Article

Towards effective interventions to address gender pay gaps: a causal approach in simulated salary band data

Thomas P. Moore[1], Camilla E. Penney[1,2*]

[1]Te Pūnaha Matatini, Aotearoa New Zealand
[2]Te Kura Aronukurangi | School of Earth and Environment, Te Whare Wānanga o Waitaha | University of Canterbury, Christchurch, Aotearoa New Zealand

Despite decades of recognition, gendered pay disparities have remained remarkably persistent in most organisations. In Aotearoa-New Zealand, Public Services such as Crown Research Institutes are required to report the summary statistics of gender and racial pay gaps. However, effectively addressing these pay gaps requires a more detailed understanding of the interplay between the factors underpinning them such as organisational role, tenure and gender. We propose a new methodology to disentangle the factors underlying organisational gender pay gaps using approaches from causal inference and Bayesian statistics. The method is designed to accommodate anonymised data collected at the salary band level. We demonstrate the technique using datasets simulated from a hierarchical organisation and reproduce pay gaps between gendered groups within a complex causal structure. We also model the effectiveness of specific interventions to quantify the degree of change required to reduce gender pay gaps. Although our approach cannot address the structural factors underlying pay gaps, it does provide a method to better understand pay gaps and possible solutions, as a crucial next step in achieving pay parity.

Introduction

There has been over 80 years of international agreement that wages should not depend on the gender of the employee (International Labour Organisation, 1951). Yet, in 2023 "men still earned more than women in most countries, in nearly all industries" (El Achkar, 2023). Aotearoa New Zealand (Aotearoa-NZ) is no exception. The aggregated gender pay gap across the motu (nation) remains at 8.2% (StatsNZ, 2024), despite equal pay having been a focus of gender equity campaigning in the country since at least the 1890s (Parker and Donnelly, 2020).

Challenges have been identified in the Aotearoa-NZ research system (Gaston, 2023; Ministry of Business, Innovation, and Employment, 2022; Ross, 2024) and the intersecting effects of pay inequity are an important factor in these discussions. Times of financial scarcity have complex effects on gender pay gaps, as was seen following the widespread adoption of austerity policies in the aftermath of the 2008 financial crisis (Karamessini and Rubery, 2013) but their longer-term implication is often increased inequality (e.g. Karamessini and Rubery, 2017, 2020). Such crises may also provide moments of potential change. The current restructuring of the research funding system (Ministry of Business, Innovation and Employment, 2024) in Aotearoa-NZ therefore demands both continued attention on pay equity and advocacy for effective interventions. In this study we address these needs by investigating a method for disentangling factors contributing to pay inequity and testing the efficacy of targeted interventions.

The current rhetorical focus of the Aotearoa-NZ government on equality of opportunity over equity of outcome (National Party and ACT New Zealand, 2023; National Party and New Zealand First, 2023) mirrors a common distinction in pay gap analysis. The most basic form of pay equity, 'equal remuneration for men and women workers for work of equal value' is enshrined in the International Labour Organization (ILO) Equal Renumeration Convention (International Labour Organisation, 1951), so represents a goal that many countries around the world, including Aotearoa-NZ, are committed to achieving. As a result, much of the literature has focused on disentangling the component of pay gaps explicitly associated with labour market discrimination (Blinder, 1973; Goldin and Polachek, 1987; Oaxaca, 1973). An alternative approach is to consider the gender pay gap as the overall difference in pay between gendered groups, whatever its cause, since the pay inequity that people actually experience is the combined result of this explicitly discrimination-based wage gap and the social factors that lead to women dominating part-time employment and being more frequently employed in jobs lower in an organisational hierarchy (Bergmann, 1974; Smith et al., 2017).

This broader interpretation of the gender pay gap to include socio-cultural controls on position and ability to work full time is consistent with the expansion of equal pay legislation in Aotearoa-NZ from more limited forms of pay equity (Ministry of Business, Innovation, and Employment, 1972), focused on 'equal remuneration for work of equal value', with gender as the only discriminant, to more recent initiatives, such as Kia Toipoto, which engage with pay

*Correspondence: camilla.penney@canterbury.ac.nz

gaps at a more structural level (although see Parker and Donnelly (2020) for a detailed discussion of the history of pay legislation in Aotearoa-NZ). This Public Service pay gaps' action plan (Te Kawa Mataaho, 2021) requires public service organisations to publish gender and ethnic pay gaps annually, with the aims of: ensuring bias does not influence starting remuneration or remuneration for employees in similar roles; developing equitable career pathways and opportunities to progress; protecting against bias and discrimination in promotion and renumeration policies and practices; building cultural competence; and normalising flexible working.

The current recommendations for calculating pay gaps in Aotearoa-NZ use summary statistics (StatsNZ, 2020). A mean or median pay value is calculated for each group (e.g., workers with a particular gender or ethnicity) in an organisation and these values are compared to identify and quantify pay gaps. Whilst this simple approach has advantages in being easy to calculate without in-depth understanding of statistical methods, it does not provide information about the interplay between factors influencing pay gaps. Effective interventions require a more nuanced understanding of the numerous factors contributing to pay inequity and their causal relationships (Caligaris, 2013; James and Brower, 2022). This more nuanced understanding has the potential to allow for cost-effective targeted interventions which, whilst not solving underlying discrimination issues, are important for continuing to work towards pay parity in times of economic scarcity. As such, collaboration between statistical scientists and the teams tasked with calculating pay gaps has potential to move organisations from identifying to addressing pay gaps.

A significant body of work has been done on disentangling the factors leading to gender pay gaps in Aotearoa-NZ universities (Barrett-Walker et al., 2023; Brower and James, 2020, 2023; James and Brower, 2022; McAllister et al., 2020). These authors used longitudinal data from the Aotearoa-NZ Performance-Based Research Fund (PBRF) to refute common hypotheses about the causes of gender pay gaps, such as parental leave or "demographic inertia" (Shaw and Stanton, 2012) and investigate intersectional pay and advancement gaps for Māori and Pasifika researchers.

Less attention, however, has been paid to appropriate methods for disentangling pay gaps in the public bodies required to report under Kia Toipoto. Here we focus on the Crown Research Institutes (CRIs), government-owned research centres that carry out scientific research "for the benefit of New Zealand" (Ministry of Business, Innovation, and Employment, 1992). At the time of writing, there are seven such institutes across the country. However, a recent review proposed the restructure of these institutions (Ministry of Business, Innovation and Employment, 2024), which are now in the process of being merged into three "public research organisations". There have also been a significant number of job cuts in many CRIs over the past year (e.g. Morton, 2024). We therefore specifically discuss the gender pay gap in CRIs, as institutes at the sharp end of Aotearoa-NZ's current research system crisis.

CRIs are also uniquely placed to take innovative approaches to address pay gaps. They are in the unusual position of having expertise in pay-gap reporting from a human resource perspective, as well as data science teams capable of implementing informative statistical methodologies, such as the one we present here. We also note that the methods we propose could be applied to any hierarchical organisation, and that the approaches we describe here could be particularly impactful if adopted in the set-up of the new public research organisations.

We focus here on binary gender pay gaps, as a simple illustration of the method and concept. This approach is not intended to imply, however, that gender is a binary. Indeed, gender-queer and non-binary individuals may face even greater pay gaps (e.g. in the Aotearoa-NZ public service sector; Te Kawa Mataaho, 2023). Additionally, ethnic or racial pay gaps, particularly, in Aotearoa-NZ, between Pākehā and Māori and Pasifika, are frequently more significant than, and intersect with, gender pay gaps (Came et al., 2020; McAllister et al., 2020; McCall, 2001). The causal structures which lead to pay discrimination are likely to be different for each of these factors, so we use a binary gender pay gap here as a minimal example. However, with real data there would be no reason not to expand the method to consider more complex, intersectional pay gaps.

We first describe the set-up of the causal structure relating gender to salary and the ways in which we hypothesise that additional organisational and societal factors affect this relationship (Causal foundations). To avoid the anonymity and privacy concerns associated with using real CRI salary and gender identity data, we then use this hypothesised causal structure to simulate pay gap data, for an organisation with a similar structure to many CRIs (Simulating a gender pay gap). We describe a method for analysing our simulated dataset using a causal inference approach (Bayesian measurement error model, Models and disparity indicators, Interventions and patriachal effects) and demonstrate the viability of disentangling various factors contributing to pay gaps and the possibility of doing so even with limited pay data (Results). The statistical details of this methodology are given in the supplementary information and in the code (available on GitHub and Zenodo[1]). Finally, we discuss the implications of our results for pay gap analysis in Aotearoa-NZ, with particular relevance to Aotearoa-NZ, particularly the current research system crisis (Discussion).

Methods

Causal foundations

In statistical causal inference, directed acyclic graphs (DAGs) are used to show the way in which different factors are assumed to interact to generate an outcome. These graphs allow us to identify which additional factors (or covariates) should be ex/included when modelling the relationship between a variable of interest (here gender) and a particular outcome (here salary). In this study,

[1]https://github.com/rgonjargon/Causal_Inference_GPG, 10.5281/zenodo.15428858

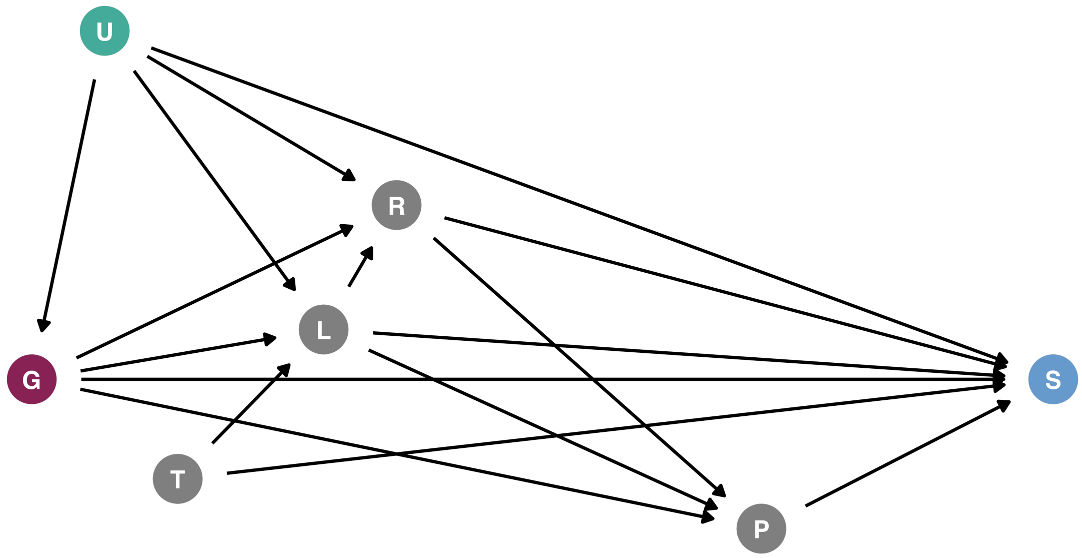


Figure 1: Directed acyclic graph (DAG) of hypothesised causal relationships between salary (S; outcome = blue) and gender (G; exposure = red). Measured variables are in grey: T = tenure; L = level; R = role type; and P = percentage FTE. Unmeasured patriarchal effects (U) are in green. Arrows indicate the direction of hypothesised causal relationships. Evaluated with DAGitty v3.1 (https://www.dagitty.net) and plotted in the DAGitty R package v0.3-1 (Textor et al., 2017). For the data simulation the magnitude, direction, and type of relationships between variables are detailed in Supplementary Material 1.

the additional factors are primarily information that might typically be collected by People and Culture – namely, position in the organisational hierarchy ("level", L), role type (R), length of time in role ("tenure", T), and percentage full-time equivalent or FTE (P). We also include the effect of unmeasured (and often unquantifiable) societal factors (U), which we term "patriarchal effects". We note that U really represents all unquantified or unquantifiable factors that influence the relationship between gender and salary, i.e., intersectional components of identity, such as race, which have been clearly demonstrated to interact with gender (Chapman and Benis, 2017; McAllister et al., 2020; McCall, 2001).

Figure 1 shows our hypothesised DAG for the gender pay gap. We propose that: 1) an employee's salary is controlled by gender (G), level (L), role type (R), tenure (T), and percentage full time equivalent (P); 2) an employee's gender affects their level, role type, and percentage FTE; 3) how long an employee has been in their role (T) affects their level in the organisational hierarchy (L); 4) percentage FTE (P) and role type (R) are controlled by level (L); and 5) patriarchal effects (U) confound some of these relationships by causally influencing gender, level, role type and salary (Figure 1). We note that this is a non-exhaustive list of factors that could influence the relationship between gender and salary but includes a sufficiently wide range of factors to demonstrate the value of using this framework in analysing salary data.

When causally estimating their relationship, gender and patriarchal effects (U) form a path that, statistically, confounds other relationships. Although U cannot be quantitatively measured, it can be considered when interpreting gender pay gap estimates. Here this is demonstrated by ex/including patriarchal effects in modelling to assess their impact. The other factors that are measured can be included directly as parameters in the statistical model (Cinelli et al., 2024).

An intersectional analysis would need to revise the DAG proposed here, to include the ways in which race, ethnicity and gender interact, both with each other and with the other parameters.

Simulating a gender pay gap

The motivation for this study comes from pay gaps in CRIs in Aotearoa-NZ. We therefore simulate pay data for an organisation with a similar hierarchical and employment structure to CRIs. Crown Research Institutes employ staff in two primary groups: research (mostly scientists) and support (which includes a wide range of staff, such as administration and senior management). For simplicity we neglect technical roles, which form a third category in some CRIs. Crown Research Institutes operate a band system for pay, with employees assigned to a particular salary band that represents their possible pay range. We are interested in simulating pay data based on these salary bands because salary band information is often available to CRI employees, for example through charge-out sheets, whilst anonymity concerns prevent access to full pay data.

We simulated the salaries for a hypothetical organisation of 500 employees with a hierarchical structure using the package 'simDAG' (Banks et al., 2014; Denz and Meiszl, 2023) in the R statistical software package (R Core Team,

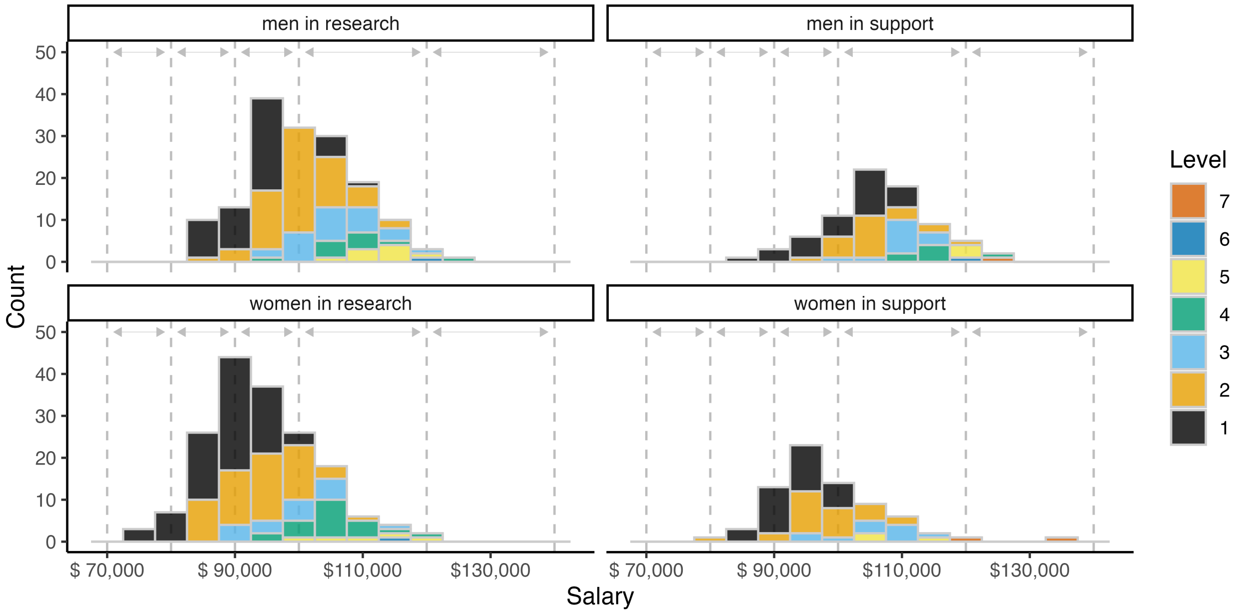


Figure 2: Simulated salary data as stacked histograms by level for different gender and role type groups. Salary bands were defined as the arrows between dashed grey lines. Level = position in the organisational hierarchy. Level 1 employees are those with no one reporting to them and level 7 employees are those that only have the CEO as a line manager.

2024). We assign the organisation-wide gender pay gap to be 5% ($4000) and the pay gap between women in research and women in support to be $2000. The resulting dataset is shown in Figure 2. The code and associated simulation details (magnitude, direction, and type of relationships between variables) are available on GitHub and an R-markdown document summarising the modelling is in the Supplementary Material 1.

Bayesian measurement error model

To estimate the gender pay gap, a Bayesian measurement error model was fitted to the simulated data. The fitting was carried out in R v4.4.1 (R Core Team, 2024) using brms v2.21.0 (Bürkner, 2017) as a front end to Stan (Stan Development Team, 2023). We use salary-band mean and approximated standard deviation (SD) as inputs to this model. We use a very simple approach to calculating these parameters, namely:

$$Salary\ band\ mean = \frac{(maximum + minimum)}{2}$$

$$Salary\ band\ SD = \frac{(maximum - minimum)}{4}$$

i.e. assuming that the salary data are Gaussian within each band and approximating the standard deviation using the range rule (Wan et al., 2014).

Our model estimates the relationship between salary and gender, as well as additional factors of role type, level, percentage FTE, tenure and patriarchal effects. Since the aim of this study is to outline a potential approach and to motivate collaboration between People and Culture and statisticians within CRIs, we leave detailed discussion of the choices made for statistical modelling to the supplementary information. Below, we briefly outline some of the simplifications used here and how they might be expanded for real data.

In the model, we specified how gender, role, and level (an ordered category that we treat as a monotonic predictor) interact. To examine the effects of tenure we divided the simulated employees into four categories, each modelled using a separate spline (to allow progression to be non-linear and different between groups): women in research; women in support; men in research and men in support.

Patriarchal effects were included in some models to remove the effect of confounding (a model is "confounded" in statistics if unaccounted variables link the independent variable, here gender, and dependent variable, here salary: e.g., men are advantaged when applying for support roles). Note that it is possible for us to explicitly remove confounding due to patriarchal effects in our models because we are using simulated data and know the form of the effects; removing these effects would not be trivial for real-world data. However, recognising the relationship between patriarchal effects and other variables would be crucial to understanding pay inequity within an organization and addressing it appropriately.

For simplicity, we neglect random effects – factors that form known groupings in the data which are likely to influence salary that we want to account for – but grouping variables of site location and team could be included in future analyses. A Gaussian likelihood was selected since, by design, the simulated salary-band data follow a normal distribution, but a skewed normal likelihood could be chosen for real data if, for example, a few employees have much higher salaries than the mean salary or certain levels receive automatic progression. Measurement error model validation

Model	Input Data	Factor Included	Confounded (U) ?
A_1	individual salaries	G,R,L,T,P	No
A_2	salary bands	G	No
A_3	salary bands	G	Yes
A_4	salary bands	G,R,L,P,T	No
A_5	salary bands	G,R,L,P,T	Yes
A_6	salary bands	G,R,L,P,T + intervention 1	No
A_7	salary bands	G,R,L,P,T + intervention 1	Yes
A_8	salary bands	G,R,L,P,T + intervention 2	No
A_9	salary bands	G,R,L,P,T + intervention 2	Yes

Table 1: Models used in gender pay gap modelling. A1 has the full salary data generated in our simulation as input, including patriarchal effects, so represents the "true" pay gaps. Factors included in our models are: G – gender, R – role type (research or support), L – level in the organisational hierarchy (1 – 7, with 1 being no reporting employees and 7 being only reporting to the CEO), T – tenure (number of years in the organisation) and P – percentage full-time equivalent. Models that are confounded do not include U, the patriarchal effects that contribute to the gender pay gap in our simulated data. Models A6- A9 are used to investigate the effects of an intervention to reduce the gender pay gap (described in Interventions and patriachal effects).

was performed using prior and posterior predictive checks, as well as effective sample sizes and r-hats (McElreath, 2018).

Models and disparity indicators

We initially used five models, including different subsets of the factors in Figure 1 (Table 1). The baseline model, A1, has the full salary data (not in bands) as an input and includes all the factors that affect salary except for patriarchal effects (U in Figure 3). This model therefore represents the "true" pay gaps in our simulated salary data. We used a series of other models based only on salary band data to explore the effects of in/excluding different factors. These models were: A2) gender only (which is usual for gender pay gap analysis); A3) gender with confounding patriarchal effects; A4) gender and factors that affect salary; A5) gender, factors that affect salary, and confounding patriarchal effects.

The results of these models are summarised in an R-markdown document that provides a reproducible framework for how gender pay gap information could be reported (Supplementary Material 2).

To assess the effects of including different factors on our model results we use six "disparity indicators", which quantify gendered pay gaps, the difference in salaries as a dollar amount: H1) across the organisation; H2) between men in support and women in support; H3) between men in research and women in research H4) between men in support and men in research; H5) between women in support and women in research and H6) between men in support and women in research. These disparity indicators were chosen to correspond to the metrics of interest for standard gender pay gap reporting (StatsNZ, 2020).

Interventions and patriachal effects

The primary motivation for identifying gender pay gaps is, at least ostensibly, to reduce them (Lips, 2003, 2013). As such, it would be beneficial for organisations to be able to assess the effectiveness of targeted interventions. A key advantage of using a causal structure that identifies how different factors interact to produce the gender pay gap is the ability to model the effects of such interventions. We therefore introduce four additional models, A6 to A9. A6 and A7 simulate the effects of an intervention where women across the entire organisation had their salary increased by the StatsNZ-calculated gender pay gap (i.e., the median difference + \$7187). A6 includes this intervention, gender and other factors affecting salary, whilst A7 also includes confounding patriarchal effects. A8 simulates the effects of an intervention where all woman in level 1 research roles were promoted to the next highest salary band. A9 is similar to A8, but also includes confounding patriarchal effects.

Results

Gender pay gap reporting

In this section, we report the results of the A4 model, which includes both gender and other factors affecting salary (without confounding). The results of the other models are summarised in Table 2. Since model A4 accounts for known factors influencing the gender-salary relationship and does not include confounding patriarchal effects (i.e., information of how the data were confounded was included in the model), recovered estimates are the most comparable to those simulated in the salary band data. Bayesian inference provides parameter estimates as posterior probability distributions, which can then be further summarised as a posterior median and 95% credible interval (Figures 3, 4, 5 and 6). Classical statistical significance metrics can be overlaid on these statistics (e.g., if 95% credible intervals do not overlap, they can be considered significantly different), although interpretation should primarily focus on the effect size and data context; if a gender pay gap is not statistically different from zero this may not indicate no difference in pay, for example.

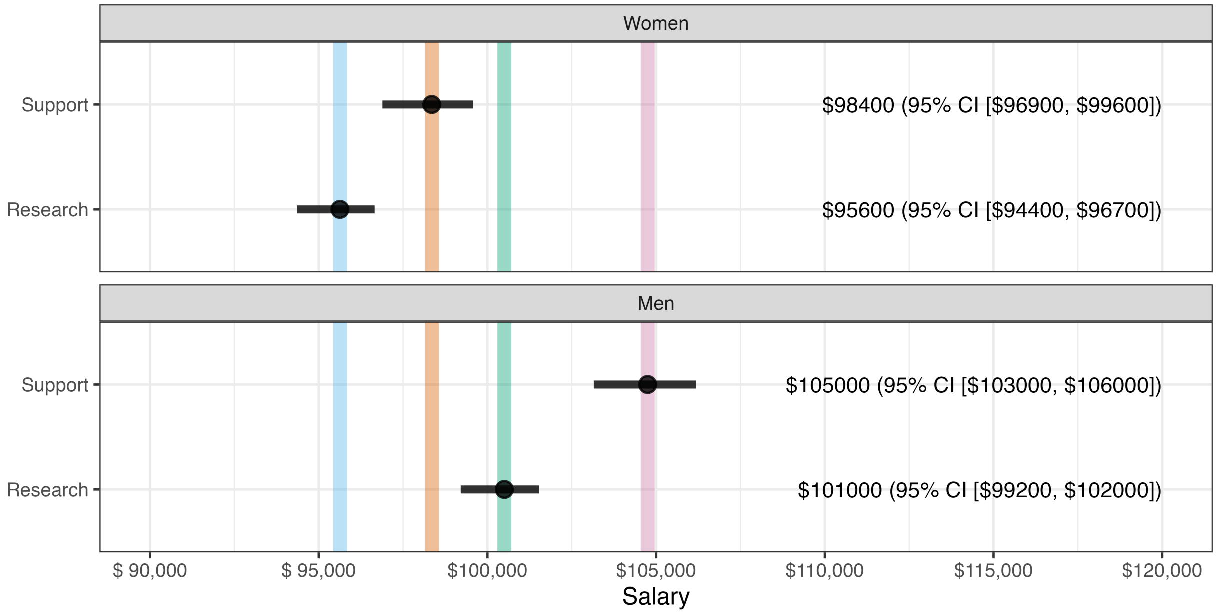


Figure 3: Conditional effect of gender and role type on salary for the A4 simulation. Estimates are displayed as intervals (median of the posterior distribution (black dot) with 95% credible interval (range)) and labelled as text. Vertical lines display median estimates: women in research (blue); women in support (orange); men in research (green); men in support (purple).

(Figures 3, 4 and 5) are conditional effect plots, which show the relationship between gender, salary, and, e.g., level (Figure 3), whilst keeping other factors at their mean values. Since the other factors are held at their mean values, the conditional effect plots are not comparable to raw data plots and so raw data are not overlaid onto these outputs. Figure 7 shows visualisations (posterior probability distributions) of the six gendered disparity indicators outlined above.

The conditional effect plot of gender and role type shows differences in salary across these groups in the hypothetical organisation, where men in support had the highest salaries, followed by men in research, women in support, and women in research (Figure 3). As mentioned in the methods, "support" includes senior management roles, as well as many low paid roles, so these hierarchies are consistent with typical CRI salary data (Figure 2). Compared across levels, salaries were lower in research than support, where differences between men and women were more pronounced (Figure 4). Salary was positively related to tenure (years of employment) for all gender and role type groups, although the form of the relationship differed: i.e., for support, progression tends towards an upper limit whereas progression in research roles is linear (Figure 5). Percentage full-time equivalent showed a negative relationship with salary, i.e., employees on part-time work had lower salaries (measured at 100% FTE) than full-time employees (Figure 6). All of these model results are consistent with the input salary data (i.e., the 'true' pay gap, Figure 1), suggesting that our model is able to qualitatively reproduce an example gendered organisational salary structure based on salary band information alone.

The six specific disparity indicators (H1–H6), which show different gender pay gaps both across the hypothetical organisation and between role-type groups, were visualised as posterior (probability) distributions (Figure 7). Accordingly, the organisational gender pay gap was estimated to be $4870 (H1); men earned more than women in support by $6410 (H2); men in research earned more than women in research by $4870 (H3); men in support earned more than men in research by $4260 (H4); woman in support earned more than woman in research by $2720 (H5); and men in support earned more than women in research by $9,130 (H6).

Disparity indicators across analyses

To show the ability of different analyses to recover simulated parameter estimates we performed a baseline analysis on salary data with no measurement error to which all other analyses were compared. This baseline analysis (A1) was able to recover the parameter estimates accurately for all disparity indicators (Table 2). Measurement error models on salary band data not including additional factors influencing salary (A2), with confounding (A3), and with both additional factors and confounding (A5) overestimated the simulated gender pay gap by amounts ranging from approximately $2000 to $4000 (Table 2). The best analysis (presented in Figures 3–6) was A4, which most accurately recovered simulated parameter estimates (regression coefficients) from salary band data (Table 2), perhaps unsurprisingly since this model is the most like the structure of the simulated data. However, even for A4 the disparity indicator H1, which is the most similar to current standards for pay-gap analysis (StatsNZ, 2020), significantly overestimates the simulated gender pay gap regression coefficient as a dollar amount (i.e., by approximately $1000

Simulation	Disparity Indicators / Hypotheses					
	H_1	H_2	H_3	H_4	H_5	H_6
Assumed structural parameters	\$4000				\$2000	
A_1: G-S covariates (baseline)	\$4239	\$5645	\$4239	\$3463	\$2056	\$7702
Measurement error models						
A_2: G-S	\$5800					
A_3: G-S not accounting for confounding	\$7734					
A_4: G-S	\$4874	\$6413	\$4873	\$4258	\$2719	\$9132
A_5: G-S covariates + not accounting for confounding	\$7020	\$7924	\$7020	\$5458	\$4554	\$12487
A_6: G-S covariates + intervention one	-\$2692	-\$640	-\$2691	\$3947	\$1896	\$1256
A_7: G-S covariates + intervention one + confounding	-\$550	\$555	-\$550	\$5229	\$4124	\$4679
A_8: G-S covariates + intervention two	-\$197	\$6621	-\$197	\$4121	-\$2697	\$3924
A_9: G-S covariates + intervention two + confounding	\$925	\$8121	\$925	\$5404	-\$1793	\$6329

Table 2: Results of disparity indicators (mean of the posterior distribution: H1-H6) for different analyses (A1-A7) and known estimates (regression coefficients). Note disparity indicators were not estimated for models that did not include those factors (A2 and A4). G-S = gender-salary relationship. H1 = mean organisational gender pay gap; H2 = how much more do men in support earn than women in support; H3 = how much more do men in research earn compared with women in research; H4 = how much more do men in support earn than women in support; H5 = how much more do women in support earn than women in research; and H6 = how much more do men in support earn than women in research.

Table 2).

The intervention of increasing the salary for women by \$7186 (the median organisational pay gap calculated using the StatsNZ methodology) effectively removed the overall gender pay gap (H1 and H2); but did not reduce the other disparity indicators to 0: H3 showed a decrease; H4 a negligible change; and H5-6 a decrease (Table 2). The second intervention, in which women in level one research were promoted the next pay band, reduced the overall gender pay gap to almost zero, but the effects on other disparity indicators were again mixed (Table 2). If confounding (patriarchal) effects are introduced in the analysis post-intervention, disparity measures increase to incorporate this information, i.e., gender pay gap estimates increase from -\$2692 to -\$550 and from -\$197 to \$925, for intervention one and two, respectively (Table 2).

Discussion

Causal inference as a tool for gender pay gap analysis

Our results demonstrate that a Bayesian measurement error model can successfully reproduce both qualitative (Gender pay gap reporting) and quantitative (Disparity indicators across analyses) aspects of the gendered pay structure of a complex, hierarchical organisation. The similarities of the pay disparity estimators in models A1 (full salary data) and A4 (band data only with the same causal structure as the simulated data) show the potential power of Bayesian causal inference approaches in understanding and disentangling pay gaps, even based on limited data.

These results also demonstrate that without considering the underlying causal structure of pay gaps we are likely to estimate pay gaps inaccurately, and to be unable to implement effective interventions to address them. Models that only link gender and salary (A2 and A3) without accounting for other factors inaccurately estimate the organisational gender pay gap for our simulated organisation (H1 and H2). Further, such models cannot provide information on occupational segregation, whereby certain roles or modes of working are disproportionately occupied by a particular gendered group, which contributes significantly to gender pay gaps (Blackwell, 2001; Roos, 1985; Wong and Charles, 2020).

The problem of using such simplified analysis – our closest analogue to the summary statistics approach that is currently standard in A-NZ (StatsNZ, 2020) – becomes particularly clear when we look at potential interventions to address the gender pay gap. Model A6 tests the effect of the most naïve intervention – simply increasing the salaries of all women in the organisation by the organisational pay gap, H1, here about \$7200. Note that the modelled intervention assumes that the organisational pay gap has been accurately diagnosed to begin with, taking the pay gap as H1 from model A1. H1 for model A6 is approximately -\$2700 (Table 2), showing that this naïve approach can indeed remove (or even reverse) the organisational gender pay gap, at least instantaneously. It would, however, be hard to argue that model A6 represents an organisation that has achieved pay equity as the other disparity indicators clearly show residual pay inequity, for example between men in support and women in research, due to not accounting for role type in the intervention.

These models also demonstrate the importance of identifying and including confounding variables in pay gap

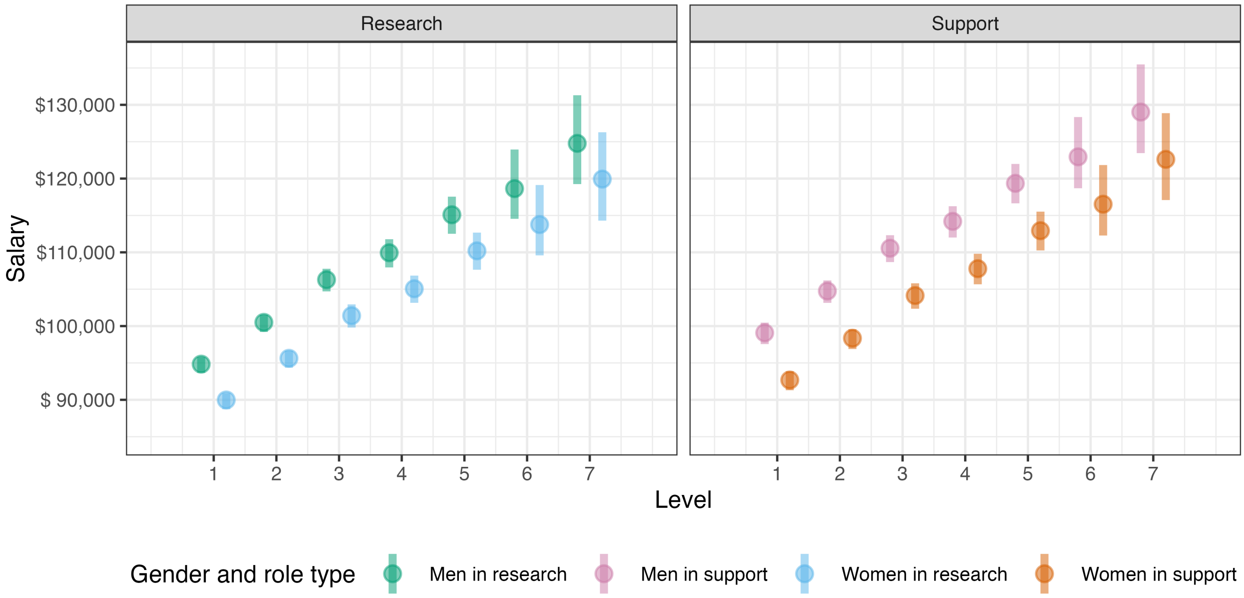


Figure 4: Conditional effect plot of gender and role type across levels for the A4 simulation: women in research (blue); women in support (orange); men in research (green); men in support (purple). Level 1 employees are those with no one reporting to them and level 7 employees are those that only have the CEO as a line manager. Points indicate median of the posterior distribution and lines the 95% credible intervals.

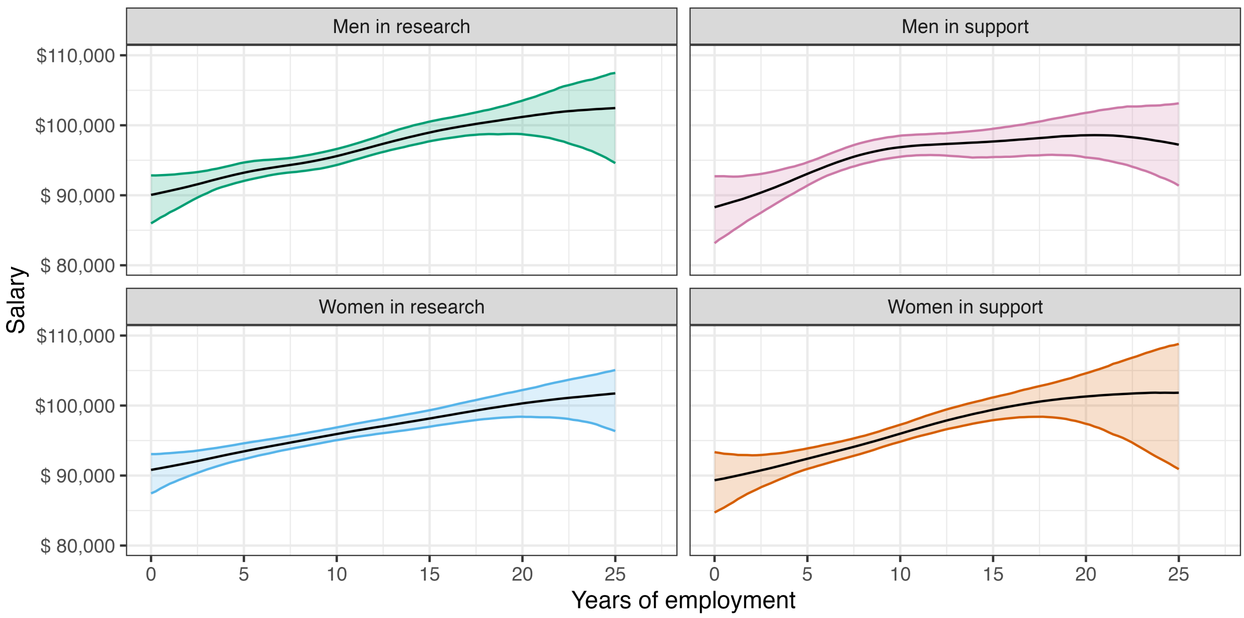


Figure 5: Conditional effect of gender and role type on salary over time for the A4 simulation. Estimates are displayed as the median of the posterior distribution (line) with 95% credible interval (ribbon): women in research (blue); women in support (orange); men in research (green); men in support (purple). The relationship between salary and years of employment was fitted as a spline: i.e., allowed to be non-linear.

analysis. Statistically, a confounding variable is one that influences both the predictor and response variables (gender and salary, respectively), meaning that the relationship between these variables is not the same as would be modelled assuming only a direct linkage. We see this principle in the results of both A3 and A5, where unmodelled patriarchal effects (U in Figure 1) lead to overestimates of the organisational pay gap and inaccurate

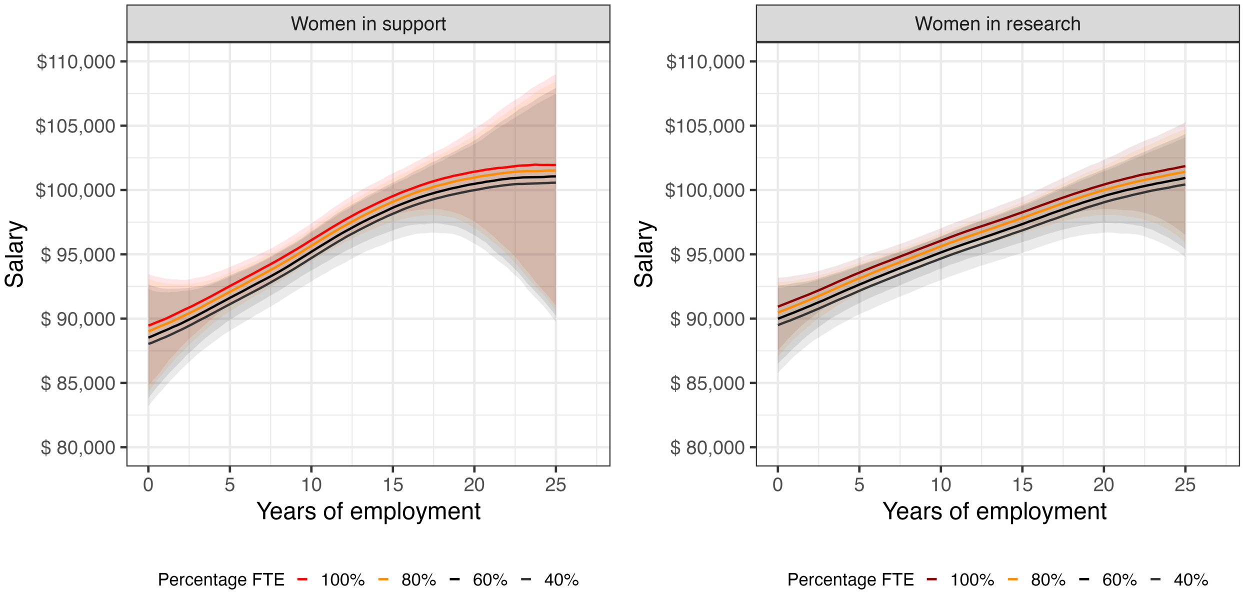


Figure 6: Conditional effect of percentage FTE on salary over time for the A4 simulation.. Estimates are displayed as the median of the posterior distribution (line) with 95% credible interval (ribbon): women in research (left); and women in support (right). The relationship between salary and years of employment was fitted as a spline: i.e., allowed to be non-linear.

estimates of other disparity indicators (Table 2). The significance of these results is that confounding leads to inaccurate understanding of the relationship between gender and salary, rather than overestimation specifically, which is likely related to the details of the measurement model. The challenge of such confounding variables is that, unlike for the simulated salary data used here, we do not in general know how these variables are functionally related. That is, confounding variables are crucial to the accuracy of pay-gap analysis but also hard to appropriately correct for (Christenfeld et al., 2004)

The challenge of incorporating confounding variables into real-world data analysis highlights a more general issue – how to choose Directed Acyclic Graph (DAG, Figure 1) or, equivalently, the hypothesised causal structure relating gender and salary. In this study, we have used a relatively simple DAG and, since we know all the factors contributing to the simulated dataset, we can model the associated pay gaps well (although note that this is not trivial, since our models A2–A9 are based on salary-band data rather than full salary data). Working out the appropriate causal structure is, therefore, key to accurately disentangling the different factors contributing to gender pay gaps. Whilst it is potentially challenging to understand which factors are causally related, and how, we also argue that a DAG, as a diagrammatic representation of an interpreted causal structure, could be used as a point of common understanding between different groups concerned with calculating gender pay gaps, such as People and Culture or Human Relations teams and statisticians or data scientists.

Our simplified DAG and salary-band based models are designed to investigate coarse pay data. However, People and Culture often gather more detailed information, including race and ethnicity data, which we have not included in our example analysis. The modelling approach we have described could be equally well applied to such data, with a DAG including the direct effects of racial discrimination and racial organisational segregation as well as the confounding effects of racism, sexism and intersectional interactions between race and gender.

Implications for addressing pay gaps in Aotearoa New Zealand

Our models demonstrate the possibility of moving beyond summary statistics towards a causal analysis of pay gaps (Figures 8 and 9). Such an analysis allows us to ask – and answer – the more complex questions required to begin effectively addressing, rather than simply quantifying, pay gaps. All of the CRIs, as scientific research organisations, have statistical science or data science teams. Thus, although the methods we discuss here might seem prohibitively complex for the teams usually tasked with pay gap reporting, such as Human Resources or People and Culture, the expertise to carry out these analyses is already present within CRIs. Indeed, a statistical causal inference approach has the potential to provide a language for communication between People and Culture, statistical scientists, and senior leadership teams tasked with identifying how pay in an organisation should ideally be related to factors such as role and time spent at the organisation.

Collaboration between People and Culture and statistical science teams would also allow exploration of intersectional factors such as race or ethnicity, which require detailed data. In some cases, despite collaboration, the full salary dataset may not be available, for example for reasons of

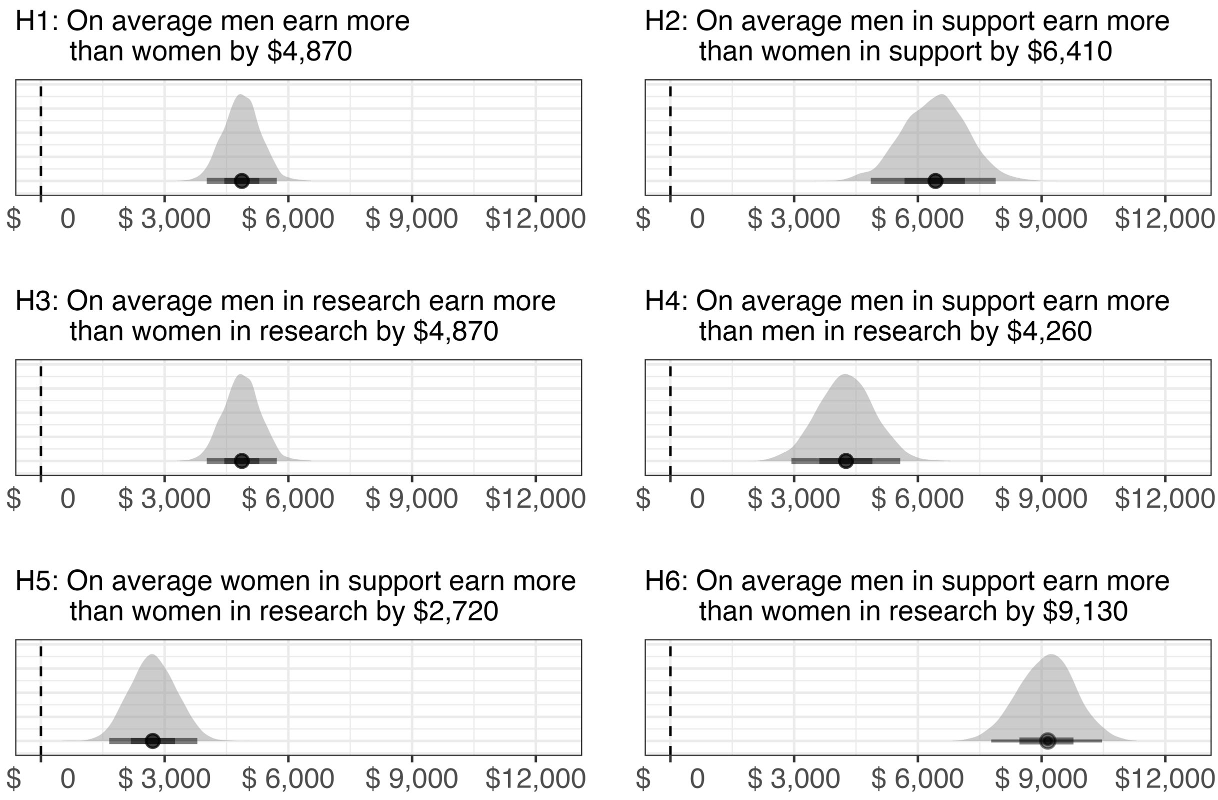


Figure 7: Posterior (probability) distributions of disparity indicators H1-6 for the hypothetical hierarchical organisation A4 simulation with the mean (point) and 95% credible interval (range) displayed. Dashed lines indicate zero, i.e., no difference in salary: a classical interpretation would suggest that if the 95% credible interval contains zero then there is no "statistically significant" difference. Results are from the A4 G-S covariates measurement error model (text in Table 2).

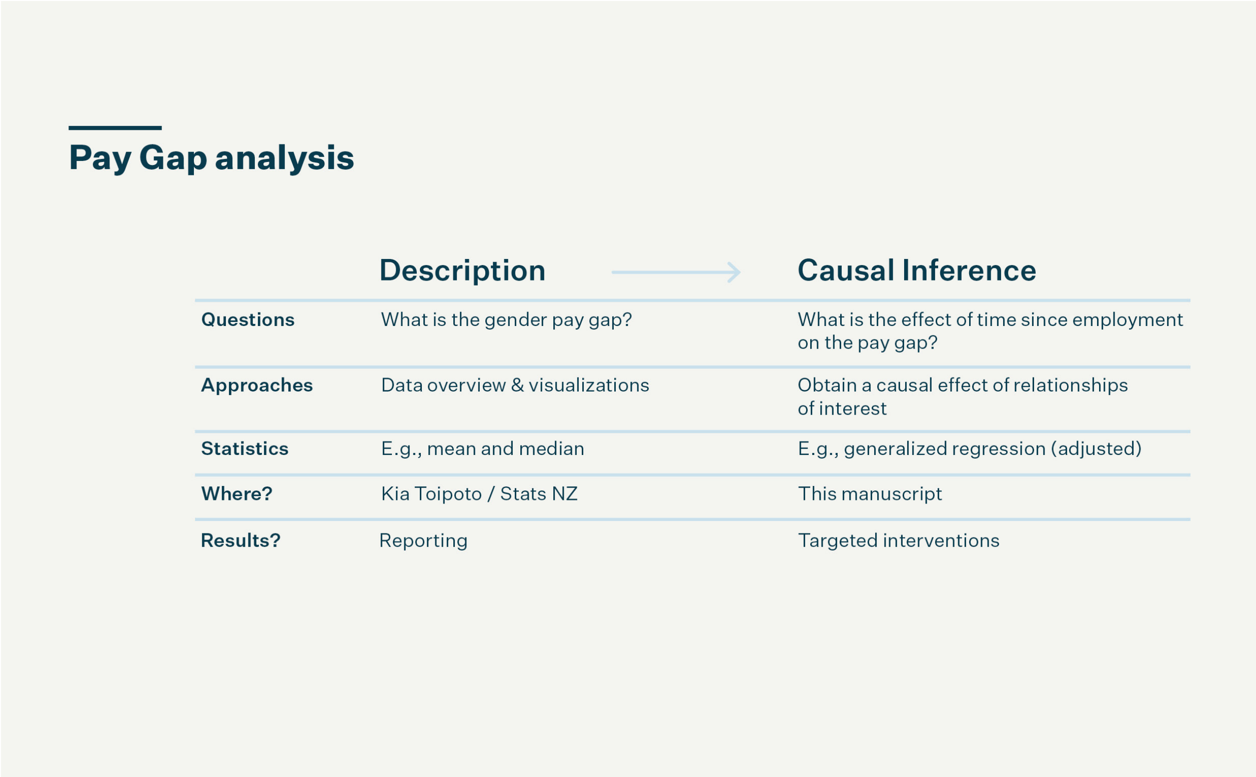
anonymity for small demographic groups. The salary-band based modelling we use accurately reproduces pay gaps despite significant loss of granularity, suggesting that detailed causal analysis is possible even where privacy concerns might prevent sharing of the full salary dataset between People and Culture and statistical scientists. The ability to work with coarse data could also be useful for identifying and addressing non-binary gender pay gaps, where the number of individuals in a category may be small.

There is little point in identifying a pay gap if no action is taken to address it. The ability to explore the effects of interventions and to identify potential unintended consequences or ineffectiveness (e.g. the increase in H3 in A6), may be particularly relevant at a time when the Aotearoa-NZ research sector is struggling for funding. Many authors have argued against the use of logics of scarcity to justify continued, or exacerbated, inequality (e.g. Karamessini and Rubery, 2013; Mehta, 2013). The methods demonstrated here may provide an explicit argument against such logics, since interventions can be targeted to be effective, at least in an instantaneous sense. We have provided the code for our model to ensure that the presented study is reproducible and to support the application of these methods to real salary data sets at CRIs, the future public research organisations and/or other organisations with comparable pay structures.

The aim of this study is not to advocate for causal analysis of pay gaps as a panacea but rather as a dramatic advance on current best practice which constitutes a pragmatic and easily realisable approach to improving pay equity across the motu. We note that, just as "causal" analysis cannot identify the true root causes of pay gaps in their historico-socio-cultural complexity,(Lips, 2013), modelling of interventions is not a panacea for ensuring that they are appropriate or effective. Our models cannot include the temporal impacts of discrimination, meaning that even an intervention designed to remedy all pay gaps in an organisation would likely be undone over time, at least without radical structural changes in both hiring practices and workplace culture (e.g. Bernard and Cooperdock 2018; McAllister et al. 2022; Naepi et al. 2020; Sensoy and DiAngelo 2017).

Conclusion

We have demonstrated a method for modelling gendered pay disparities using Bayesian measurement error models. This method can reproduce disparity indicators in synthetic salary data for a hypothetical organisation with a complex, hierarchical structure. Our modelling demonstrates that current methods of gender pay gap analysis, which do not account for indirect linkages between gender and salary (for example through gendered distributions of role type),



Pay Gap analysis

	Description	Causal Inference
Questions	What is the gender pay gap?	What is the effect of time since employment on the pay gap?
Approaches	Data overview & visualizations	Obtain a causal effect of relationships of interest
Statistics	E.g., mean and median	E.g., generalized regression (adjusted)
Where?	Kia Toipoto / Stats NZ	This manuscript
Results?	Reporting	Targeted interventions

Figure 8: Conceptual diagram of the difference between description (reporting) and causal inference. Only a causal understanding allows targeted interventions. Adapted from Laubach et al. (2021).

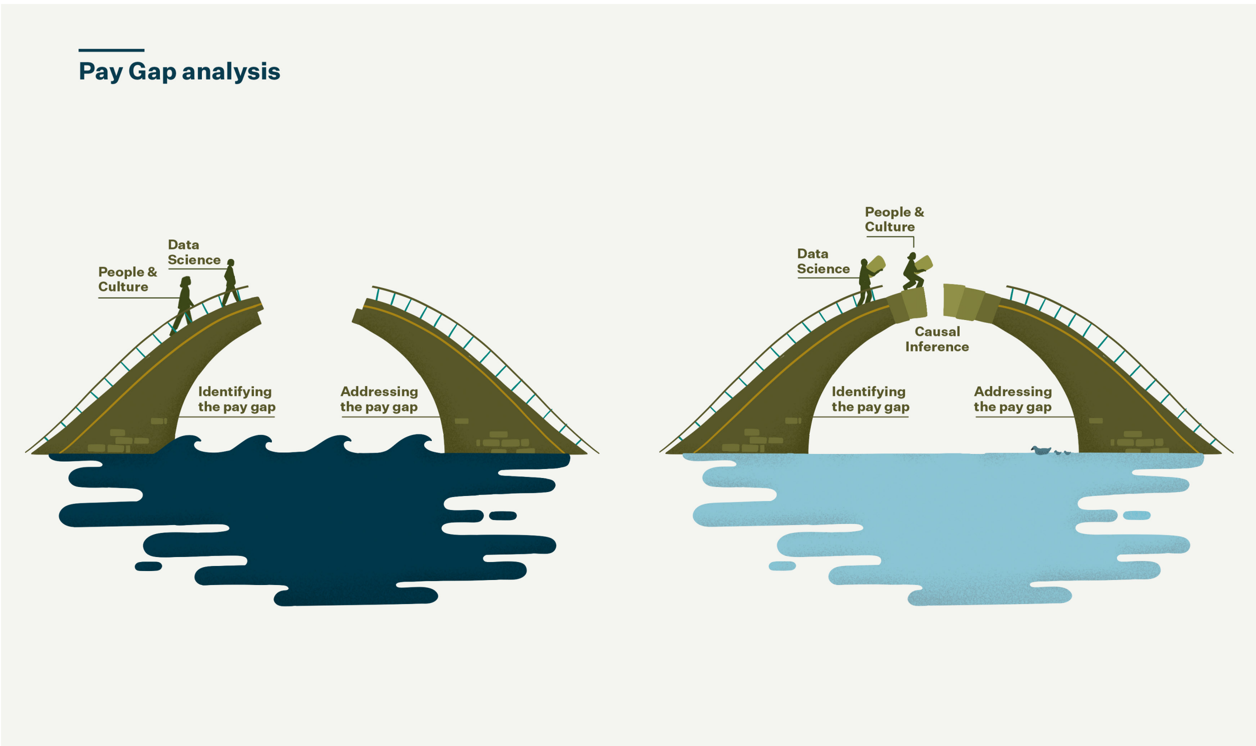


Figure 9: Collaboration between People and Culture and Data Science has the potential to move Aotearoa-New Zealand's research institutes from identifying pay gaps to addressing them.

may not accurately capture the relationship between gender and salary. A causal-inference approach can account for the complex interplay of factors controlling salary and to allow the testing of interventions designed to reduce pay gaps. Our results show that naïve approaches to designing such interventions are unlikely to lead to equitable pay structures. The modelling approaches demonstrated here do, however, provide a method for testing and designing such interventions.

The main conclusion of this study, though, is that collaboration between different parties interested in gender pay gaps, such as human resources and data science teams in CRIs, has the potential to result in a step change in pay equity in Aotearoa-NZ. Maintaining a focus on equity in times of crisis is crucial in order not to lose hard-won gains. We therefore strongly encourage implementation of these approaches on real datasets to improve the experiences of workers across the research sector.

Acknowledgements

The authors would like to acknowledge the members of Te Tira Maurikura and Te Pūnaha Matatini for numerous insightful conversations that informed the development of this manuscript. We thank Jean Donaldson and Toi Āria for design of the Figures 8 and 9 and the Complexity Community of Inquiry in Te Pūnaha Matatini for the funding which supported Toi Āria's contribution. Particular thanks to Ashley Mortensen, Linley Jesson, Will Godsoe and Olivia Truax for their reading of, and helpful comments on, the study. We would also like to thank two anonymous reviewers and the editor, Tara McAllister, for comments which helped to significantly improved this manuscript. Thank you finally to Ben Dickson for his hard work in formatting the manuscript.

References

Banks, J., Carson II, J. S., Nelson, B. L. and Nicol, D. M. (2014), *Discrete-Event System Simulation*, Prentice Hall International Series in Industrial and Systems Engineering, 5 edn, Pearson Education Limited, Edinburgh.

Barrett-Walker, T., Buelow, F., Te Atu O Tu MacDonald, L., Brower, A. and James, A. (2023), 'Stochastic Modelling of Intersectional Pay Gaps in Universities', *Royal Society Open Science* **10**(10), 230615. https://dx.doi.org/10.1098/rsos.230615

Bergmann, B. (1974), 'Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex', *Eastern Economic Journal* **1**(2), 103–110. https://www.jstor.org/stable/40315472

Bernard, R. E. and Cooperdock, E. H. G. (2018), 'No Progress on Diversity in 40 Years', *Nature Geoscience* **11**(5), 292–295. https://doi.org/10.1038/s41561-018-0116-6

Blackwell, L. (2001), 'Occupational Sex Segregation and Part-time Work in Modern Britain', *Gender, Work & Organization* **8**(2), 146–163. https://doi.org/10.1111/1468-0432.00126

Blinder, A. S. (1973), 'Wage Discrimination: Reduced Form and Structural Estimates', *The Journal of Human Resources* **8**(4), 436. https://doi.org/10.2307/144855

Brower, A. and James, A. (2020), 'Research Performance and Age Explain Less than Half of the Gender Pay Gap in New Zealand Universities', *PLOS ONE* **15**(1), e0226392. https://doi.org/10.1371/journal.pone.0226392

Brower, A. and James, A. (2023), 'Sticky Floors, Double-Binds, and Double Whammies: Adjusting for Research Performance Reveals Universities' Gender Pay Gap is Not Disappearing', *SAGE Open* **13**(3), 21582440231192323. https://doi.org/10.1177/21582440231192323

Bürkner, P.-C. (2017), '**Brms** : An *R* Package for Bayesian Multilevel Models Using *Stan*', *Journal of Statistical Software* **80**(1). https://doi.org/10.18637/jss.v080.i01

Caligaris, S. (2013), A Causal Graphs-based Approach for Assessing Gender Gaps with an Application to Child Health & Nutrition in China, Dottorato di Ricerca in Statistica, Universit'a degli Studi di Milano - Bicocca. https://boa.unimib.it/retrieve/e39773b2-a76d-35a3-e053-3a05fe0aac26/phd_unimib_744674.pdf

Came, H., Badu, E., Ioane, J., Manson, L. and McCreanor, T. (2020), 'Ethnic (Pay) Disparities in Public Sector Leadership in Aotearoa New Zealand', *International Journal of Critical Indigenous Studies* **13**(1), 70–85. https://doi.org/10.5204/ijcis.v13i1.1331

Chapman, S. J. and Benis, N. (2017), 'Ceteris Non Paribus: The Intersectionality of Gender, Race, and Region in the Gender Wage Gap', *Women's Studies International Forum* **65**, 78–86. https://doi.org/10.1016/j.wsif.2017.10.001

Christenfeld, N. J. S., Sloan, R. P., Carroll, D. and Greenland, S. (2004), 'Risk Factors, Confounding, and the Illusion of Statistical Control', *Psychosomatic Medicine* **66**(6), 868–875. https://doi.org/10.1097/01.psy.0000140008.70959.41

Cinelli, C., Forney, A. and Pearl, J. (2024), 'A Crash Course in Good and Bad Controls', *Sociological Methods & Research* **53**(3), 1071–1104. https://doi.org/10.1177/00491241221099552

Denz, R. and Meiszl, K. (2023), 'simDAG: An R Package to Simulate Simple and Complex (Longitudinal) Data from a DAG and Associated Node Information'. https://cran.r-project.org/web/packages/simDAG/index.html

El Achkar, S. (2023), 'Equal Pay for Work of Equal Value: Where Do We Stand in 2023?'.

https://ilostat.ilo.org/blog/equal-pay-for-work-of-equal-value-where-do-we-stand-in-2023/

Gaston, N. (2023), 'Starved of Funds and Vision, Struggling Universities Put NZ's Entire Research Strategy at Risk', *The Conversation* . https://theconversation.com/starved-of-funds-and-vision-struggling-universities-put-nzs-entire-research-strategy-at-risk-207708

Goldin, C. and Polachek, S. (1987), 'Residual Differences by Sex: Perspectives on the Gender Gap in Earnings', *The American Economic Review* **77**(2), 143–51. https://www.jstor.org/stable/1805442

International Labour Organisation (1951), 'Equal Remuneration Convention'. https://www.treaties.mfat.govt.nz/search/details/t/1465/

James, A. and Brower, A. (2022), 'Levers of Change: Using Mathematical Models to Compare Gender Equity Interventions in Universities', *Royal Society Open Science* **9**(9), 220785. https://doi.org/10.1098/rsos.220785

Karamessini, M. and Rubery, J. (2017), The Challenge of Austerity for Gender Equality in Europe: A Consideration of Eight Countries at the Center of the Crisis, *in* R. Connelly and E. Kongar, eds, 'Gender and Time Use in a Global Context', Palgrave Macmillan US, New York, pp. 51–74. https://doi.org/10.1057/978-1-137-56837-3_3

Karamessini, M. and Rubery, J. (2020), Prolonged Austerity and Gender Equality The Cases of Greece and the UK Compared, *in* A. Simonazzi, M. C. Marcuzzo, A. Palumbo and P. Villa, eds, 'Economic Policy, Crisis and Innovation: Beyond Austerity in Europe', Routledge Studies in the European Economy, Routledge, Taylor & Francis Group, London. https://doi.org/10.4324/9780429291142

Karamessini, M. and Rubery, J., eds (2013), *Women and Austerity*, 0 edn, Routledge. https://doi.org/10.4324/9780203066294

Laubach, Z. M., Murray, E. J., Hoke, K. L., Safran, R. J. and Perng, W. (2021), 'A Biologist's Guide to Model Selection and Causal Inference', *Proceedings of the Royal Society B: Biological Sciences* **288**(1943), 20202815. https://doi.org/10.1098/rspb.2020.2815

Lips, H. M. (2003), 'The Gender Pay Gap: Concrete Indicator of Women's Progress Toward Equality', *Analyses of Social Issues and Public Policy* **3**(1), 87–109. https://doi.org/10.1111/j.1530-2415.2003.00016.x

Lips, H. M. (2013), 'The Gender Pay Gap: Challenging the Rationalizations. Perceived Equity, Discrimination, and the Limits of Human Capital Models', *Sex Roles* **68**(3-4), 169–185. https://doi.org/10.1007/s11199-012-0165-z

McAllister, T. G., Naepi, S., Wilson, E., Hikuroa, D. and Walker, L. A. (2022), 'Under-Represented and Overlooked: Māori and Pasifika Scientists in Aotearoa New Zealand's Universities and Crown-Research Institutes', *Journal of the Royal Society of New Zealand* **52**(1), 38–53. https://doi.org/10.1080/03036758.2020.1796103

McAllister, T., Kokaua, J., Naepi, S., Kidman, J. and Theodore, R. (2020), 'Glass Ceilings in New Zealand Universities', *Mai Journal* **9**(3), 272–285. https://doi.org/10.20507/MAIJournal.2020.9.3.8

McCall, L. (2001), 'Sources of Racial Wage Inequality in Metropolitan Labor Markets: Racial, Ethnic, and Gender Differences', *American Sociological Review* **66**(4), 520. https://doi.org/10.2307/3088921

McElreath, R. (2018), *Statistical Rethinking: A Bayesian Course with Examples in R and Stan 2nd Edition*, Chapman and Hall/CRC. https://doi.org/10.1201/9780429029608

Mehta, L. (2013), Introduction, *in* 'The Limits to Scarcity: Contesting the Politics of Allocation', Routledge, pp. 1–8. https://doi.org/10.1201/9780429029608

Ministry of Business, Innovation, and Employment (1972), 'Equal Pay Act'. https://www.legislation.govt.nz/act/public/1972/0118/latest/DLM407770.html

Ministry of Business, Innovation, and Employment (1992), 'Crown Research Institutes Act'. https://www.legislation.govt.nz/act/public/1992/0047/latest/dlm265144.html

Ministry of Business, Innovation, and Employment (2022), 'Te Ara Paerangi Future Pathways White Paper'. https://www.mbie.govt.nz/assets/te-ara-paerangi-future-pathways-white-paper-2022.pdf

Ministry of Business, Innovation and Employment (2024), 'Terms of Reference: Science System Advisory Group'. https://www.mbie.govt.nz/dmsdocument/28134-terms-of-reference-science-skills-advisory-group

Morton, J. (2024), 'GNS Science Proposes Job Cuts in Major Shake-Up', *NZ Herald* . https://www.nzherald.co.nz/nz/sad-day-for-science-job-cuts-proposed-at-gns-science-77-roles-affected/4QC32LHSKNGHLEEVHK6BJCG6SM/

Naepi, S., McAllister, T. G., Thomsen, P., Leenen-Young, M., Walker, L. A., McAllister, A. L., Theodore, R., Kidman, J. and Suaaliia, T. (2020), 'The Pakaru 'Pipeline': Māori and Pasifika Pathways within the Academy', *The New Zealand Annual Review of Education* **24**, 142–159. https://doi.org/10.26686/nzaroe.v24i0.6338

National Party and ACT New Zealand (2023), 'Coalition Agreement New Zealand National Party & ACT New Zealand'.
https://www.nzdoctor.co.nz/sites/default/files/2023-11/National_ACT_Agreement.pdf

National Party and New Zealand First (2023), 'Coalition Agreement New Zealand National Party & New Zealand First'.
https://www.nzfirst.nz/coalition-agreement

Oaxaca, R. (1973), 'Male-Female Wage Differentials in Urban Labor Markets', *International Economic Review* **14**(3), 693.
https://doi.org/10.2307/2525981

Parker, J. and Donnelly, N. (2020), 'The Revival and Refashioning of Gender Pay Equity in New Zealand', *Journal of Industrial Relations* **62**(4), 560–581.
https://doi.org/10.1177/0022185620929374

R Core Team (2024), 'R: A Language and Environment for Statistical Computing', R Foundation for Statistical Computing.
https://www.r-project.org/

Roos, P. A. (1985), *Gender and Work: A Comparative Analysis of Industrial Societies*, Suny Press.

Ross, J. (2024), 'Are Better Days Ahead for New Zealand's Storm-Lashed Universities?', *Times Higher Education* .
https://www.timeshighereducation.com/depth/are-better-days-ahead-new-zealands-storm-lashed-universities

Sensoy, Ö. and DiAngelo, R. (2017), '"We Are All for Diversity, but . . .": How Faculty Hiring Committees Reproduce Whiteness and Practical Suggestions for How They Can Change', *Harvard Educational Review* **87**(4), 557–580.
https://doi.org/10.17763/1943-5045-87.4.557

Shaw, A. K. and Stanton, D. E. (2012), 'Leaks in the Pipeline: Separating Demographic Inertia from Ongoing Gender Differences in Academia', *Proceedings of the Royal Society B: Biological Sciences* **279**(1743), 3736–3741.
https://doi.org/10.1098/rspb.2012.0822

Smith, M., Layton, R. and Stewart, A. (2017), 'Inclusion, Reversal, or Displacement: Classifying Regulatory Approaches to Pay Equity', *Comparative Labor Law & Policy Journal* **39**(1), 211.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2858046

Stan Development Team (2023), 'Stan Modeling Language Users Guide and Reference Manual'.
https://mc-stan.org/docs/2_36/reference-manual-2_36.pdf

StatsNZ (2020), *Organisational Gender Pay Gaps: Measurement and Analysis Guidelines*, Stats NZ Tatauranga Aotearoa, Wellington, N.Z.
https://www.stats.govt.nz/assets/Uploads/Methods/Organisational-gender-pay-gaps-measurement-and-analysis-guidelines/organisational-gender-pay-gaps-measurement-analysis-guidelines.pdf

StatsNZ (2024), 'Labour Market Statistics (Income): June 2024 Quarter'.
https://www.stats.govt.nz/information-releases/labour-market-statistics-income-june-2024-quarter/

Te Kawa Mataaho (2021), 'Kia Toipoto'.
https://www.publicservice.govt.nz/system/public-service-people/pay-gaps-and-pay-equity/kia-toipoto

Te Kawa Mataaho (2023), 'Gender Deep Dive'.
https://www.publicservice.govt.nz/assets/DirectoryFile/Deep-Dive-Rainbow-Gender.pdf

Textor, J., Van Der Zander, B., Gilthorpe, M. S., Liśkiewicz, M. and Ellison, G. T. (2017), 'Robust Causal Inference Using Directed Acyclic Graphs: The R Package 'Dagitty'', *International Journal of Epidemiology* p. dyw341.
https://doi.org/10.1093/ije/dyw341

Wan, X., Wang, W., Liu, J. and Tong, T. (2014), 'Estimating the Sample Mean and Standard Deviation from the Sample Size, Median, Range and/or Interquartile Range', *BMC Medical Research Methodology* **14**(1), 135.
https://doi.org/10.1186/1471-2288-14-135

Wong, Y. L. A. and Charles, M. (2020), Gender and Occupational Segregation, *in* 'Companion to Women's and Gender Studies', pp. 303–325.
https://doi.org/10.1002/9781119315063.ch16