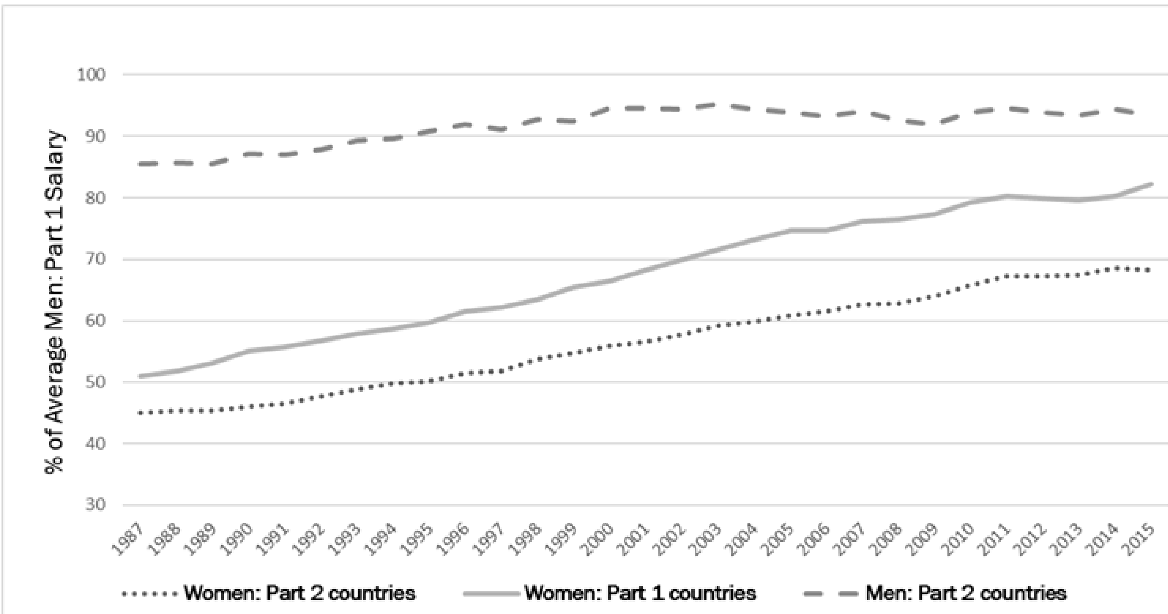
Table entries correspond to the number and average salary of staff in different grades in different years, relative to the average salary of a GA staff in 1987-1990 (normalized at 100).

Figure 4: The Aggregate Gap: Women's Mean Salaries Over Time, Relative to Men's



The figure plots the average salary of female staff as a percentage of the average salary of male staff from 1987 to 2015. For example, in 1996, the average female staff earned 60% as much as the average male staff.

Figure 5: The Aggregate Gap: Mean Salaries by Subgroups Over Time, Relative to Men from Part 1 Countries



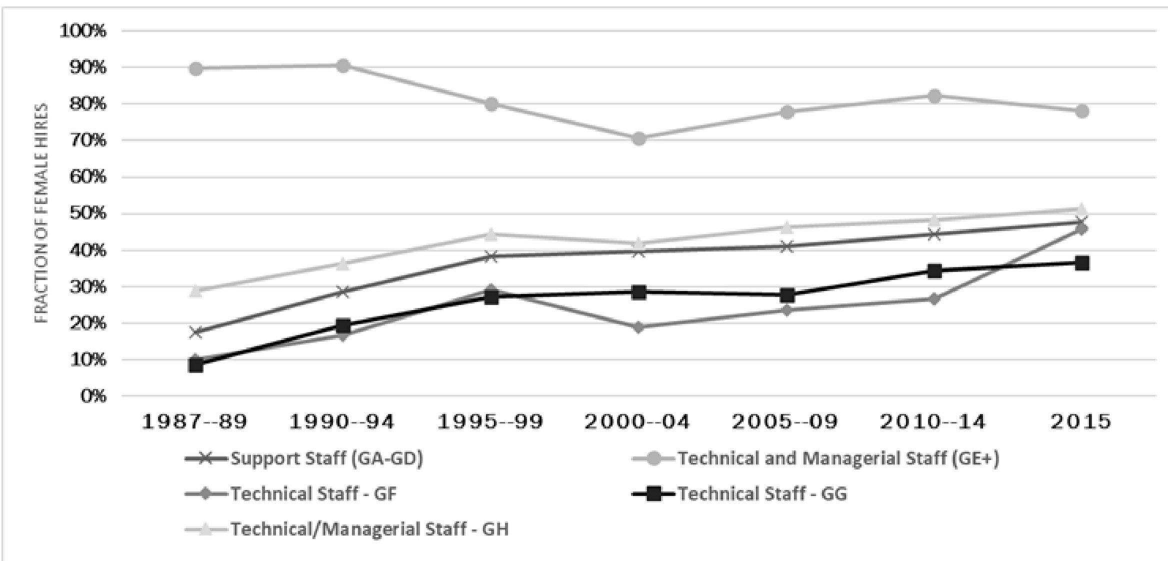
The figure plots the average salary of staff from three groups (Women from Part 1 and Part 2 countries and Men from Part 2 countries, as a percentage of the average salary of male staff from Part 1 countries from 1987 to 2015. Part 1 countries joined the WBG as lenders and are typically high income. Part 2 countries joined the WBG as borrowers and are typically low or middle income. For example, in 2003, the average female staff from Part 2 countries earned 60% as much as the average male staff from a Part 1 country.

Occupational segregation among new staff at the WBG has declined significantly. In 1988, women comprised 20 percent of all staff hired into technical and managerial positions, and this number increased to 48 percent by 2015 (Figure 6). This strong convergence toward parity is not fully completed for mid-career hires (GG grade), who were still more than 60 percent male in 2015. On the other hand, among ACS staff hires , women remain the predominant gender to this day, decreasing in share from 92 percent to 78 percent. This evolution runs paralell to change in the U.S. labor market, where the gap in the incidence of traditionnally male managerial and professional jobs among women and men had virtually disappeared by 2011 (Blau and Kahn (2017)).

In contrast to the large differences in hiring across grades, entry salary differences by gender are relatively modest within a grade. Figure 7 presents the relative entry salaries of women and men at the main entry grades (male entry salary are normalized as the constant line). The largest gap is found in mid-level technical and managerial staff hires (grade GG). Female employees enter at a salary deficit that fluctuates around US$3,000 , but this category of hires is less homogeneous, as it requires various levels of previous work experience. For entry-level technical staff, the gap is modest (GF, around US$500) or absent (GE). For support staff, a modest entry salary gap opened up in the middle of the period but disappeared by 2015.

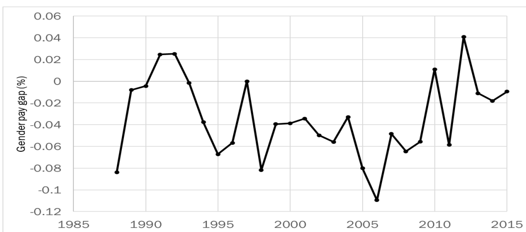
Relatively to men, women's salary paths show clear declines over years of tenure for most –but not all– entry positions. Figure 8 reports salary paths for the three most common entry grades at the WBG. The graphics aggregate staff according to their tenure, mixing cohorts who joined the institution in different years. The female deficit in annual salary after 15 years is US$2,500 for GB, US$2,000 for GF and non-existent for GG. The evolution of salaries is heavily influenced by the speed and frequency of promotions, but differences in performance ratings are also rewarded

Figure 6: Fraction of Women among New Hires at the WBG, 1987–2015, by Entry Grade

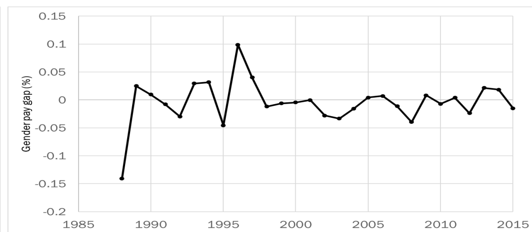


The figure plots the proportion of women among staff hired at different grades from 1987 to 2015. For example, between 1995 and 1999, 80% of employees hired at grades GA-GD were female.

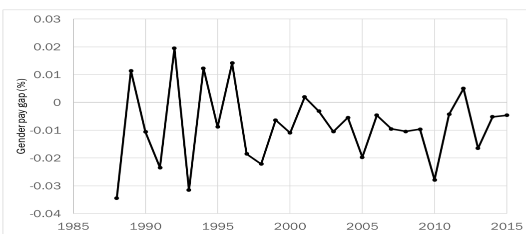
Figure 7: Women's Entry Salaries by Grade and Year of Entry, Selected Grades, 1987–2015, Relative to Men's



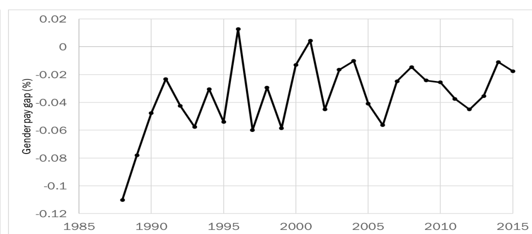
(a) Support Staff (GA–GD)



(b) Technical Staff (GE)



(c) Technical Staff (GF)



(d) Technical Staff (GG)

The figure plots the relative difference between the average salary of new female staff and the average salary of new male staff from 1987 to 2015, by entry grade. For example, in 2000, the average new female GF hire earned 1% less than the average new male GF hire.

with higher raises in non promotion years.

The salary paths described above will be affected by the gradual attrition of staff over their careers. In fact, all grades combined, less than 60 percent of staff remain at the WBG 15 years or longer (Figure 9). Women are more likely to stay than men in all categories, but the differences are less than five percentage points everywhere. Importantly for the validity of our exercise, we do not find gender disparities in how leavers compare to stayers on average. Leavers of both genders seem to have slightly higher salaries (Table 3) and slightly lower performance ratings than stayers (Table 4).

4 Implementation of the decomposition

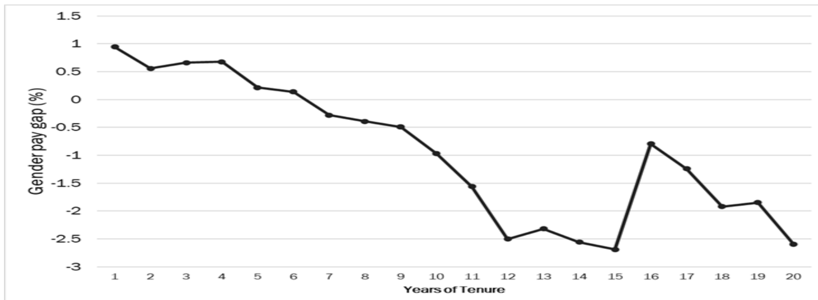
4.1 Microsimulation Algorithm

The microsimulation algorithm we designed takes as inputs the incumbent salary distribution at t_0 and distributions of entry salaries, salary increases and exit probabilities for the different groups of employees in each year. Section 4.2 describes how we model and estimate these input distributions from the data. We simulate salaries for each group of employees from the first year when data are available until the year in which the pay gap of interest is measured. The algorithm updates salaries from year $t-1$ to year t in the following way: [26]

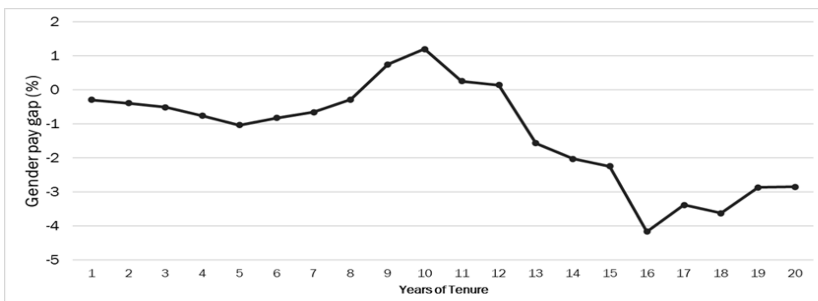
1. Store the salaries carried over from $t-1$ as a histogram with N bins. In our

[26] The idea for microsimulations dates back to Orcutt (1957) and modern applications are reviewed in O'Donoghue (2014). The primary use of this technique is to simulate distributional effects of public policies but it is also employed in the context of simulation-based estimation of structural microeconomic models and in some cases to perform decompositions (see Keane and Wolpin (2010) for such a decomposition of black-white differences in women's economic decisions, or Joubert (2015) for a decomposition of the causes of informality). We are not aware of microsimulations tools applied to modeling an individual firm's workforce.

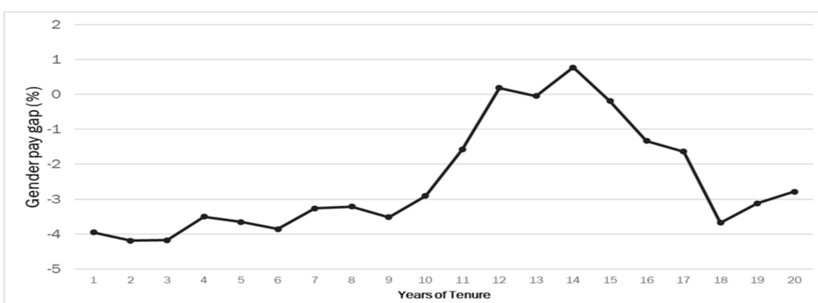
Figure 8: Women's salary paths by Grade and Year of Entry, Selected Grades, 1987–2015, Relative to Men's



(a) Staff hired at GB level (31.9% of hires)



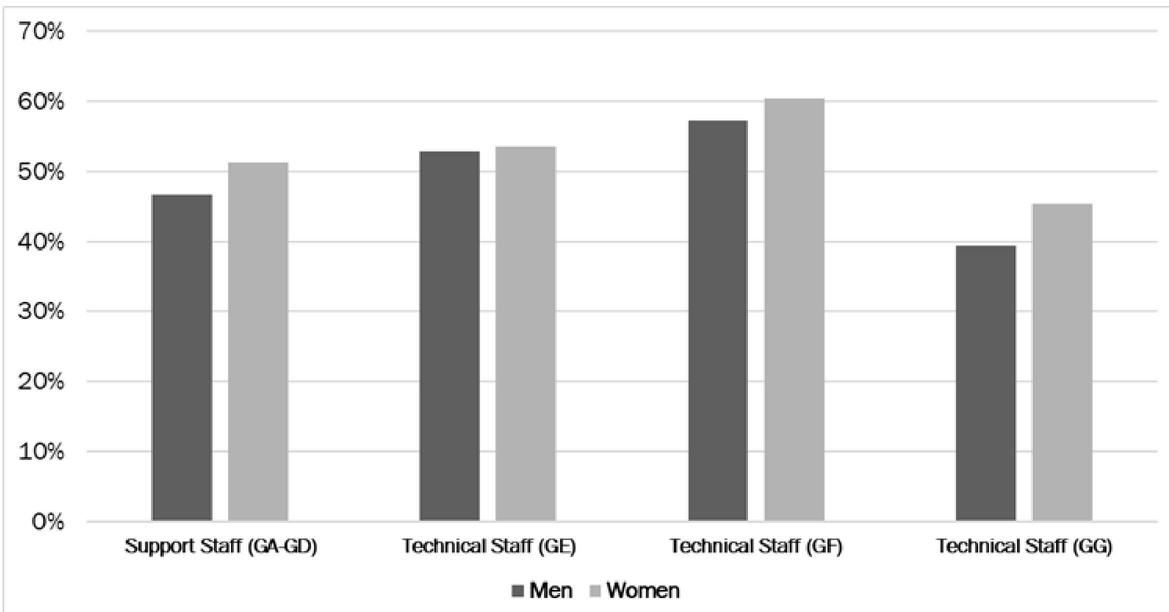
(b) Staff hired at GF level (23.5% of hires)



(c) Staff hired at GG level (33.3% of hires)

The figure plots the relative difference between the average salary of female staff and the average salary of new male staff by years of WBG tenure and entry grade. The sample corresponds to all staff hired at each entry grade between 1987 and 1995. For example, women hired at grade GF who were still in the organization after 14 years of tenure earned 2% less on average than their male peers.

Figure 9: Fraction of Staff Who Remain at the WBG After 15 Years, by entry grade



The figure plots the fraction of staff still employed at the WBG 15 years after they were hired, for each hiring grade. The sample consists of all staff hired at each entry grade between 1987 and 2000.

Table 3: Salary Ratios of Leavers over Stayers by Gender

Hiring Grade	**Tenure**	**Gender**			
		Male		**Female**	
		Ratio	*sd*	Ratio	*sd*
GA-GD					
	5	0.99	*0.08*	1.03	*0.07*
	10	0.98	*0.10*	1.03	*0.08*
	15	1.01	*0.17*	1.02	*0.10*
GE					
	5	0.98	*0.04*	1.00	*0.03*
	10	1.02	*0.06*	1.00	*0.05*
	15	1.03	*0.10*	1.03	*0.08*
GF					
	5	1.00	*0.03*	1.00	*0.02*
	10	1.00	*0.04*	0.98	*0.04*
	15	0.99	*0.06*	1.00	*0.06*
GG					
	5	1.02	*0.03*	1.03	*0.03*
	10	1.01	*0.03*	1.02	*0.04*
	15	1.01	*0.05*	1.02	*0.06*

Table entries are ratios of the average salaries of leavers over stayers for each group defined by an entry grade, number of years of tenure and gender. The standard deviations of the ratio x/y is approximated by $(E(x)/E(y))^2 + (Var(x)/E(x)^2 + Var(y)/E(y)^2)$

Table 4: Performance Ratings Ratios of Leavers over Stayers by Gender

Hiring Grade	**Tenure**	**Gender**			
		Male		**Female**	
		Ratio	*sd*	Ratio	*sd*
GA-GD					
	5	0.97	*0.04*	0.97	*0.04*
	10	0.97	*0.01*	0.97	*0.02*
	15	1.00	*0.01*	1.00	*0.01*
GE					
	5	0.97	*0.04*	0.97	*0.05*
	10	0.97	*0.01*	0.97	*0.01*
	15	1.00	*0.01*	1.00	*0.01*
GF					
	5	0.97	*0.05*	0.97	*0.05*
	10	1.00	*0.02*	0.97	*0.01*
	15	0.97	*0.01*	0.97	*0.01*
GG					
	5	0.97	*0.05*	0.97	*0.04*
	10	0.97	*0.01*	1.00	*0.02*
	15	1.00	*0.01*	0.97	*0.01*

Table entries are ratios of the average accumulated performance ratings of leavers over stayers for each group defined by an entry grade, number of years of tenure and gender. The standard deviations of the ratio x/y is approximated by $(E(x)/E(y))^2 + (Var(x)/E(x)^2 + Var(y)/E(y)^2)$

application, $N = 500$ and each bin represents USD 1,000 in annual salaries (narrower bins increased run time but did not improve our ability to fit the data significantly).

2. For each bin in the histogram, and for each employee in the bin, draw a salary increase for year t from the input distribution estimated from the data.[27] Then determine the new bin to which each individual belongs after the salary raise.

3. Add entry salaries to the post-raises salary histogram obtained in the previous step. Entry salaries are obtained by drawing n times from the input entry salary distribution estimated from the data, where n is the actual number of entrants in year t from that group.

4. Determine for each bin of the resulting histogram (i.e. post-raises and post-entries) how many employees exit the sample at the end of t.[28]

5. Go to $t + 1$.

4.2 Parametric assumptions

The simulation algorithm can be implemented with various levels of parametrization, depending on the amount of observations in the data set and the statistical regularity of the firm's processes. Our proposed specification attempts to balance the stability, simplicity, and ease of implementation of the decomposition - by incorporating some parametric assumptions - against the risk of specification error. The

[27]Specifically, we recreate a continuous distribution by drawing a salary from a uniform distribution over the support of the bin before applying the salary raise. An alternative is to apply the raises to the midpoint salary w_m in the bin. In our application, the bins are very narrow so that the two procedures are essentially equivalent

[28]In practice, we draw a random number for each employee and compare it to their probability of exiting. That is, the actual number of exits is matched in expectation only.

stability of the decomposition is assessed by obtaining bootstrapped confidence intervals around the estimated decomposition factors (see section 4.3). We propose to assess the joint impact of specification error directly on the output of the simulation algorithm through a validation exercise described in section 4.3. We also repeat the exercise using alternative specifications in addition to our preferred one.

We stratify the estimation by year and group of employee, where groups are defined by gender, nationality groups and entry position.[29] Therefore, we do not impose any restrictions on the way the processes change over time or how they differ across groups of employees. Also, the number of entrants for each group g (now including nationality and entry position, in addition to gender) and year t, denoted as n^h_{gt}, is directly taken from the data.

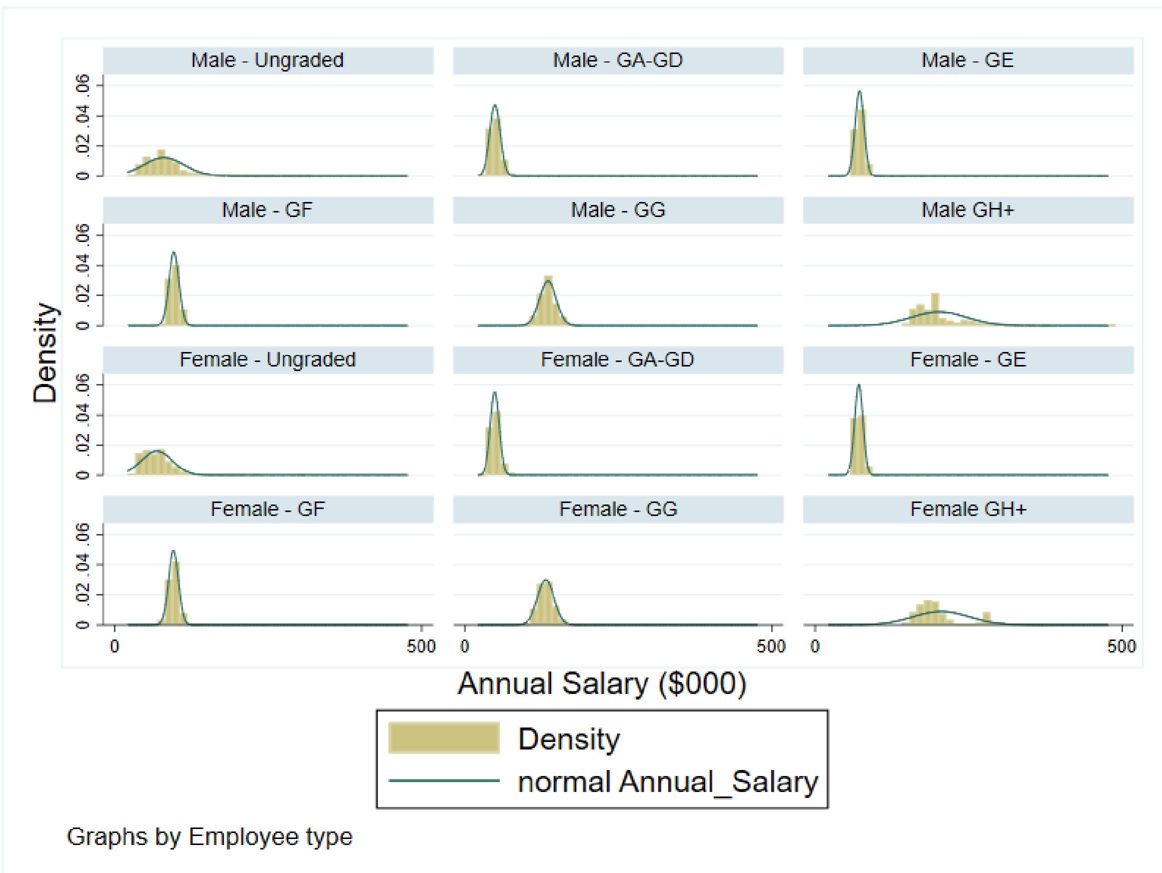
For a given position and gender, the salary of a new hire, w^h_{igt}, is modeled as a normally distributed random variable with mean μ^h_{gt} and standard deviation σ^h_{gt}. Figure 10 shows that the distributions of entry salaries, aggregated over nationality groups, appear well approximated by a normal distribution for all groups with adequate sample sizes.[30] While a formal Shapiro-Wilkins test rejects normality for a majority of the cells, the normal distribution is an efficient way to summarize our data such that a relatively small set of numbers need to be passed on to the simulation algorithm. We show in section 4.3 that the misspecification error generated by the use of a normal distribution does not preclude a good approximation of the average gender pay gap.

If our microsimulations had failed to achieve a satisfactory level of fit using

[29]Nationalities are grouped according to the World Bank classification "Part 1" and "Part 2" described in section 3.

[30]External hires at the very top positions of the institution (grades GH and above) are not as well approximated, due to the small cell sizes. However, note that the bulk of the top management is hired into the institution at lower grades and promoted internally, and thus belongs to the well approximated group.

Figure 10: Distribution of Salaries at Entry by Entry Grade 2010–2015



The figure plots histograms of entry salaries for each grade and gender and a fitted normal distribution. The sample consists of all staff hired between 2010 and 2015 at grades GA through GL.

these parametric assumptions, a non-parametric alternative would have been to directly import from the data histograms of entry salaries for each the 336 groups of employees * years. A disadvantage of this approach is that our "composition of hires" counterfactual varies the number of entrants from each group, while keeping the distribution of salaries among them constant. This is easily simulated when using a parametric distribution by randomly drawing the counterfactual number of entry salaries, whereas scaling histograms by a non-integer factor would involve reallocating "fractions of employees" across bins. As for non-parametrically estimated continuous densities (say with a kernel density estimator), we are not aware of convenient ways to "draw" from such distributions, as would be needed in the context of a micro-simulation algorithm.

The choice of a normal distribution may seem unusual because salaries are usually modeled using a log-normal distribution to capture a right-skewness. However, remember that we are estimating salary distributions within cells that are homogeneous with respect to employer, gender, tenure and type of position, which removes many sources of skewness. As a robustness check we redo the exercise assuming that entry salaries follow a log-normal distribution and obtain very similar decomposition results (Table 5, column 3).

Salary decreases are not observed at the WBG but some years saw widespread salary freezes. To capture this feature of the data, salary growth was modeled as a two-step process. With some probability p_{gt}, an employee's salary stays constant in a given year.[31] This probability is estimated by the fraction of 0-increases among group g in year t. Conditional on a strictly positive raise, we model the growth rate of salaries r_{gt} using a log-normal distribution with mean μ^r_{gt} and standard deviation σ^r_{gt}. Figure 11 shows that this parametric assumption is not unreasonable overall.[32]

[31]Salary raises are also set to zero if there is only one employee in a given cell.

[32]We also attempted to estimate a specification in which the raise rate is allowed to depend on the

Figure 11: Distribution of Log-Salary Increases by Entry Grade 2010–2015

The figure plots histograms of log-salary increases for each grade and gender and a fitted normal distribution. The sample consists of all staff hired between 2010 and 2015 at grades GA through GL.

Table 5: Decomposition of the Aggregate Gender Gap: Robustness to Alternative Parametric Specifications

	Baseline	**Probit Exits**	**Log-normal Entry Salaries**
legacy	10.7	10	10.5
salary_growth	5.0	5.2	5.5
retention	0.4	1.8	0.7
entry_salaries	7.0	7.2	7.7
grade_composition	76.7	75.5	75.3

This table presents decomposition results for the aggregate gender pay gap using two alternative specifications. In column 2, exits are modeled using a probit specification, with current salary as the lone regressor. In column 3, entry salaries are assumed to follow a log-normal distribution.

Finally, exits are determined by a linear probability model in which the probability of exiting the sample is allowed to depend on the current salary:[33]

$$1_{(exit_{it}=1/t,g,w_{it})} = \alpha^{l}_{gt} + \beta^{l}_{gt} w_{it} + \varepsilon^{l}_{gt} \tag{4.1}$$

The linear probability model has the advantage of a significantly faster simulation run time compared to a non-linear probability model without compromising our ability to fit the aggregate gender pay gap. However, we generated our main result once using a probit model for exits and obtained almost identical decomposition factors (Table 5, column 2).

salary level, but the coefficient on that additional regressor was not tightly estimated for all cells. The resulting simulated raises led to some improbably large salary increases and a worse overall simulation fit than in our preferred specification.

[33] For cells with fewer than 20 employees, or fewer than 5 exits, β^{l}_{gt} was set to 0.

4.3 Estimation and validation

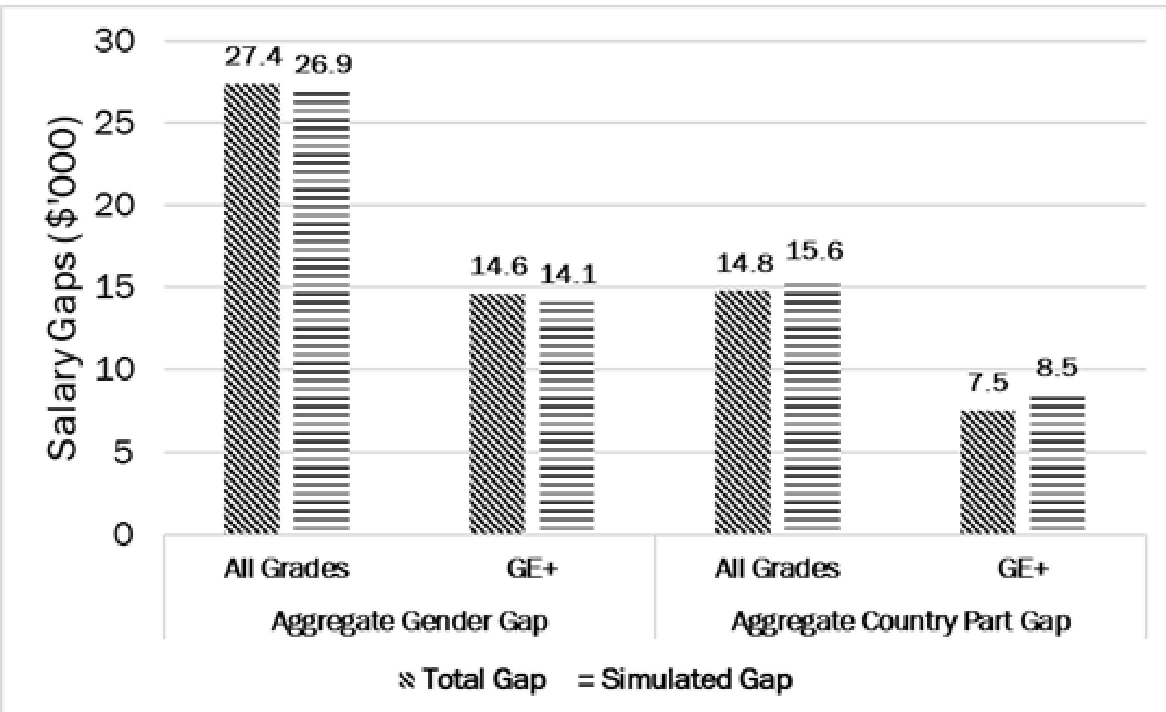
In our simulation algorithm, all dynamic linkages operate through current salaries, which are observable. As a result, we can separately estimate all the input processes for entry salaries, raises and exits before aggregating them into the current salary distribution. We obtain confidence intervals around the decomposition factors, as well as the parameters of the entry, exit and salary growth processes, by bootstrapping the estimation, simulation and decomposition exercise 250 times.[34] The decomposition factors and confidence intervals are reported in section 5.

A benefit of estimating input processes separately is that the end-of-period distributions of salaries are not used in estimation. Therefore, they can be used to jointly validate –out-of-sample– the simulation algorithm and the specifications of the input processes. In particular, our main object of interest –the difference between the average salaries of male and female employees in 2015– is not targeted in estimating our input parameters. We start by verifying that this difference is accurately replicated after simulating salary distributions from 1987 to 2015. In each simulated period, the algorithm potentially combines simulated incumbents with simulated exits, raises, and new entrants. Therefore, the simulated 2015 distribution compounds errors from each of these elements over 25 simulation periods. Figure 12 compares the pay gaps measured in the data in 2015 with the corresponding statistics obtained by simulating men and women's salaries in 2015, using the algorithm described in Section 4.3. Bars 1 and 2, which correspond to the aggregate gender gap measured for all employees, show a simulation error of only US$500 out of a gap of US$27,400. After restricting to technical staff (grades GE and above), the simulation error remains at US$500 out of a gap of US$14,600.

We also compare, for each group of employees, the simulated end-of-period

[34]The confidence intervals stop moving significantly after 100 bootstraps

Figure 12: Actual vs Simulated Salary Gaps



The figure compares pay gaps computed from actual versus simulated salary data for different subgroups. Odd-numbered columns correspond to the data, whereas even-numbered columns correspond to simulations. Columns 1 through 4 plot salary differences between the average male and female staff. Columns 5 through 8 plot salary differences between staff from Part 1 countries and Part 2 countries. Part 1 countries joined the WBG as lenders and are typically high income. Part 2 countries joined the WBG as borrowers and are typically low or middle income. Columns 1, 2, 5 and 6 consider all staff, while Columns 3, 4 and 7, 8 restricts the sample to staff hired at GE or higher.

distributions of salaries with their data counterpart between 1987 and 2015. Figures 13 and 14 present the difference between the mean and the standard deviations of the distribution of salaries of each group of employees in the data and the simulations, respectively. Looking at the means first, we find that simulations can fit the data very well, with no systematic pattern of under- or over-prediction and gaps that are typically less than 1 or 2 percent. Ungraded employees (first column) who are a very heterogeneous group with short tenures and were the object of important reforms in the years 1998-2000, are relatively less well captured. Regarding the standard deviations, the dispersion is slightly higher in the simulations than in the data but discrepancies remain small at around 5 percent on average after running the simulation algorithm for 28 periods (1987 to 2015).

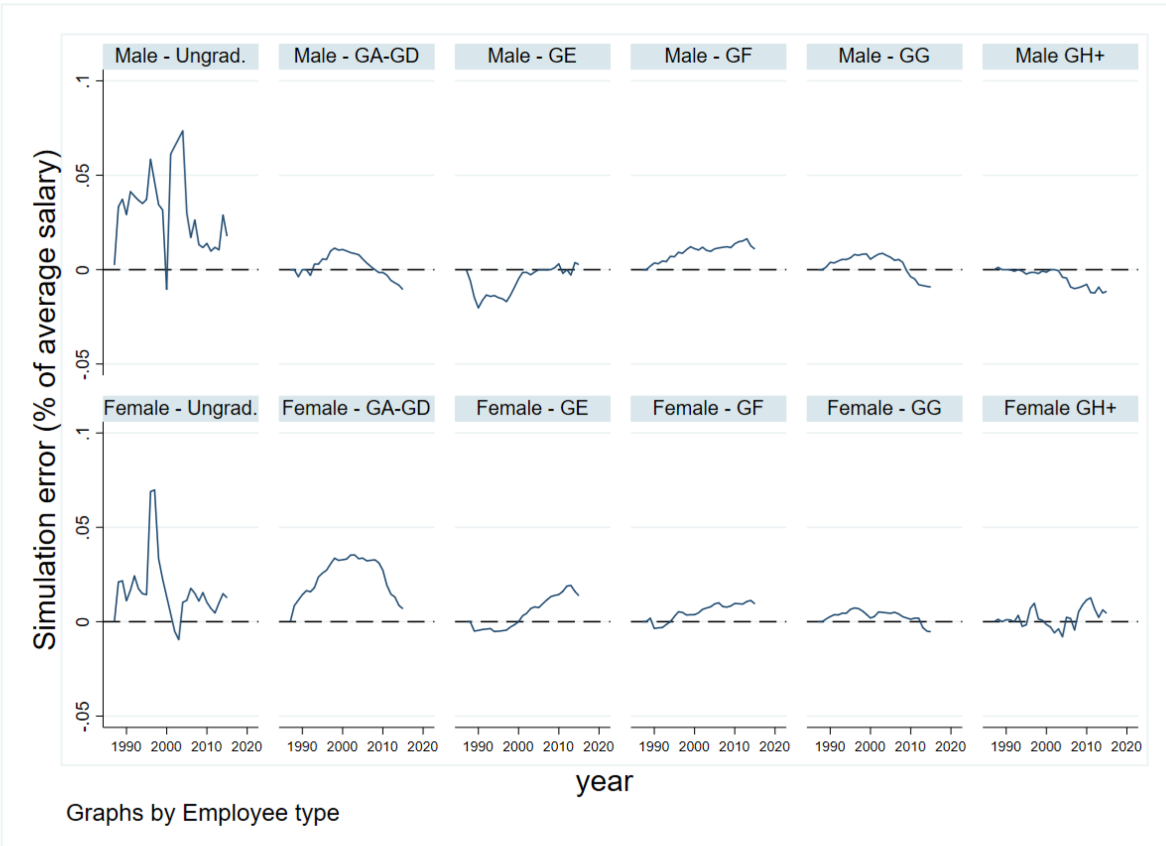
5 Results

5.1 Decomposition of the Aggregate Gender Pay Gap

The main results of the aggregate gender pay gap decomposition are shown in Table 6. The decomposition (including the input estimation and micro-simulations steps) was replicated 250 times on bootstraps of our main sample. The first column of Table 6 contains the mean of the share of the total gap that is explained by each factor. Columns two and three report the confidence interval obtained by computing the 5th and 95th percentiles of the distribution of bootstrapped estimates. Confidence intervals are narrow around the point estimates for all factors.

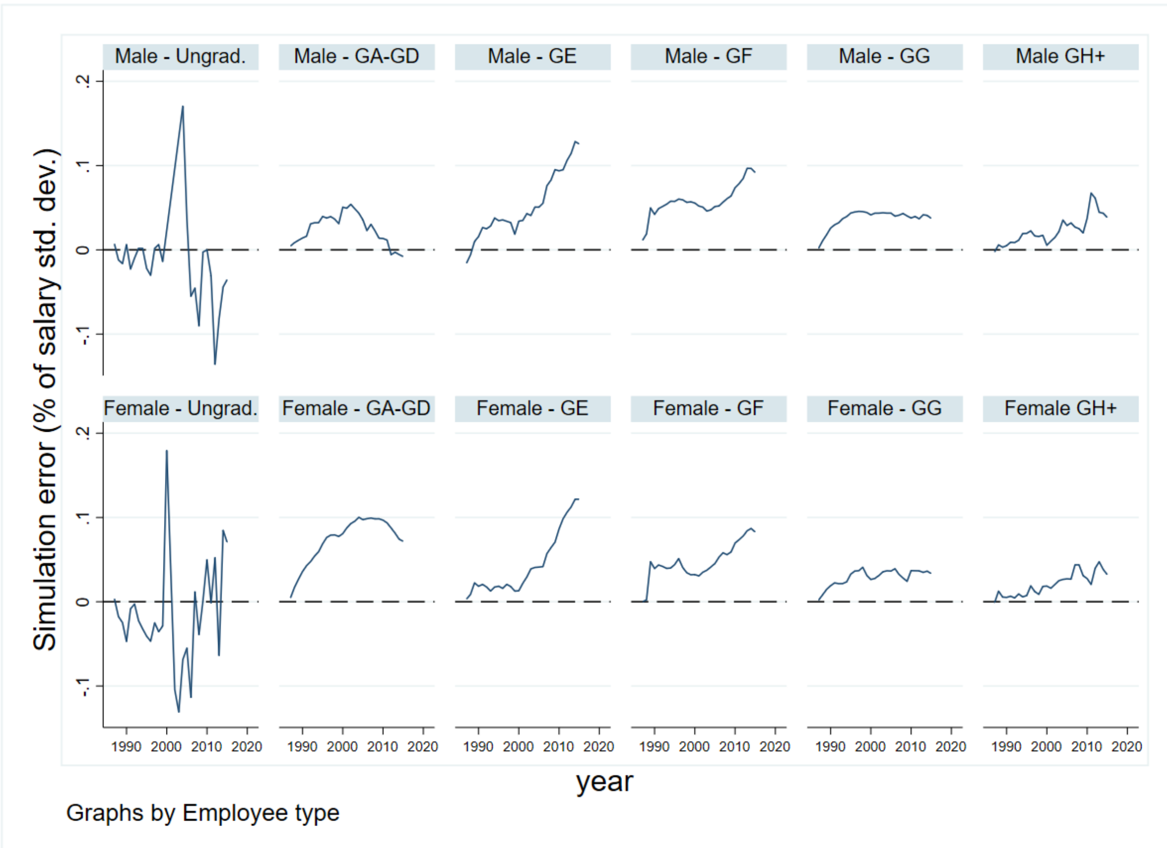
We find that retention plays a small role in explaining the aggregate gender pay gap. When we set the parameters governing the probability, conditional on their salary, with which women exit the institution, in each year and for each entry grade,

Figure 13: Fit of the Simulated Salary averages by Employee Group



The figure shows the percentage difference between simulated and actual average salaries by gender, entry grade and year. For a given subgroup, this corresponds to $\frac{\bar{w}_t^{sim} - \bar{w}_t^{data}}{\bar{w}_t^{data}}$, where $\bar{w}_t^{sim}$ is the average simulated salary for that subgroup in year t and $\bar{w}_t^{data}$ is the average salary for that subgroup in the data in year t.

Figure 14: Fit of the Simulated Salary Standard Deviations by Employee Group



The figure shows the percentage difference between simulated and actual standard deviations of salaries by gender, entry grade and year. For a given subgroup, this corresponds to $\frac{\sigma(w_t^{sim})-\sigma(w_t^{data})}{\sigma(w_t^{data})}$, where $\sigma(w_t^{sim})$ is the standard deviation of simulated salaries for that subgroup in year t and $\sigma(w_t^{data})$ is the standard deviation of salaries for that subgroup in the data in year t.

Table 6: Decomposition of the Aggregate Gender Gap: Bootstrapped Distributions of the Percentage Contribution of each Decomposition Factor.

		All Grades			Technical Staff		
		mean	***p5***	***p95***	**mean**	***p5***	***p95***
legacy		10.7	*9.2*	*12.2*	15.6	*12.5*	*18.7*
salary growth		5.0	*3.5*	*6.5*	6.6	*3.0*	*10.3*
retention		0.4	*-4.7*	*4.5*	-8.8	*-21.0*	*1.5*
entry salaries		7.0	*5.6*	*8.3*	9.6	*7.1*	*12.4*
grade composition		76.7	*72.5*	*81.4*	76.8	*67.9*	*87.0*
Total pay gap	%		22.9			11.3	
	000USD		27.4			14.6	

The table entries are percent contributions for each decomposition factor. The distributions of the factor contributions are estimated by simulating the decomposition using each of 250 bootstraps of the simulation input parameters.

to be equal to that of men, the overall gap diminishes by less than 1 percent. This finding does not imply that exit rates between men and women are identical. Rather, our descriptive results show that, for both genders, the average leaver is similar to the average stayers, in terms of current salaries (Table 3). As a result, exit patterns do not affect the salary distributions of men and women differently in our exercise.

This result relies on our assumption about the salary progressions forgone by leavers. As discussed in Section 4.3, the salary progression a leaver would have experienced (conditional on last salary), had she stayed, is estimated from the salary progressions experienced by stayers. Bias to our results would ensue if that assumption were violated differently for men and for women; for example, if men were more likely than women to take into account idiosyncratic career prospects when deciding to leave the institution.

Salary growth and entry salaries also play a modest role. After giving women the same parameters as men in the salary growth model, the aggregate gender gap decreases by 5 percent. Further equating entry salaries within each entry grade accounts for another 7 percent. The latter is easily reconciled with our descriptive findings: entry salaries exhibit very modest differences across genders for most entry grades. As for salary growth, we do find a significant gap in favor of men for the two main entry grades (GB and GF). However, the gap opens up slowly over the career and thus applies to the relatively small number of employees who stay for a long time at the bank.

The bulk of the gap (77 percent) corresponds to the historical occupational segregation by gender among hires. Section 3.4 shows that women have been disproportionately hired at lower grades, accounting for nearly 80 percent of support staff hired in 2015. Within technical grades (GE and above), a larger proportion of men was hired in the past, but the gap has become much smaller in recent years. It is unlikely, but not impossible, that gender differences in hiring grades reflect in part differences in the willingness or ability to negotiate a higher grade for a given job.[35]

To investigate whether there are hiring composition effects within technical and managerial staff, we implement the decomposition on the sample of employees hired at grades GE and higher. The decomposition factors are shown in columns 4, 5 and 6 of Table 6. The composition of hires remains the most important factor even after excluding ACS staff. That is, the relative mix of entry (GE–GF) versus mid-career (GG–GH) external hires among men and among women generates most of the gap among grades GE and above. Indeed, Figure 6 shows that mid-career hires are still only about 40 percent female. Although this proportion is much higher than it was in 1987, there has also been a strong increase in the proportion of women in entry-

[35]The hiring procedure described in section 3 does not allow the hiring manager any discretion in assigning a grade to a given job, and that grade is set and approved before the hiring manager is able to interview (and therefore observe the gender of) the candidate.

level technical grades (GF). The latter trend has reduced the pay gap of the whole institution but contributed positively to the pay gap among technical and managerial staff.

The remainder of the gap (11 percent) corresponds to differences between the stock of men and women hired before 1987, for whom we do not have entry salary, salary growth, or retention information.

5.2 Comparison to decompositions of the U.S. gender pay gap

While the decomposition factors we propose are original and do not map directly into more traditional decomposition results, some parallels can be drawn with what is known about the gender pay gap in the U.S., as recently summarized in Blau and Kahn (2017). As noted in section 3, the unconditional gender gap for full-time U.S. workers in 2014 was 21%, versus the aggregate gap of 22.8% we measure at the WBG. Conceptually, the WBG's aggregate gap is also unconditional in the sense that it averages over employees of all skill and occupation levels but note that it is mechanically free of industry effects, which accounted for 17.6% of the U.S. gender gap in 2010.

Blau and Kahn (2017) also control for schooling, experience and occupation to explain a total of 68% of the gap. The remaining 32%, or about 6pp, are differences in the returns to these characteristics. The "explained" part of the gap, which is measured by counterfactually equating schooling, job characteristics and experience across genders, bears similarities with our "composition of hires" counterfactual. Entry grades at the WBG are assigned based on the job to be performed, and have specific schooling and prior experience requirements. Therefore, counterfactually equating the number of hires by entry grade across genders is akin to making their schooling attainments and initial occupations similar. In addition, this counterfactual also brings the ratio of new versus more experienced employees within each gender closer together. If we combine this and the "retention" counterfactual, the tenures accumulated at the WB by female and male employees would also be made to be similar. Taken together, these two factors explain 77.1% of the gap, which is of a similar magnitude to the 68% explained part of the U.S. gap mentioned earlier. The "residual gap" can be likened to our entry wage and salary growth factors, which capture the salary returns to initial characteristics and accumulated tenure, respectively.

5.3 Quantifying Pipeline Effects

Pay gaps reflect changes in hiring and compensation patterns with significant lags. The reason is that employees often stay in the same firm for decades, and individual salary paths exhibit high levels of auto-correlation. This inertia makes it difficult to accurately diagnose the causes of pay differentials, and is indeed one of the main motivations for the approach we propose.

In particular, some of the sources of the current pay gap at the WBG may have already vanished. For example, the fraction of women among all technical and managerial hires (GE and above) has almost reached parity, with a gap remaining only for mid-career hires (GG and GH). More generally, our descriptive results show that an organization such as the WBG is highly non stationary, with strong trends affecting the processes that generate pay gaps. Taking this inertia into account could be very important for policy lest differentials may be over corrected.

By simulating our algorithm forward, we can quantify the inertia exhibited by pay gaps. In doing so, we use the parameters estimated using the last year of the data for all the processes. This procedure allows us to project how much of the gender pay gap would close on its own in the coming years, and at what pace, if nothing changed in the organization.

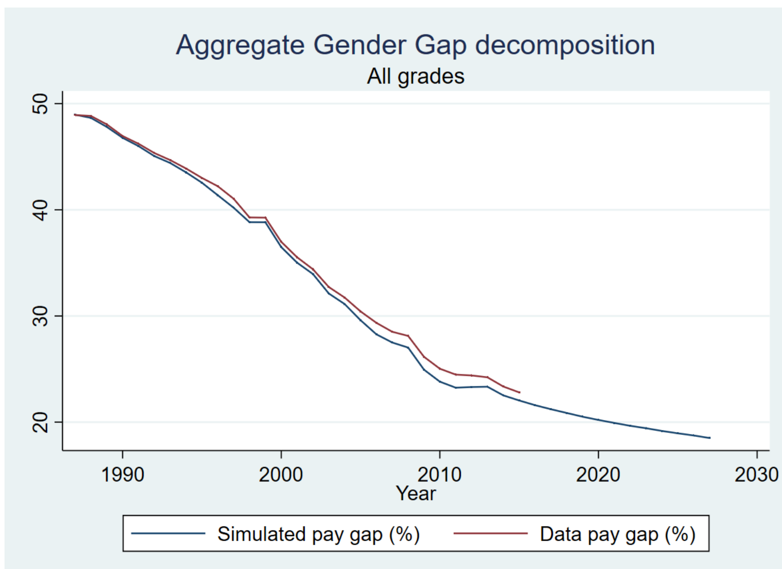
Figure 15 shows that under that scenario, the aggregate gender pay gap would continue to close but at a slower pace, from 23 percent to around 18 percent over the next decade. Interestingly, Figure 16 shows that this improvement would not come from improvements in the pay gap among technical staff. Gender pay differences for that group level off at 10 percent, reflecting that a higher proportion of women are hired at the entry technical position (GF) relative to men. Since hiring parity has not improved among support staff either (Figure 6), the projected gains in the aggregate gap are likely explained by the diminishing overall fraction of support staff, among which women are still largely over represented.

5.4 Aggregate Nationality Gap

To provide perspective and illustrate alternative applications of the methodology we consider how pay gaps associated with nationality compare with pay gaps associated with gender. That is, we decompose pay gaps between Part 1 and Part 2 employees (Table 7). The aggregate country part gap in 2015 is smaller than the aggregate gender gap at US$14,800, and well approximated by the simulations (see Figure 12, bars 5 and 6).

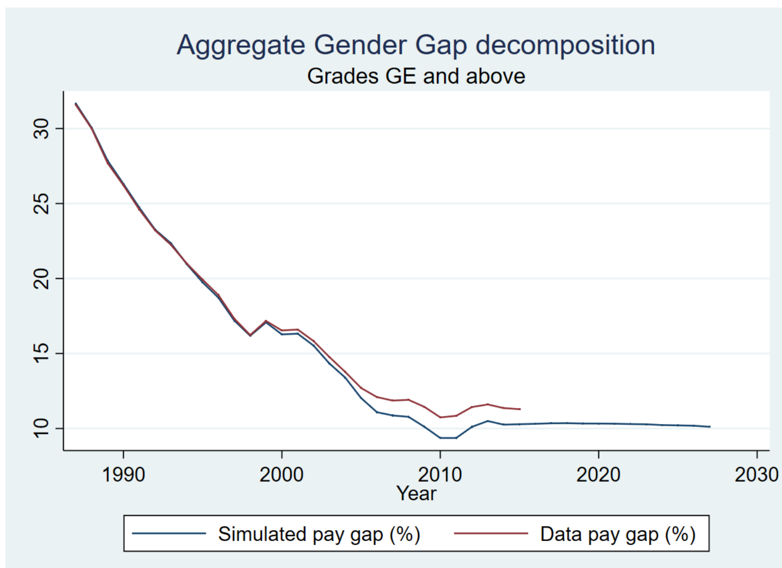
As with the aggregate gender pay gap, the bulk of the nationality gap (62 percent) corresponds to the fact that Part 2 employees are, on average, hired at lower grades. Equating exit probabilities across Part 1 and Part 2 employees further closes 14 percent of the gap. Equating salary growth closes an additional 6 percent, and equating entry salaries accounts for another 11 percent. The remainder of the gap can be attributed to differences predating 1987. Within the technical and managerial category (Table 7, columns 4-6), the composition of hires factor dominates as well.

Figure 15: Forward Projection of the Aggregate Gender Gap - All Staff



The figure shows the gender pay gap among all staff in the data between 1987 and 2015 and the simulated gender pay gap between 1987 and 2025. Simulations after 2015 obtained by iterating the microsimulation algorithm forward using the input parameters estimated in year 2015 for years 2016-2025.

Figure 16: Forward Projection of the Aggregate Gender Gap - Technical Staff



The figure shows the gender pay gap among staff hired at grade GE or higher in the data between 1987 and 2015 and the simulated gender pay gap between 1987 and 2025. Simulations after 2015 obtained by iterating the microsimulation algorithm forward using the input parameters estimated in year 2015 for years 2016-2025.

Table 7: Decomposition of the Aggregate Nationality Gap: Bootstrapped Distributions of the Percentage Contribution of each Decomposition Factor.

		All Grades			**Technical Staff**		
		mean	***p5***	***p95***	**mean**	***p5***	***p95***
legacy		*7.7*	*6.5*	*9.3*	17.0	*13.4*	*21.9*
salary growth		5.8	*3.4*	*8.1*	10.6	*4.5*	*18.1*
retention		14.0	*6.3*	*21.7*	4.2	*-15.1*	*21.3*
entry salaries		10.5	*8.1*	*13.1*	12.9	*7.7*	*19.5*
grade composition		61.7	*54.6*	*68.8*	55.0	*40.5*	*69.8*
Total pay gap	%		13.1%			5.9%	
	000USD		14.8			7.5	

The table entries are percent contributions for each decomposition factor. The distributions of the factor contributions are estimated by simulating the decomposition using each of 250 bootstraps of the simulation input parameters.

5.5 Robustness to the Decomposition Order

The decomposition results we have presented shut down sources of gender gaps in a specific order: retention, salary growth, entry salaries conditional on grade, entry grade, and pre-1987 legacy. This order is the reverse of the order in which each factor is simulated in our algorithm, but is otherwise arbitrary.

To test the importance of this choice, we implement our decomposition in all 120 possible different orders and report the distribution of contributions obtained for each factor in Tables 8 and 9. Although the order does affect the exact contribution of each factor, our results remain qualitatively identical and quantitatively very similar for all decompositions, with the nuance that differential retention tends to have a modest negative contribution to gender pay gaps.

Table 8: Decomposition of the Aggregate Gender Gap: Robustness to the Decomposition Order

		All Grades			Technical Staff		
		mean	**p5**	**p95**	**mean**	**p5**	**p95**
legacy		14.8	7.0	21.7	31.1	15.3	50.4
salary growth		2.9	-0.2	5.2	2.8	-2.5	6.9
retention		-7.4	-16.1	1.3	-20.0	-34.7	-7.0
entry salaries		8.5	7.2	9.9	13.0	9.5	17.3
grade composition		81.1	75.9	85.6	73.1	63.4	83.6
Total pay gap	%		22.9			11.3	
	000USD		27.4			14.6	

The table entries are percent contributions for each decomposition factor. The decomposition was obtained for each of all 120 possible decomposition orders. The moments reported in the table are computed over the set of results thus obtained.

Table 9: Decomposition of the Aggregate Nationality Gap: Robustness to the Decomposition Order

		All Grades			Technical Staff		
		mean	**p5**	**p95**	**mean**	**p5**	**p95**
legacy		8.8	6.9	10.6	20.1	15.3	25.0
salary growth		4.8	3.7	6.1	6.7	3.2	11.1
retention		16.6	15.0	18.4	8.9	2.6	15.3
entry salaries		11.5	10.3	12.9	14.3	10.9	18.0
grade composition		58.2	55.6	60.4	50.0	40.2	59.1
Total pay gap	%		13.1%			5.9%	
	000USD		14.8			7.5	

The table entries are percent contributions for each decomposition factor. The decomposition was obtained for each of all 120 possible decomposition orders. The moments reported in the table are computed over the set of results thus obtained.

6 Conclusion

We have proposed a methodology to decompose aggregate pay gaps in firms. We used it to analyze salary differences between genders at the WBG between 1987 and 2015. Aggregate gaps can be construed as an interplay between composition and compensation effects and our decomposition results suggest that composition effects play the major role in explaining the aggregate gap in 2015. A policy implication of this finding is that if firms are pressured or constrained to minimize aggregate gender pay gaps, it will generate incentives for them to outsource low-paid jobs held by women to external partners. Such regulation could favor changes to the scope of firms in the direction of greater segmentation by job type.

Identifying where the bulk of the aggregate gender gap comes from, and how persistent it will be—ongoing differences in the gender composition of hires, in the case of the WBG—can also guide data collection and further investigation in the most promising direction. For instance, we do not know whether occupational segregation emerges at the point of hiring (equally qualified men and women have applied, but men are chosen more often) or at the point of job applications (fewer qualified women apply relative to men). If differences arise at the application stage, a very different kind of policy would be required (such as an outreach program) than if the application pool is balanced (anti-discrimination training and penalties).

Lastly, our methodology can help quantify the fraction of current aggregate pay gaps that are the results of past hiring imbalances that no longer exist, akin to a star whose light we perceive after its source has died. In the case of our data, around 20% of the gap would disappear over 10 years, as more gender-balanced cohorts rise through the ranks. At least in the context of the WBG, the bulk of the gender gap will not therefore close on its own. Still, recognizing and quantifying these pipeline effects is necessary to design and target appropriate remedial policies. Considering that it takes three decades to train a female executive, such considerations are particularly relevant for a growing literature which considers gender pay gaps at the top of the wage distribution (e.g. Gayle et al. (2012), Kunze and Miller (2017), Flabbi et al. (2019), Bertrand et al. (2019), Dalvit et al. (2021) among others) where male-female convergence over recent decades has been much slower than at the middle of the distribution (Blau and Kahn (2017)).

References

Albrecht, J., Björklund, A. and Vroman, S. (2003), 'Is there a glass ceiling in sweden?', *Journal of Labor economics* **21**(1), 145–177.

Albrecht, J., Bronson, M. A., Thoursie, P. S. and Vroman, S. (2018), 'The career dynamics of high-skilled women and men: Evidence from sweden', *European Economic Review* **105**, 83–102.

Angelov, N., Johansson, P. and Lindahl, E. (2016), 'Parenthood and the gender gap in pay', *Journal of Labor Economics* **34**(3), 545–579.

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007), 'Is there a glass ceiling over europe? exploring the gender pay gap across the wage distribution', *ILR Review* **60**(2), 163–186.

Baker, G., Gibbs, M. and Holmstrom, B. (1994*a*), 'The internal economics of the firm: Evidence from personnel data', *The Quarterly Journal of Economics* **109**(4), 881–919.

Baker, G., Gibbs, M. and Holmstrom, B. (1994*b*), 'The wage policy of a firm', *The Quarterly Journal of Economics* **109**(4), 921–955.

Baker, G. and Holmstrom, B. (1995), 'Internal labor markets: Too many theories, too few facts', *The American Economic Review* **85**(2), 255–259.

Barnet-Verzat, C. and Wolff, F.-C. (2008), 'Gender wage gap and the glass ceiling effect: a firm-level investigation', *International Journal of Manpower* **29**(6), 486–502.

Bertrand, M., Black, S. E., Jensen, S. and Lleras-Muney, A. (2019), 'Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway', *The Review of Economic Studies* **86**(1), 191–239.

Bertrand, M., Hsieh, C.-T. and Tsivanidis, N. (2021), Contract labor and firm growth in india, Technical report, No. w29151. National Bureau of Economic Research.

Bielby, W. T. and Baron, J. N. (1984), 'A woman's place is with other women: Sex segregation within organizations', *Sex segregation in the workplace: Trends, explanations, remedies* pp. 27–55.

Blau, F. D. and DeVaro, J. (2007), 'New evidence on gender differences in promotion rates: An empirical analysis of a sample of new hires', *Industrial Relations: A Journal of Economy and Society* **46**(3), 511–550.

Blau, F. D. and Kahn, L. M. (2017), 'The gender wage gap: Extent, trends, and explanations', *Journal of Economic Literature* **55**(3), 789–865.

Blinder, A. S. (1973), 'Wage discrimination: reduced form and structural estimates', *Journal of Human resources* pp. 436–455.

Bourguignon, F., Ferreira, F. H. and Leite, P. G. (2003), 'Conditional cash transfers, schooling, and child labor: Micro-simulating brazil's bolsa escola program', *The World Bank Economic Review* **17**(2), 229–254.

Bourguignon, F., Ferreira, F. H. and Leite, P. G. (2008), 'Beyond oaxaca–blinder: Accounting for differences in household income distributions', *The Journal of Economic Inequality* **6**(2), 117–148.

Cohen, P. N. and Huffman, M. L. (2007), 'Working for the woman? female managers and the gender wage gap', *American Sociological Review* **72**(5), 681–704.

Dalvit, N., Patel, A. and Tan, J. (2021), 'Intra-firm hierarchies and gender gaps', *Labour Economics* p. 102029.

Doeringer, P. B. and Piore, M. J. (1985), 'Internal labor markets and manpower analysis', *Armonk,NY: Sharpe* .

Doeringer, P. B. and Piore, M. J. (2020), *Internal labor markets and manpower analysis: with a new introduction*, Routledge.

Fernandez, R. M. and Campero, S. (2017), 'Gender sorting and the glass ceiling in high-tech firms', *Ilr Review* **70**(1), 73–104.

Filmer, D., King, E. and Van De Walle, D. (2005), 'Testing for pay and promotion bias in an international organization', *International Journal of Manpower* **26**(5), 404–420.

Flabbi, L. (2010), 'Gender discrimination estimation in a search model with matching and bargaining', *International Economic Review* **51**(3), 745–783.

Flabbi, L., Macis, M., Moro, A. and Schivardi, F. (2016), Do female executives make a difference? the impact of female leadership on gender gaps and firm performance, Technical report, National Bureau of Economic Research.

Flabbi, L., Macis, M., Moro, A. and Schivardi, F. (2019), 'Do female executives make a difference? the impact of female leadership on gender gaps and firm performance', *The Economic Journal* **129**(622), 2390–2423.

Fortin, N., Lemieux, T. and Firpo, S. (2011), Decomposition methods in economics, *in* 'Handbook of labor economics', Vol. 4, Elsevier, pp. 1–102.

Gayle, G.-L., Golan, L. and Miller, R. A. (2012), 'Gender differences in executive compensation and job mobility', *Journal of Labor Economics* **30**(4), 829–872.

Gerhart, B. (1990), 'Gender differences in current and starting salaries: The role of performance, college major, and job title', *ILR Review* **43**(4), 418–433.

Giuliano, L., Levine, D. I. and Leonard, J. (2005), 'Race, gender, and hiring patterns: Evidence from a large service-sector employer', *Working manuscript* .

Gobillon, L., Leturcq, M., Meurs, D. and Roux, S. (2014), 'Elite institutions, fields of study and the gender wage gap: case study of a large firm', *File downloaded on March* **13**, 2017.

Guvenen, F., Kaplan, G. and Song, J. (2014), The glass ceiling and the paper floor: Gender differences among top earners, 1981-2012, Technical report, National Bureau of Economic Research.

G.V., X. (2017), 'Can we talk about the gender pay gap?', *The Washington Post* .
URL: *https://www.washingtonpost.com/graphics/2017/business/women-pay-gap/#auth-xaq*

Hersch, J. and Viscusi, W. K. (1996), 'Gender differences in promotions and wages', *Industrial Relations: A Journal of Economy and Society* **35**(4), 461–472.

Jones, D. R. and Makepeace, G. H. (1996), 'Equal worth, equal opportunities: pay and promotion in an internal labour market', *The Economic Journal* pp. 401–409.

Joubert, C. (2015), 'Pension design with a large informal labor market: evidence from chile', *International Economic Review* **56**(2), 673–694.

Keane, M. P. and Wolpin, K. I. (2010), 'The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women', *International Economic Review* **51**(3), 851–892.

Kleven, H., Landais, C. and Søgaard, J. E. (2018), Children and gender inequality: Evidence from denmark, Technical report, National Bureau of Economic Research.

Kunze, A. and Miller, A. R. (2017), 'Women helping women? evidence from private sector data on workplace hierarchies', *Review of Economics and Statistics* **99**(5), 769–775.

Malkiel, B. G. and Malkiel, J. A. (1973), 'Male-female pay differentials in professional employment', *The American Economic Review* **63**(4), 693–705.

Matsa, D. A. and Miller, A. R. (2011), 'Chipping away at the glass ceiling: Gender spillovers in corporate leadership', *American Economic Review* **101**(3), 635–39.

Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International economic review* pp. 693–709.

O'Donoghue, C. (2014), *Handbook of microsimulation modelling*, Emerald Group Publishing.

Orcutt, G. H. (1957), 'A new type of socio-economic system', *The review of economics and statistics* pp. 116–123.

Oyer, P. and Schaefer, S. (2011), Personnel economics: Hiring and incentives, *in* 'Handbook of labor economics', Vol. 4B, Elsevier, pp. 1769–1823.

Petersen, T. and Morgan, L. A. (1995), 'Separate and unequal: Occupation-establishment sex segregation and the gender wage gap', *American Journal of Sociology* **101**(2), 329–365.

Petersen, T. and Saporta, I. (2004), 'The opportunity structure for discrimination', *American Journal of Sociology* **109**(4), 852–901.

Ransom, M. and Oaxaca, R. L. (2005), 'Intrafirm mobility and sex differences in pay', *ILR Review* **58**(2), 219–237.

Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee (2017), Technical report, European Commission.

Waldman, M. (2012), 13. theory and evidence in internal labor markets, *in* 'The handbook of organizational economics', Princeton University Press, pp. 520–572.