I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12896

Not Much Bounce in the Springboard: On the Mobility of Low Pay Workers

Gail Pacheco
Alexander T. Plum
Peter J. Sloane

JANUARY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12896

Not Much Bounce in the Springboard: On the Mobility of Low Pay Workers

Gail Pacheco
NZWRI (Auckland University of Technology)

Alexander T. Plum
NZWRI (Auckland University of Technology)

Peter J. Sloane
Swansea University, Adelaide University and IZA

JANUARY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Not Much Bounce in the Springboard: On the Mobility of Low Pay Workers*

Estimating economic earnings mobility is imperative for understanding the degree to which low pay employment is a temporary or long-term position. The current literature estimates transition probabilities between low and higher pay. This study extends the focus to identify the underlying pecuniary wage change via construction of an intermediate pay zone marginally above low pay. Utilising monthly administrative data we find that individuals with a strong attachment to the low pay sector have a very low probability of shifting into higher pay. Further, these individuals also have a substantially greater risk of experiencing a low pay-no pay cycle relative to those who are intermediate or higher paid. Notably, this finding is only uncovered using within year variation in wages to reveal intensity of labour market attachment.

JEL Classification: J62, J31, C33, C55

Keywords: low pay dynamics, transition probability, state dependence, dynamic models, administrative data

Corresponding author:
Alexander T. Plum
New Zealand Work Research Institute
Auckland University of Technology
Private Bag 92006
Auckland 1142
New Zealand

E-mail: alexander.plum@aut.ac.nz

* We want to thank Lixin Cai for valuable comments.

1. Introduction

The discussion on growing inequality within advanced economies has intensified in recent years. Not only do many citizens consider high levels of inequality as a violation of their moral values, but it is also argued that inequality might have the potential to harm economic growth (IMF 2019). Estimating the degree of economic mobility clarifies whether belonging to the 'left behind' is a temporary or permanent (long-term) position. This discussion also establishes the research strand that looks at the labour market trajectories of the low-paid, their chances of entering higher pay and the risk of becoming unemployed. Understanding whether low-paid jobs open a gateway to higher-paid jobs is crucial for determining their potential to facilitate economic mobility, and much research has been conducted on this topic recently.

The majority of studies in this space draw a particularly positive picture of the earnings prospects of the low-paid. For example, Uhlendorff (2006) concludes for West German men 'that being low-paid does not have any adverse effects on future employment prospects' [p. 17]. Concerning the British labour market, Cai et al. (2018) present evidence for a stepping stone effect of low pay. Buddelmeyer et al. (2010) find for Australia that 'among men there appears to be no significant difference between low-paid and high-paid employment' [p. 46].[1]

The prevailing strategy in many prior studies is to define the wage categories of low pay and higher pay using a sharp threshold. One consequence of this approach is that it is unclear what the magnitude of the wage change is when transitioning between the categories of low and higher pay. To investigate the scale of transition, this study utilises a buffer wage range (labelled *intermediate-paid*) that covers wages that are marginally above the low pay threshold. This permits closer inspection as to the likelihood of making a significant shift out of low pay, relative to those that shift around the threshold point at the upper boundary of low pay.

To estimate intertemporal labour market transitions, the annual mean monthly wage constitutes our outcome variable.[2] The individuals' position in the wage distribution is used to separate the sample into low-paid (lowest decile), intermediate-paid (between lowest decile and first quartile) and higher-paid (above the first quartile). Following the prevailing identification strategy in the low pay literature, we include as regressor the lagged labour market position (labelled as *Base model*).

Additionally, we know from recent research from Pacheco & Plum (2019) that earnings prospects are highly heterogeneous across the low-paid and negatively associated with their attachment to the low pay sector. More specifically, Pacheco & Plum (2019) utilise population-wide monthly administrative data and define low pay attachment (as a categorical variable) by the relative share of months employed in the low pay sector. The categories used are either not

[1] Similar findings are presented by Clark & Kanellopoulos (2013) and Mosthaf (2014), in which the authors show that the on average a low-paid worker has a higher probability of moving between two consecutive periods from low pay into higher pay than rather staying on low pay.
[2] We use annual mean monthly wages instead of time-point specific wages (e.g., monthly wages) to reduce the impact of transitory monthly wage shocks.

working at all in the low pay sector; working for less than half of the employed months in the low pay sector (weak low pay attachment); or working for at least half of the employed months in the low pay sector (strong low pay attachment). The authors show that the employment prospects for those with strong low pay attachment differs significantly from those with a weak low pay attachment. We therefore incorporate an adjusted version of their concept by replacing the lagged dependent variable in the second model (labelled as *Intensity model*) by intensity markers which refer to the relative share the individual was employed in one of the three labour market positions (low, intermediate or high pay). To generate these markers, we calculate the individual's position in the wage distribution for each employed month and therefore provide a very granular description of their employment history.

The labour market transitions are estimated using the standard approach in the literature of a dynamic multinomial random effects logit model (for a detailed discussion, see Cai (2019)). We find striking differences across various dimensions between the *Base* and the *Intensity models*:

- The *Base model* indicates that, compared to higher pay, low pay reduces the chance of being on higher pay in the next year by on average 20 percentage points.[3] This probability differential is exacerbated in the *Intensity model*: being on low pay in all employed months lowers the chance of receiving a mean monthly wage that belongs to the upper three quartiles on average by 86 percentage points compared to being on higher pay in all employed months.
- In the *Base model*, we do not find any significant difference in the effect of experiencing low pay and intermediate pay on the chances of becoming higher-paid. In the *Intensity model*, being on intermediate pay for all employed months lifts the chance of receiving a higher pay in the next year by on average 25 percentage points compared to being on low pay in all employed months.
- When we compare goodness-of-fit statistics, we find that compared to the *Base model*, the *Intensity model* has a five percentage points higher share of correct predictions.
- While the explanatory power of the lagged labour market related variables increase when moving from the *Base model* to the *Intensity model*, the explanatory power of the remaining covariates (e.g., qualification, ethnicity) declines.

To complete our study, we analyse the relationship between labour market position and employment stability. With few exceptions, the empirical literature does not find evidence for a low pay-no-pay cycle, that is the risk of future unemployment is not found to be significantly different between those on low pay and those on higher pay (e.g., Buddelmeyer et al. 2010, Cai et al. 2018). In the *Base model*, we also find very small effects of past low pay employment, defined as a binary indicator, on the chances of experiencing non-employment spells. However, the *Intensity model* indicates that being continuously on low pay significantly lowers the chance

[3] This number is in line with what is found in the economic literature. For example, Cai et al. (2018, Table 3) calculated for British males using the BHPS that on average those on low pay have a 20.0 percentage points lower probability turning into higher pay than those who were already higher-paid.

of being continuously employed in the future on average by 13.6 percentage points compared to when being continuously higher-paid employed.

Our findings indicate that low pay employment *itself* seems not to foster human capital accumulation to improve future earning progression. This finding is in line with the conclusion of Stewart (2007): 'If unemployed individuals' employment prospects are to be permanently improved, they need to find jobs where they can augment their skills (for example, through on-the-job training), raise their productivity and move up the pay distribution.' [p. 529] – and this extends to the low-paid.

The remainder of this paper is structured as follows: Section 2 contains a brief summary of the literature advancements in this space in the last two decades, section 3 encompasses an overview of the administrative data and key descriptives; Section 4 presents the econometric model; while results are shown in Section 5, followed by conclusions.

2. Literature review

The seminal paper of Stewart & Swaffield (1999) shaped the understanding of why analysing the degree of persistence in low pay is of high relevance. Though cross-sectional figures might indicate that inequality is on the rise, it is unclear how it affects individuals' earnings mobility as either 'incidence of permanent low pay has increased' or 'transitory fluctuations in earnings have increased' [Stewart & Swaffield 1999, p. 24]. In a worst-case scenario, low pay workers might be trapped in a 'low pay no-pay cycle'.

In the past two decades, analysing the degree of persistence in low pay and the inter-relation of low pay and unemployment gained much attention in the economic literature. There are numerous low pay studies analysing the labour market dynamics for different countries, periods, and different definitions of low pay.[4] The respective findings are summarized as the following:

1) Uhlendorff (2006), Cappellari (2007), Clark and Kanellopoulos (2013), Cai (2014), Mosthaf (2014), Fok et al. (2015) and Cai et al. (2018) present evidence that being on low pay in the past genuinely increases the likelihood of experiencing low pay in the future.
2) Numerous studies also show that low-paid jobs 'lead to a higher-paid job in the future' (Uhlendorff 2006, p. 18). In particular, it is often found that the probability of moving to higher pay is higher than the probability of experiencing further instances of low pay. Therefore, low-paid jobs are deemed as 'stepping stones' (Cai et al. 2017, p. 283).[5]

[4] Theoretical explanations on the impact of low pay are scarce and can be broadly divided into the two groups of human capital accumulation and signalling theory. Due to its unclear theoretical impact, it is a research topic that needs to be assessed empirically.

[5] See also Cai (2014) or Mosthaf (2014). Though not explicitly stating it, the study of Clark & Kanellopoulos (2013) shows that the probability moving from low pay to higher pay is exceeded substantially by the risk staying on low pay.

3) Regarding the future risk of experiencing unemployment, findings are not consistent across the literature. For example, using the first six waves (1991-1996) of the British Household Panel Survey (BHPS) Stewart (2007, p. 511) finds that low-wage employment has 'almost as large an adverse effect as unemployment on future prospects and the difference in their effects is found to be insignificant'. However, Cai et al. (2018), who make use of the 18 waves of BHPS, do not find any evidence in support of a low pay no-pay cycle.

Though there have been numerous empirical investigations of low pay, common features of these studies are that no further differentiation of the higher-paid group is undertaken. However, Plum (2019) has recently shown that for the British labour market there is a mass of observations around the low pay threshold. Moreover, that wages vary substantially within the higher-paid group, ranging from one to four times the cut off value of the low pay threshold. These descriptive findings for the British labour market raise the question of whether the majority of the transitions between low and higher pay are driven by (small) pecuniary changes around the low pay threshold.

3. Data and descriptive statistics

3.1. Data

In our empirical analysis we use administrative data from Statistics NZ's Integrated Data Infrastructure (IDI).[6] The IDI contains population-wide longitudinal microdata about individuals, households, and organizations. These data are sourced from government, and non-government agencies, as well as Statistics NZ surveys. The data are confidentialised by means of assigning a unique identifier to each individual.

The spine of the IDI is the Central Linking Concordance (CLC), which contains a list of all unique identifiers, which enables the researcher to link across multiple data-sets within the IDI. For our purposes, we link the CLC with three sources – (1) birth record data from the Department of Internal Affairs (DIA), (2) monthly tax data from Inland Revenue (IR) to gauge income information over time, and (3) the 2013 Census survey, which provides a range of individual and household characteristics, such as educational attainment.

To reduce the impact of birth cohort-specific effects, we focus our analysis on males born in 1975[7]. Our period of analysis covers ten years, starting from 2005. Therefore, the age range of our sample is 30 to 40 years. This reduces the influence of new labour market entrants, as well as those exiting for retirement purposes.[8] We also use DIA birth record information to identify birth year of each of their children. Next, we link these individuals with information provided

[6] See the appendix for details of the disclaimer associated with use of the IDI.
[7] In a robustness analysis, we reran our estimation for each birth cohort born between 1970 and 1980 but findings were consistent across all cohorts.
[8] Compared to prime-aged worker, labour market entrants and those close to the retirement age might differ not only in their wage progression but most likely also in their unobservable characteristics.

by the 2013 Census on the ethnicity to identify members of the three largest ethnic groups, New Zealand European, Māori, and Pacific peoples.

The 2013 Census also provides individual information on the highest qualification.[9] When determining the wage distribution, we use the whole sample, regardless of educational attainment. However, in our empirical estimation we drop the most highly qualified individuals. This exclusion from the population of interest is necessary, as the underlying assumption of the applied random effects model is that the covariates are uncorrelated with the individual-specific time-invariant error term. The random effects capture unobservable differences in aspects like motivation or ability and thus is likely linked with the individuals' qualification – and therefore might violate the independence assumption. As a consequence, in our estimation sample, we form three qualification-related categories: no school qualification, Level 1-4, and Level 5 or 6 (these are all below a bachelor's level qualification).

Finally, we use information from IR on wages and salaries. For our analysis, we use monthly gross earnings before tax that come from wages and salaries. IR records are on an individual-employer level, and as an employee might be holding multiple jobs or change his job during a month, there could be more than one IR record entry per month per individual. For simplification, we aggregate wages across all employers for each month. Our observation window spans from 2005 to 2015, and we keep individuals with seven or more consecutive years of income from wages and salaries who have at least one month of employment per year.[10] Furthermore, IR records provide information on whether the individual received income from additional sources[11]. We therefore include information on income from benefits and ACC (Accidents Claimants Compensation) in our regression.

3.2. Descriptive statistics

Our final sample consists of 67,251 observations, and 86 percent are of NZ European ethnicity, 11 percent Māori, and 3.5 percent Pacific peoples (see Table 1).[12] With respect to educational qualification, roughly every fifth individual has no qualification, and every second between Level 1 and 4. Three out of ten have Level 5 or 6 qualifications. In terms of labour market status, we observe that 85 percent of our population of interest are employed throughout the year and just 3.4 percent for less than six months. Further, the vast majority did not receive any benefits or ACC claims.

---

[9] One drawback is that we do not observe changes in qualification. To minimize the relevance of this aspect, we restricted the sample to individuals of age 30 to 40 with a minimum number of employment spells per year (see explanation on the identification of employment spells further below).

[10] As the number of non-employment spells might be endogenous, we re-ran the estimation and trimmed the sample to continuously employed; the findings were not affected.

[11] In total, IR records have information on the following seven income sources: (1) wages and salaries, (2) benefits, (3) ACC claims, (4) paid parental leave, (5) withholding payments, (6) pensions, and (7) student loans. As the first three of these sources dominate, we do not include information on income sourced from avenues (4) to (7).

[12] As an individual might state multiple ethnicities, we prioritise them, giving highest priority to Māori, then Pacific peoples and last NZ European (we do not include individuals who state being of Asian, MELAA or other ethnicity).

Table 1: Descriptive statistics

	N	*Share*
Ethnicity		
NZ European	57,681	85.8%
Māori	7,188	10.7%
Pacific peoples	2,382	3.5%
Qualification		
No qualification	12,384	18.4%
Level 1-4	34,605	51.5%
Level 5-6	20,262	30.1%
Months employed		
Emp: 12 months	57,558	85.6%
Emp: 10-11 months	4,224	6.3%
Emp: 6-9 months	3,207	4.8%
Emp: <6 months	2,262	3.4%
Number of children		
No children	23,445	34.9%
1 child	13,860	20.6%
2 children	18,264	27.2%
3 children	8,112	12.1%
4+ children	3,570	5.3%
Benefit recipient		
No benefits	63,498	94.4%
Benefits 1-6 months	2,190	3.3%
Benefits >6 months	1,563	2.3%
ACC		
No ACC	64,278	95.6%
ACC 1-3 months	2,229	3.3%
ACC >3 months	744	1.1%

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251

There exist different strategies to identify low pay[13] and in this study, we apply a relative approach (akin to the relative thresholds used by others, e.g. Uhlendorff (2006), Cappellari (2007) and Cai (2014)). We define men with mean monthly earnings belonging to the 10th lowest percentile as low-paid. Then, we define those being on intermediate pay when receiving a mean monthly wage above the low pay threshold but belonging to the lowest quartile. Finally, those with a mean monthly wage that belongs to the top three quartiles are defined as higher-paid. Mean monthly earnings are calculated on an annual basis and thus, we have up to ten labour

[13] The OECD (1997) sets the low pay threshold at two thirds of the annual earnings. However, this approach does not offer any straight forwarded pattern to define those on intermediate pay. Moreover, as we do not have any information on working hours, changes in the monthly wages might be due to an adjustment in the working time. However, as OECD Stats homepage (https://stats.oecd.org/) data pointed out, approximately 95 percent of prime aged men in New Zealand are working full time.

market state observations per individual. Note that the three labour market positions are mutually exclusive and an individual can only be in one of them.

Table 2: Labour market transition

	Low pay$_t$	Intermediate pay$_t$	Higher pay$_t$	Share$_{t-1}$
Low pay$_{t-1}$	52.2%	26.5%	21.3%	11.0%
Intermediate pay$_{t-1}$	16.5%	49.4%	34.1%	17.7%
Higher pay$_{t-1}$	3.5%	8.7%	87.8%	71.3%
Share$_t$	11.2%	17.9%	71.0%	

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251

Table 2 provides an overview of the labour market distribution[14] and shows the transition probabilities between labour market states across two consecutive years. We can see that 21.3 percent of those on low pay at $t-1$ manage to climb up the salary ladder and move into higher pay at t. However, every second low pay worker is found to remain in the low pay sector in the succeeding year.

Deriving broad indicators of the labour market position is a standard approach in the econometric literature and is often simply a result of restrictions regarding the availability of further information. However, Pacheco & Plum (2019) show that wages are especially volatile at the lower tail of the earnings distribution. The IR data records wages and salaries on the monthly level, and we use this information to derive for each individual the month specific position in the wage distribution. Using this information, we generate relative indicators on the intensity of the respective labour market position for each year.

We consider three labour market positions: j is equal to 1 if the individual is on low pay, 2 on intermediate pay and 3 on higher pay. For individual $i = 1, \dots, N$ in year t and month m, we introduce the following labour market related dummy variables:

$$\mathrm{em}_{it_mj} = \begin{cases} 1 & \text{employed in } j \\ 0 & \text{else} \end{cases} \tag{1}$$

and

$$\mathrm{em}_{it_m} = \begin{cases} 1 & \text{employed} \\ 0 & \text{else} \end{cases} \tag{2}$$

Based on this, we construct the following indicator:

$$\mathrm{int}_{itj} = \frac{\sum_m \mathrm{em}_{it_mj}}{\sum_m \mathrm{em}_{it_m}} \tag{3}$$

[14] As the sample is unbalanced, the share of the respective labour market status deviate from the cut-off points.

By construction, $\text{int}_{itj} \in \{0, \ldots, 1\}$ and reflects individual i's relative share of employed months in year t working in the labour market position j. Table 3 shows the distribution of the labour market intensity variables. On average, those on low pay at t spend 61.4 percent of their employed months in the low pay sector. Conversely, those on low pay t spend 20.9 percent of their employed months receiving a higher-paid wage. Looking at those individuals who were higher-paid in t, much less heterogeneity is found: on average, they receive a higher-paid wage in 85.4 percent of their employed months.

Table 3: Labour market intensity

	Low pay intensity	Intermediate pay intensity	Higher pay intensity
Low pay$_t$	0.614	0.178	0.209
	(0.367)	(0.219)	(0.292)
Intermediate pay$_t$	0.252	0.445	0.303
	(0.296)	(0.272)	(0.298)
Higher pay$_t$	0.030	0.116	0.854
	(0.103)	(0.209)	(0.247)

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Std dev in brackets.

Furthermore, we find that for those on low pay in t, 38 percent spend less than fifty percent of their employed months in the low pay sector – while for 62 percent, they spend a minimum of every second employed month being low-paid. To get an impression of how the transition probability into higher pay differs between both groups, Table 4 differentiates the first row of Table 2 accordingly. We find substantial differences: those with a weak low pay intensity in $t-1$ (Low pay intensity$_{t-1}$<50%) have a five times higher share of transitions into higher pay than those with a stronger low pay attachment (Low pay intensity$_{t-1}$≥50%). Likewise, noticeable differences are found with respect to the probability of staying on low pay: it is more than two times higher for those with a strong low pay intensity.

Table 4: Labour market transition and low pay intensity

	Low pay$_t$	Intermediate pay$_t$	Higher pay$_t$	Share$_{t-1}$
Low pay$_{t-1}$ and				
Low pay intensity$_{t-1}$< 50%	29.0%	30.7%	40.4%	40.0%
Low pay intensity$_{t-1}$≥ 50%	67.8%	23.7%	8.6%	60.0%

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. By constructing, there are no observations with 'Low pay intensity$_t$=0' and being identified as 'Low pay$_t$'.

4. Econometric specification

4.1. Base model

To identify state dependence in low pay, we apply a dynamic random effects multinomial logit model that has been established in numerous other low pay studies (e.g. Uhlendorff (2006), Mosthaf (2014), Fok et al. (2015), Cai et al. (2018), Pacheco & Plum (2019)). Note that the outcome variable ($y_{it} = j$) refers to the labour market position of individual i in year t, which is derived on i's position in the mean monthly wage distribution. This implies that the individual does not have to be employed in each month. The second part of the study deals with the interrelation of labour market position and number of employed months.

Thus, the probability of individual i being in the labour market state y_{it} at time point $t \in \{1, \dots, T\}$ can be written as:

$$P(y_{it} = j | y_{it-1}, X_i, \alpha_{ij}) = \frac{exp(X_i' \beta_j + y_{it-1}' \gamma_j + \alpha_{ij})}{\sum_{k=1}^{3} exp(X_i' \beta_k + y_{it-1}' \gamma_k + \alpha_{ij})} \quad (4)$$

X_i refers to a vector of explanatory variables. These include categorical variables on ethnicity (NZ European, Māori, Pacific peoples), educational attainment (no qualification, Level 1-4, Level 5 or 6) [15], number of employed months in $t-1$ (12 months, 10-11 months, 6-9 months, <6 months), number of children (no children, 1 child, 2 children, 3 children, 4+ children), being a benefit recipient in $t-1$ (no benefits, 1-6 months receiving benefits, ≥7 months receiving benefits) and receiving ACC in $t-1$ (no ACC, 1-3 months receiving ACC, ≥4 months receiving ACC).[16] y_{it-1} is a vector of dummy variables with respect to the lagged labour market position. Additionally, as individuals may differ in unobservables such as motivation or ability (Heckman, 1981a), to control for unobserved heterogeneity, we include a time-invariant error term α_{ji}. As we are interested in studying the upward mobility of low pay, low pay is set as the reference category, and therefore coefficient vectors β_1, γ_1 and α_{i1} in equation (4) are set equal to zero.

Due to a correlation between the time-invariant error term and the initial conditions problem, the labour market position in the initial period might not be randomly distributed (Heckman 1981b). As Skrondal and Rabe-Hesketh (2014) have pointed out, not accounting for unobserved heterogeneity and its correlation with the initial labour market position might result in biased estimations. To take care of the "initial conditions problem", we follow the suggestion of Wooldridge (2005) by applying a conditional random-intercept model:

$$\alpha_{ji} = y_{i(t=0)}' \lambda + \kappa_{ji} \quad (5)$$

It is assumed that the random effects are normally distributed $\kappa_{ji} \sim N\left(0, \sigma_{\kappa_j}^2\right)$ and are correlated by ρ_κ. After substituting equation (8) into (7) the likelihood function for individual i takes the following form:

$$L_i = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \prod_{j=2}^{3} \left\{ \frac{exp\left(X_i' \beta_j + y_{it-1}' \gamma_j + y_{i(t=0)}' \lambda + \kappa_{ji}\right)}{1 + \sum_{j=2}^{3} exp\left(X_i' \beta_j + y_{it-1}' \gamma_j + y_{i(t=0)}' \lambda + \kappa_{ji}\right)} \right\}^{d_{ijt}} f(\kappa) d\kappa \quad (6)$$

[15] As robustness, we re-ran separate regression based on ethnicity (for NZ European and Māori) and educational attainment but without any major impact on the findings.
[16] We have not included control variables on job characteristics (e.g. industry of the employer, firm seize) as individuals might change employer/occupation during the year.

Note that d_{ijt} equals 1 if individual i is in state j at time point t and zero otherwise. To integrate out the random effects, we use maximum simulated likelihood (MSL). Using random numbers based on prime numbers (also called Halton draws, see Train 2009), two times R standard uniform distributed draws are derived and transformed by the inverse cumulative standard normal distribution. For each draw, the likelihood is derived for each observation, multiplied over all individuals and time-points and finally averaged over all draws (using 75 draws):

$$MSL = \prod_{i=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{t=1}^{T_i} P_{it}(\kappa_1^r, \kappa_2^r) \right\} \quad (7)$$

4.2. Intensity model

In the intensity model, we replace the vector of the lagged labour market position y_{it-1} in equation (6) by an intensity marker $\text{int}_{i(t-1)j}$ which refers to the share of employed months individual i spent in $t-1$ in labour market position j:[17]

$$P(y_{it} = j | \text{int}_{i(t-1)j}, X_i, \xi_{ij}) = \frac{exp(X_i' \beta_j + \text{int}_{i(t-1)2} \vartheta_{j2} + \text{int}_{i(t-1)3} \vartheta_{j3} + \text{int}_{i(t-1)2} \text{int}_{i(t-1)3} \vartheta_{j4} + \xi_{ij})}{\sum_{k=1}^{3} exp(X_i' \beta_k + \text{int}_{i(t-1)2} \vartheta_{k2} + \text{int}_{i(t-1)3} \vartheta_{k3} + \text{int}_{i(t-1)2} \text{int}_{i(t-1)3} \vartheta_{k4} + \xi_{ij})} \quad (8)$$

To control for the initial conditions problem, we apply the following transformation:

$$\xi_{ij} = \text{int}_{i(t=0)2} o_{j2} + \text{int}_{i(t=0)3} o_{j3} + \text{int}_{i(t=0)2} \text{int}_{i(t-1)3} o_{j4} + \varepsilon_{ji} \quad (9)$$

We assume that the random effects are normally distributed $\varepsilon_{ji} \sim N\left(0, \sigma_{\varepsilon_j}^2\right)$ and the correlation parameter is ρ_ε. To derive the likelihood contribution of individual i, equation (12) needs to be plugged into (11). Again, we use MSL to integrate out the random-effects.

5. Results

5.1. Model performance

As interpreting the coefficients is not straightforward, or meaningful for the research objectives at hand, they are not discussed in the main text of this paper. However, the relevant coefficients and estimation results are provided in the Appendix (see Table A 1 for the *Base model* and Table A 2 for the *Intensity model*).

An assumption underlying both models is that random effects capture the unobserved heterogeneity between individuals. As shown in Table 5, irrespective of the model used, we can see that the variances of both random effect error terms $\left(\sigma_{\kappa_j}^2, \sigma_{\varepsilon_j}^2\right)$ are significantly different from zero at the 1 percent level and that they are positively correlated. When comparing across the

[17] To compare directly the base model with the labor intensity model across, we did not change the outcome variable. This is a potential open task for future research.

base and intensity model, we find that the variances of the random effects with respect to the probability becoming higher-paid employed $\left(\hat{\sigma}^2_{\kappa_3} > \hat{\sigma}^2_{\varepsilon_3}\right)$ are significantly (at the 1 percent level) larger in the base model. This finding indicates that some of the persistent differences between individuals was attributed to unobserved individual characteristics in the base model and is now captured by including a more detailed lagged labour market position in the intensity model. We also find a significantly (at the 1 percent level) stronger positive correlation $(\hat{\rho}_\varepsilon > \hat{\rho}_\kappa)$ of the random effects in the intensity model.

Table 5: Random effects

	Base Model		*Intensity model*
$\hat{\sigma}^2_{\kappa_2}$	0.871	$\hat{\sigma}^2_{\varepsilon_2}$	0.877
	(0.064)		(0.056)
$\hat{\sigma}^2_{\kappa_3}$	3.194	$\hat{\sigma}^2_{\varepsilon_3}$	2.312
	(0.146)		(0.097)
$\hat{\rho}_\kappa$	0.628	$\hat{\rho}_\varepsilon$	0.757
	(0.027)		(0.017)
N	67,251		67,251

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015.

Next, we look at several goodness-of-fit statistics to compare the performance of both models. As indicators, we have chosen the log-likelihood, the two information criteria AIC and BIC[18] and the share of correct predictions. Table 6 shows that the intensity model outperforms the base model in each category. For example, the share of correct predictions climbs from 74.6 percent in the base model to 79.4 percent in the intensity model.

Table 6: Goodness-of-fit statistics

	Base Model	*Intensity model*
Log Likelihood	-37,429.25	-33,724.21
AIC	74,944.5	67,542.41
BIC	75,336.5	67,970.87
Correct predictions	0.7458	0.7944
N	67,251	67,251

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015.

5.2. Transition probabilities into higher pay

The key objective here is to extend the literature investigating transition probabilities of the low-paid into higher pay, by focusing on the magnitude of wage change experienced via introduction of the 'intermediate-paid' category in our identification strategy.

[18] Note that the degrees of freedom (df) in the *Intensity model* is slightly higher as we included the interaction of the intensity variables: df = 43 (*Base model*), df = 47 (*Intensity model*)

To investigate the chances of moving into higher pay, we calculate the average partial effect of becoming higher-paid in t, differentiated according to the labour market position in $t-1$. For the base model, we use low pay at $t-1$ as the reference category. As Table 7 shows, those on higher pay have on average a 20 percentage points higher chance being higher-paid employed in the following period compared to someone who was low-paid.

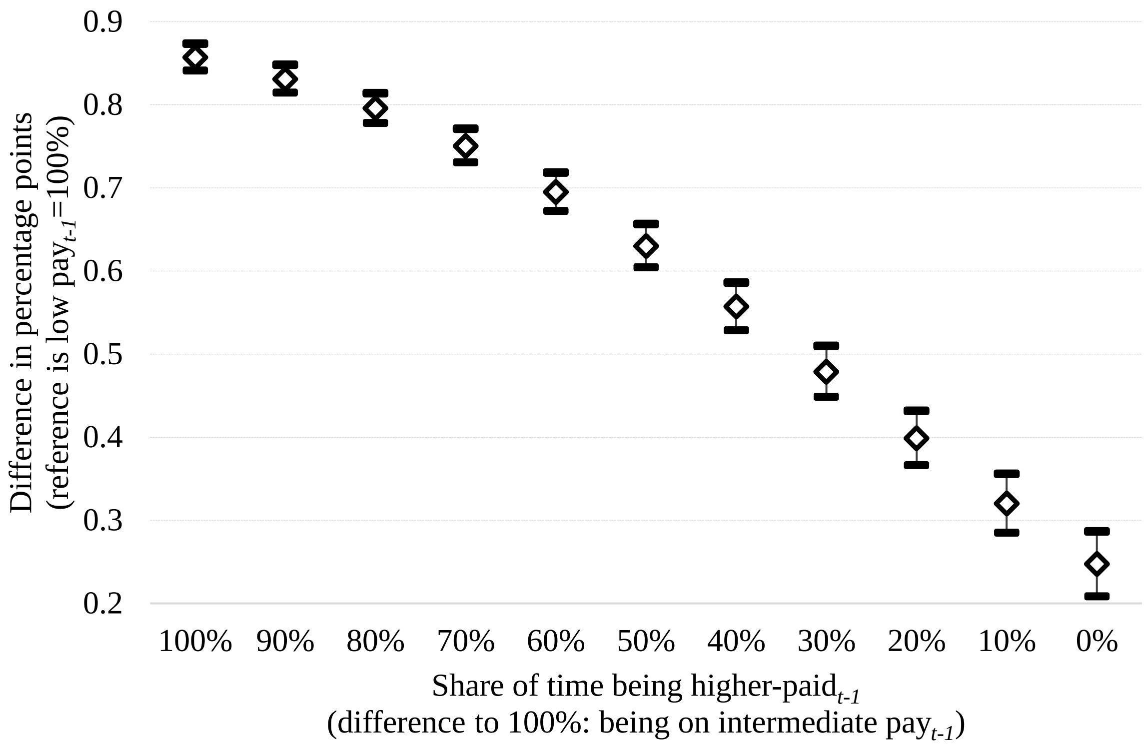
Table 7: Average partial effect of becoming higher-paid employed (Base Model)

Labour market position$_{t-1}$ (reference is *Low pay*$_{t-1}$)	
Intermediate pay$_{t-1}$	-0.006 [-0.024; +0.011]
Higher pay$_{t-1}$	+0.200 [+0.179; +0.221]

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Numbers in [] show the 95% conf. interval.

Next, we switch to the intensity model. As reference category, we choose being on low pay at $t-1$ in all of the employed months. First, we calculate the average partial effect becoming higher-paid in t for different levels of time spent being higher-paid in $t-1$, running from 100 percent to 0 percent in 10 percentage point increments. The remaining time spent (relative to 100 percent) is on intermediate pay. Therefore, the average partial effect of 0 percent refers to the case comparing being all the employed time on intermediate pay compared to all the employed time on low pay. Figure 1 visualizes the respective average partial effects, including the 95 percent confidence interval. The numbers differ fundamentally to those found in the base model. For example, those who were all the employed months higher-paid have, on average a 86 percentage points higher chance of staying on higher pay compared to when being on low pay in all employed months. This number drops to 25 percentage points when being on intermediate pay for all employed months.

Figure 1: Average partial effect of becoming higher-paid employed (Intensity Model)

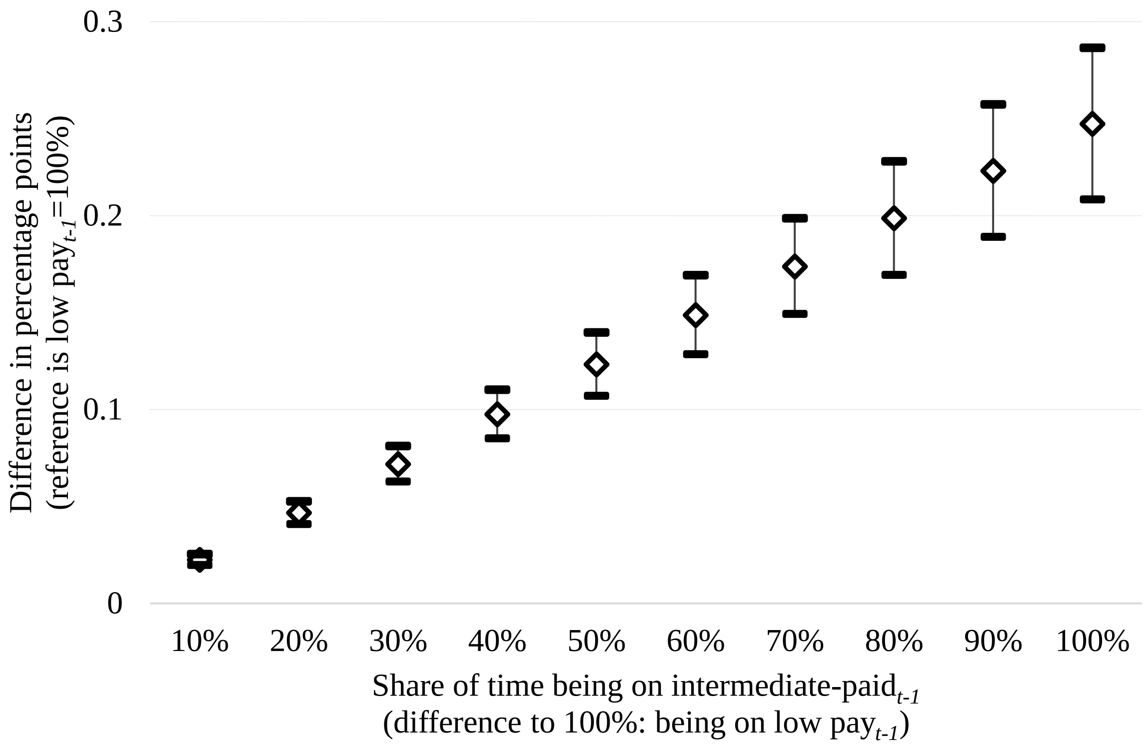


Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Lower and upper bar show the 95% conf. interval.

We also calculate the average partial effects becoming higher-paid between when working on intermediate / low pay in the previous period. As in the previous example, the reference category is being on low pay in the previous period in each employed month. The average partial effect is calculated compared to the case the individual was on intermediate pay for a certain share of the employed months and in the other months on low pay.

As Figure 2 shows, if the individual was receiving intermediate pay for 10 percent of their employed months and the remaining months on low pay, the average probability becoming higher-paid is 2 percentage points higher compared to being on low pay in all their employed months. Additionally, if the individual spent half of their employed months on intermediate pay, the average partial effect climbs to 12 percentage points.

Figure 2: Average partial effect of becoming higher-paid employed (Intensity Model)



Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Lower and upper bar show the 95% conf. interval.

5.3. Average partial effects of covariates

Table 6 indicated that including the past labour market position in a more granular manner improves the predictive power of the model. Next, we want to analyse how this affects the partial effects of the covariates. Table 8 shows the average partial effects of the main covariates on the probability becoming higher-paid employed. In line with our expectations and the economic literature in this space, we find that compared to NZ European, Māori and Pacific peoples have a lower chance becoming higher-paid employed. For example, in the base model, the probability declines on average by 5 percentage points for Māori (and 3.6 percentage points for Pacific peoples). However, when applying the intensity model, the APE drops substantially to 2.7 percentage points (1 percentage point). Comparable changes can be found when looking at the covariates with respect to qualification, receiving benefits or the employment intensity. For example, according to the base model, holding a Level 5 or 6 qualification improves on average the probability of experiencing higher pay employment in the subsequent period by 12.7 percentage points compared to having no qualification. Notably, in the intensity model, the average partial effect drops by two-third to 3.2 percentage points.

Moreover, we find that when moving from the base to the intensity model the average partial effect of some covariates are not significantly different from zero at the 5 percent level anymore. For example, being a benefit recipient at $t-1$ has on average no significant impact on the probability becoming higher-paid employed – this is in comparison to finding a strong significant impact in the base model.

Table 8: Average partial effects of covariates on becoming higher-paid employed

	Base Model			*Intensity model*		
	Mean	[95% conf. interval.]		Mean	[95% conf. interval.]	
Qualification (reference is no qualification)						
Level 1-4	0.091	0.073	0.108	0.033	0.023	0.044
Level 5-6	0.127	0.107	0.146	0.032	0.021	0.044
Months employed$_{t-1}$ (reference is employed for 12 months)						
10-11 months$_{t-1}$	-0.102	-0.119	-0.086	-0.048	-0.060	-0.036
6-9 months$_{t-1}$	-0.132	-0.153	-0.112	-0.062	-0.077	-0.047
<6 months$_{t-1}$	-0.180	-0.209	-0.152	-0.053	-0.072	-0.034
Benefit recipient$_{t-1}$ (reference is receiving no benefits)						
1-6 months$_{t-1}$	-0.044	-0.063	-0.026	0.004	-0.010	0.017
≥7 months$_{t-1}$	-0.113	-0.149	-0.076	0.008	-0.016	0.031
ACC$_{t-1}$ (reference is receiving no ACC)						
1-3 months$_{t-1}$	-0.030	-0.046	-0.014	-0.015	-0.028	-0.002
>3 months$_{t-1}$	-0.006	-0.033	0.022	0.042	0.021	0.062
Ethnicity (reference is NZ European)						
Māori	-0.050	-0.068	-0.031	-0.015	-0.027	-0.003
Pacific peoples	-0.036	-0.067	-0.006	-0.010	-0.029	0.010

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251.

5.4. The low pay – no pay cycle

Several studies have put their focus on the interrelation of low pay and no pay, asking whether individuals on low pay are stuck in a cycle of low pay and unemployment (e.g., Stewart 2007, Buddelmeyer et al. 2010, Cai et al. 2018). So far, findings in this strand of literature are mixed. Cai et al. (2018) state for Great Britain that 'those who are on low pay are roughly equally likely to transition into unemployment as those who are on higher pay' [p. 296] and they find the difference to be statistically insignificant. Buddelmeyer et al. (2010) findings for Australia are that 'on the probability of experiencing unemployment in the future (…) there appears to be no significant difference between low-paid and high-paid employment' [p. 46]. However, in contrast, Stewart (2007) shows in his study for the British labour market that 'low-wage employment is found to have almost as large an adverse effect as unemployment on future prospects and the difference in their effects is found to be insignificant' [p. 511]. One explanation for the different findings is how the group of the unemployed is defined. Stewart (2007) restricts his sample to the repeated unemployed who have 'an intervening spell of employment' [p. 520] and drops those with a continuing spell without employment.

When estimating labour market transitions, the above listed studies use as an identification strategy labour market-related information on the time point of interview to create mutually exclusive labour market positions. However, in our econometric model, we are not looking at a labour market outcome at a specific time-point but instead at mean monthly wages. Therefore, the relationship we analyse here is whether those on low pay have a greater risk of facing less stable employment patterns compared to those individuals on intermediate or higher pay.

One drawback of the administrative data is the limited information on whether the individual was searching for a job during non-employment or whether he turned towards self-employment. However, given that the construction of our sample panel is restricted to individuals that always have a minimum degree of labour market attachment (at least 1 month per year), there is limited likelihood of our population of interest containing individuals with long-term continuous non-employment spells. This definition is in line with the approach of Stewart (2007).

To disentangle the intertemporal relationship of non-employment and accounting for the impact of the position in the wage distribution, we use a dynamic random effects multi-nomial logit model to estimate state dependence in non-employment. Regarding our dependent variable, we construct a categorical employment intensity variable, which takes the value 1 if the individual was receiving wages and salaries for all 12 months in the respective year, the value 2 if receiving for 10 or 11 months and 3 if receiving for less than 10 months.[19] The explanatory variables include the lagged and initial categorical employment intensity variable.[20]

To estimate the impact of the past position in the wage distribution on the employment stability, we construct two models. The first model, which we label *Base model*, includes the labour market position based on the mean monthly wages for the period $t-1$ and $t=0$. The second model, which we label *Intensity model*, includes the intensity for each labour market position for the period $t-1$ and $t=0$ (including interaction effects). In both models, the reference category is receiving wages & salaries for 12 months at t. Estimation results can be found in the Appendix (see Table A 3 and Table A 4).

Before looking at the partial effects, we compare the goodness-of-fit statistics of both models (see Table 9). Regarding the log-likelihood and the two information criteria, we find marginal changes that indicate a better model fit for the intensity model. With respect to the share of correct predictions, the intensity model provides a slightly higher share, though the increase is just 0.2 percentage points. We know from the descriptive statistics that 85.6 percent of the observations are employed all twelve months; however, the share of correct predictions is very

[19] Due to a low number of observations, we had to collapse those being employed between 9 and 6 months and those for less than six months together.

[20] We also include the same covariates as in low pay models: ethnicity (NZ European, Māori, Pacific peoples), educational attainment (no qualification, Level 1-4, Level 5 or 6), number of children (no children, 1 child, 2 children, 3 children, 4+ children), being a benefit recipient in $t-1$ (no benefits, 1-6 months receiving benefits, ≥7 months receiving benefits) and receiving ACC in $t-1$ (no ACC, 1-3 months receiving ACC, ≥4 months receiving ACC),

close to that number, indicating that the statistical model hardly improves predictability of unemployment.

Table 9: Goodness-of-fit statistics
(Number of employed months)

	Base model	*Intensity model*
Log Likelihood	-28668.63	-28461.16
AIC	57427.27	57020.31
BIC	57837.5	57467.01
Correct predictions	0.8597	0.8620
N	67,251	67,251

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015.

Next, we compare the partial effect of the lagged number of employed months on the number of employed months in t of both models. We find in both models strong indications for state dependence and the differences between the models are small. For example, Table 10 shows that an individual who was employed for nine or less months in $t-1$ has in the base model (intensity model) on average a 18 percentage points (17.1 percentage points) higher risk being employed for nine or less months in the subsequent period compared to someone who was employed for all twelve months in the previous period.

Table 10: Average partial effect of number of employed months$_{t-1}$

	Base Model			*Intensity Model*		
	Months employed$_t$					
	12	*10-11*	*≤ 9*	*12*	*10-11*	*≤ 9*
Months employed$_{t-1}$ (reference is employed for 12 months)						
10-11 months$_{t-1}$	-0.103	+0.042	+0.060	-0.098	+0.040	+0.058
	[-0.119; -0.087]	[+0.032; +0.053]	[+0.049;+0.072]	[-0.113; -0.083]	[+0.030; +0.050]	[+0.047; +0.070]
≤9 months$_{t-1}$	-0.238	+0.058	+0.180	-0.224	+0.053	+0.171
	[-0.264; -0.212]	[+0.046;+0.070]	[+0.155; +0.205]	[-0.249; -0.199]	[+0.041; +0.065]	[+0.147; +0.196]

Notes: Data sourced from IDI (2018). Authors' calculations. Based on a random subsample of population of interest N = 62,484. Time period = 2007 to 2013. Numbers in parenthesis refer to standard errors.

Finally, we turn to the impact of the lagged position in the earnings distribution on the number of employed months. Starting with the base model, we can see in Table 11 that being on higher pay in the previous period on average increases the chances to being employed all 12 months in t and reduces the risk of being employed for less than ten months compared to being on low pay. However, the effect is very small: compared to low pay, being higher-paid in $t-1$ lifts the chances to be continuously employed in t on average by 1.9 percentage points.

Table 11: Average partial effect of months employed
(Base model)

	Months employed$_t$		
	12	*10-11*	*≤ 9*
Labour market position$_{t-1}$ (reference is *Low pay*$_{t-1}$)			
Intermediate pay$_{t-1}$	+0.002 [-0.006; +0.011]	-0.006 [-0.012;+0.000]	+0.004 [-0.002;+0.010]
Higher pay$_{t-1}$	+0.019 [+0.011; +0.028]	-0.013 [-0.019;-0.007]	-0.006 [-0.012;-0.001]

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Numbers in [] show the 95% conf. interval.

Finally, we derive the partial effects for the intensity model (see Table 12). Reference category is being on low pay in the previous period for all employed months. We compare the probability of the number of employed months in t with those being for a share of x months higher-paid and $(1-x)$ intermediate-paid in $t-1$. We find that those on higher pay for each employed month in the previous period have on average a 13.6 percentage points higher chance being employed for all 12 months in t compared to when being on low pay in each employed month. This number is rather stable across different levels of higher pay and intermediate-pay intensities. Further, we find that those on higher pay have a lower risk being employed for nine or less months in the preceding year compared to someone on low pay.

To sum up these findings, the intensity model reveals a stronger impact of the previous position in the wage distribution on the number of employed months compared to the base model.

Table 12: Average partial effect of months employed (Intensity model)

		Months employed$_t$		
Higher pay$_{t-1}$	*Intermediate pay$_{t-1}$*	*12*	*10-11*	*≤ 9*
100%	0%	+0.136 [+0.112;+0.158]	-0.067 [-0.083;-0.05]	-0.068 [-0.082;-0.054]
90%	10%	+0.126 [+0.104;+0.148]	-0.063 [-0.079;-0.046]	-0.064 [-0.078;-0.05]
80%	20%	+0.119 [+0.097;+0.140]	-0.059 [-0.075;-0.043]	-0.059 [-0.074;-0.046]
70%	30%	+0.112 [+0.092;+0.133]	-0.056 [-0.071;-0.039]	-0.056 [-0.07;-0.043]
60%	40%	+0.108 [+0.086;+0.128]	-0.054 [-0.068;-0.037]	-0.054 [-0.068;-0.039]
50%	50%	+0.104 [+0.085;+0.127]	-0.052 [-0.068;-0.037]	-0.052 [-0.067;-0.039]
40%	60%	+0.104 [+0.085;+0.127]	-0.052 [-0.068;-0.037]	-0.052 [-0.067;-0.039]
30%	70%	+0.108 [+0.086;+0.128]	-0.054 [-0.07;-0.037]	-0.054 [-0.068;-0.039]
20%	80%	+0.112 [+0.089;+0.134]	-0.056 [-0.071;-0.039]	-0.056 [-0.071;-0.041]
10%	90%	+0.118 [+0.093;+0.142]	-0.059 [-0.075;-0.041]	-0.059 [-0.075;-0.043]
0%	100%	+0.125 [+0.097;+0.150]	-0.061 [-0.082;-0.043]	-0.063 [-0.079;-0.046]

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251. Numbers in [] show the 95% conf. interval.

6. Conclusions

While there are numerous studies investigating the degree of persistence in low pay, not much attention has been paid to the question about where those individuals who exit the low pay sector move to and the magnitude of the wage change experienced for those moving out of low pay? For this reason, we introduce an intermediate pay zone that is marginally above low pay. To estimate intertemporal labour market transitions, we discuss two different identification strategies. The first one mirrors the prevailing method by looking at annual changes in the mean monthly wage. The second approach is to account for the intensity an individual was employed in a certain labour market position, especially with respect to the low pay sector.

Our database are linked New Zealand administrative data sources. We use birth record data from DIA to construct our population of interest, which are males born in 1975 (and their children). These are linked with the 2013 Census survey, which holds information on individual's characteristics. Finally, we add monthly income information provided from IR. We define men with a mean monthly wage at the lowest percentile as low-pay, those above but belonging to the lowest quartile as intermediate pay and higher-pay else. For the empirical analysis, we follow the standard approach in the economic literature by applying dynamic random effects multinomial logit models.

Our empirical models lead to contrasting conclusions. For example: the *Base model* indicates that (compared to higher pay) the average partial effect of low pay is -20 percentage points; while it is -86 percentage points in the *Intensity model* if the individual was employed in all months in the low pay sector. Using the goodness-of-fit statistics as a quality indicator, we find that the latter model has a five percentage points higher share of correct predictions compared to the first model. These stark differences are also found when looking at employment stability, showing that those with a high share of low-pay spells have an elevated risk of experiencing non-employment spells.

Overall, our findings indicate that in the New Zealand context low pay employment itself seems not to foster human capital accumulation sufficiently enough to improve future earning progression substantially and stabilize employment patterns. Future research should investigate whether this is a country specific finding or whether this phenomenon is present in other developed countries.

References

Buddelmeyer, H., Lee, W. S., & Wooden, M. (2010). Low-Paid Employment and Unemployment Dynamics in Australia. *Economic Record*, 86(272), 28-48.

Cai, L. (2014). State-Dependence and Stepping-Stone Effects of Low-Pay Employment in Australia. *Economic Record*, *90*(291), 486-506.

Cai, L. (2019). Is there a low-pay no-pay cycle in Australia? A note on Fok, Scutella and Wilkins (2015). *Empirical Economics*, 1-19.

Cai, L., Mavromaras, K., & Sloane, P. (2018). Low Paid Employment in Britain: Estimating State-Dependence and Stepping Stone Effects. *Oxford Bulletin of Economics and Statistics*, *80*(2), 283-326

Cappellari, L. (2007). Earnings mobility among Italian low-paid workers. *Journal of Population Economics*, 20(2), 465-482.

Clark, K., & Kanellopoulos, N. C. (2013). Low pay persistence in Europe. *Labour Economics*, 23, 122-134.

Fok, Y. K., Scutella, R., & Wilkins, R. (2015). The Low-Pay No-Pay Cycle: Are There Systematic Differences across Demographic Groups? *Oxford Bulletin of Economics and Statistics*, *77*(6), 872-896.

Heckman, J. J. (1981a). Heterogeneity and state dependence. In *Studies in Labor Markets* (pp. 91-140). University of Chicago Press.

Heckman, J. J. (1981b). The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process, in C. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge.

International Monetary Fund (IMF). 2019. *IMF's Work on Income Inequality*. Retrieved from: https://www.imf.org/external/np/fad/inequality/

Mosthaf, A. (2014). Do Scarring Effects of Low-Wage Employment and Non-Employment Differ BETWEEN Levels of Qualification? *Scottish Journal of Political Economy*, 61(2), 154-177.

Plum, A. (2019). The British low-wage sector and the employment prospects of the unemployed. *Applied Economics*, *51*(13), 1411-1432.

Pacheco, G., & Plum, A. (2019) When there is no way up: Reconsidering low-paid jobs as stepping stones. *mimeo*.

Skrondal, A., & Rabe-Hesketh, S. (2014). Handling initial conditions and endogenous covariates in dynamic/transition models for binary data with unobserved heterogeneity. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 63(2), 211-237.

Stewart, M. B. (2007). The interrelated dynamics of unemployment and low-wage employment. *Journal of Applied Econometrics*, 22(3), 511-531.

Stewart, M. B., & Swaffield, J. K. (1999). Low pay dynamics and transition probabilities. *Economica*, 66(261), 23-42.

Uhlendorff, A. (2006). *From no pay to low pay and back again? A multistate model of low pay dynamics*. IZA Discussion Papers, No. 2482, http://nbn-resolving.de/urn:nbn:de:101:1-2008051370

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1), 39-54.

Appendix

Table A 1: Regression results – Base model

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Dependent variable:	**Intermediate pay$_t$**				**Higher pay$_t$**			
Labour market position$_{t-1}$								
Low pay	*reference category*							
Intermediate pay	0.877	0.050	17.46	0.000	0.607	0.055	11.02	0.000
Higher pay	0.629	0.053	11.87	0.000	1.804	0.055	32.64	0.000
Labour market position$_{t=0}$								
Low pay	*reference category*							
Intermediate pay	0.320	0.067	4.81	0.000	0.444	0.097	4.56	0.000
Higher pay	0.619	0.065	9.57	0.000	2.123	0.089	23.87	0.000
Qualification								
No qualification	*reference category*							
Level 1-4	0.048	0.056	0.85	0.394	0.712	0.076	9.35	0.000
Level 5-6	0.052	0.065	0.80	0.421	1.035	0.086	12.04	0.000
Months employed$_{t-1}$								
12 months	*reference category*							
10-11 months	-0.710	0.057	-12.54	0.000	-1.271	0.060	-21.18	0.000
6-9 months	-1.020	0.063	-16.28	0.000	-1.661	0.068	-24.35	0.000
<6 months	-1.336	0.074	-18.01	0.000	-2.133	0.084	-25.34	0.000
Number of children								
No children	*reference category*							
1 child	0.114	0.056	2.04	0.041	0.204	0.063	3.26	0.001
2 children	0.048	0.055	0.88	0.382	0.178	0.063	2.81	0.005
3 children	-0.003	0.071	-0.04	0.971	0.087	0.082	1.07	0.286
4+ children	0.007	0.092	0.08	0.938	-0.015	0.112	-0.13	0.896

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Benefit recipient$_{t-1}$								
No benefits				*reference category*				
1-6 months	-0.111	0.069	-1.61	0.108	-0.433	0.080	-5.38	0.000
≥7 months	-0.850	0.089	-9.53	0.000	-1.406	0.125	-11.22	0.000
ACC$_{t-1}$								
No ACC				*reference category*				
1-3 months	-0.311	0.080	-3.90	0.000	-0.468	0.083	-5.68	0.000
>3 months	-0.182	0.128	-1.42	0.155	-0.179	0.135	-1.33	0.185
Ethnicity								
NZ European				*reference category*				
Māori	-0.047	0.066	-0.71	0.480	-0.433	0.091	-4.77	0.000
Pacific peoples	-0.170	0.115	-1.48	0.139	-0.419	0.155	-2.70	0.007
_cons	0.235	0.076	3.11	0.002	-0.280	0.103	-2.72	0.007
$\sigma^2_{\kappa_2}$	0.871	0.064	13.65	0.000				
$\sigma^2_{\kappa_3}$	3.194	0.146	21.93	0.000				
ρ_κ	0.628	0.027	23.15	0.000				
Log likelihood					-37429.25			

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251.

Table A 2: Regression results – Intensity model

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Dependent variable:		**Intermediate pay$_t$**				**Higher pay$_t$**		
Labour market position$_{t-1}$								
Intermediate pay intensity	3.248	0.115	28.27	0.000	4.214	0.141	29.80	0.000
Higher pay intensity	1.197	0.085	14.07	0.000	6.165	0.102	60.44	0.000
Intermediate pay intensity x Higher pay intensity	-2.042	0.328	-6.22	0.000	-2.663	0.338	-7.88	0.000
Labour market position$_{t=0}$								
Intermediate pay intensity	0.872	0.149	5.86	0.000	1.584	0.197	8.06	0.000
Higher pay intensity	0.528	0.093	5.65	0.000	1.464	0.119	12.29	0.000
Intermediate pay intensity x Higher pay intensity	-1.209	0.374	-3.23	0.001	-3.218	0.453	-7.11	0.000
Qualification								
No qualification				*reference category*				
Level 1-4	0.051	0.056	0.91	0.363	0.371	0.070	5.29	0.000
Level 5-6	0.063	0.066	0.96	0.337	0.371	0.080	4.66	0.000
Months employed$_{t-1}$								
12 months				*reference category*				
10-11 months	-0.607	0.057	-10.58	0.000	-0.920	0.063	-14.71	0.000
6-9 months	-0.903	0.063	-14.22	0.000	-1.238	0.072	-17.23	0.000
<6 months	-1.071	0.077	-13.85	0.000	-1.260	0.094	-13.47	0.000
Number of children								
No children				*reference category*				
1 child	0.082	0.057	1.44	0.149	-0.008	0.062	-0.13	0.900
2 children	0.022	0.057	0.39	0.693	-0.022	0.063	-0.36	0.720
3 children	0.017	0.073	0.24	0.810	0.003	0.081	0.04	0.966
4+ children	-0.027	0.093	-0.29	0.774	-0.121	0.110	-1.10	0.271

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Benefit recipient$_{t-1}$								
No benefits				*reference category*				
1-6 months	-0.047	0.070	-0.67	0.502	0.002	0.085	0.02	0.982
≥7 months	-0.563	0.092	-6.14	0.000	-0.323	0.134	-2.41	0.016
ACC$_{t-1}$								
No ACC				*reference category*				
1-3 months	-0.326	0.081	-4.05	0.000	-0.391	0.084	-4.66	0.000
>3 months	-0.151	0.130	-1.16	0.246	0.340	0.138	2.46	0.014
Ethnicity								
NZ European				*reference category*				
Māori	-0.078	0.066	-1.18	0.239	-0.210	0.082	-2.55	0.011
Pacific peoples	-0.197	0.113	-1.75	0.080	-0.250	0.137	-1.82	0.068
_cons	-0.514	0.081	-6.37	0.000	-2.938	0.116	-25.38	0.000
$\sigma^2_{\varepsilon_2}$	0.877	0.056	15.62	0.000				
$\sigma^2_{\varepsilon_3}$	2.312	0.097	23.78	0.000				
ρ_ε	0.757	0.017	43.27	0.000				
Log likelihood				-33724.21				

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251.

Table A 3: Regression results – Base model

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Dependent variable:	**Employed for 10-11 months$_t$**				**Employed for less than 10 months$_t$**			
Labour market position$_{t-1}$								
Low pay				*reference category*				
Intermediate pay	-0.116	0.059	-1.97	0.049	0.059	0.055	1.07	0.285
Higher pay	-0.285	0.056	-5.08	0.000	-0.161	0.054	-2.98	0.003
Labour market position$_{t=0}$								
Low pay				*reference category*				
Intermediate pay	-0.034	0.074	-0.45	0.652	-0.071	0.075	-0.94	0.345
Higher pay	-0.180	0.068	-2.66	0.008	-0.370	0.068	-5.40	0.000
Months employed$_{t-1}$								
12 months				*reference category*				
10-11 months	0.857	0.059	14.53	0.000	1.155	0.056	20.56	0.000
<10 months	1.244	0.058	21.31	0.000	2.191	0.052	42.22	0.000
Months employed$_{t=0}$								
12 months				*reference category*				
10-11 months	0.622	0.072	8.70	0.000	0.656	0.075	8.74	0.000
<10 months	0.438	0.060	7.33	0.000	0.694	0.061	11.44	0.000
Qualification								
No qualification				*reference category*				
Level 1-4	-0.374	0.055	-6.85	0.000	-0.261	0.057	-4.55	0.000
Level 5-6	-0.412	0.062	-6.68	0.000	-0.315	0.065	-4.88	0.000
Number of children								
No children				*reference category*				
1 child	-0.081	0.053	-1.53	0.127	0.018	0.054	0.34	0.734
2 children	-0.176	0.052	-3.39	0.001	-0.020	0.053	-0.38	0.705
3 children	-0.108	0.067	-1.62	0.104	0.127	0.067	1.89	0.058

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
4+ children	0.009	0.089	0.10	0.921	0.186	0.090	2.06	0.039
Benefit recipient$_{t-1}$								
No benefits				*reference category*				
1-6 months	0.302	0.073	4.13	0.000	0.190	0.068	2.79	0.005
≥7 months	0.190	0.101	1.87	0.061	0.433	0.088	4.94	0.000
ACC$_{t-1}$								
No ACC				*reference category*				
1-3 months	0.425	0.077	5.50	0.000	0.437	0.077	5.71	0.000
>3 months	-0.187	0.131	-1.42	0.155	-0.495	0.122	-4.05	0.000
Ethnicity								
NZ European				*reference category*				
Māori	0.416	0.064	6.53	0.000	0.370	0.067	5.53	0.000
Pacific peoples	-0.118	0.119	-1.00	0.319	-0.027	0.120	-0.23	0.819
_cons	-2.635	0.087	-30.15	0.000	-3.034	0.090	-33.62	0.000
$\sigma^2_{\kappa_2}$	0.911	0.065	14.05	0.000				
$\sigma^2_{\kappa_3}$	1.124	0.074	15.19	0.000				
ρ_κ	0.857	0.028	30.5	0.000				
Log likelihood					-28668.63			

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251.

Table A 4: Regression results – Intensity model

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
Dependent variable:	**Employed for 10-11 months$_t$**				**Employed for less than 10 months$_t$**			
Labour market position$_{t-1}$								
Intermediate pay intensity	-0.993	0.129	-7.72	0.000	-1.041	0.120	-8.70	0.000
Higher pay intensity	-1.105	0.084	-13.23	0.000	-1.170	0.079	-14.80	0.000
Intermediate pay intensity x Higher pay intensity	1.026	0.315	3.25	0.001	1.100	0.309	3.56	0.000
Labour market position$_{t=0}$								
Intermediate pay intensity	-0.586	0.163	-3.60	0.000	-0.518	0.158	-3.27	0.001
Higher pay intensity	-0.289	0.098	-2.94	0.003	-0.521	0.096	-5.42	0.000
Intermediate pay intensity x Higher pay intensity	1.268	0.368	3.45	0.001	0.506	0.372	1.36	0.174
Months employed$_{t-1}$								
12 months				*reference category*				
10-11 months	0.811	0.059	13.69	0.000	1.094	0.056	19.42	0.000
<10 months	1.173	0.059	19.85	0.000	2.097	0.053	39.76	0.000
Months employed$_{t=0}$								
12 months				*reference category*				
10-11 months	0.576	0.070	8.24	0.000	0.593	0.072	8.22	0.000
<10 months	0.358	0.061	5.9	0.000	0.558	0.060	9.28	0.000
Qualification								
No qualification				*reference category*				
Level 1-4	-0.307	0.054	-5.72	0.000	-0.182	0.055	-3.30	0.001
Level 5-6	-0.298	0.061	-4.87	0.000	-0.172	0.063	-2.75	0.006
Number of children								
No children				*reference category*				
1 child	-0.040	0.053	-0.76	0.446	0.080	0.053	1.51	0.132
2 children	-0.122	0.052	-2.35	0.019	0.057	0.052	1.08	0.278
3 children	-0.067	0.066	-1.01	0.312	0.186	0.066	2.82	0.005

	Coef.	Std. Err.	z	P>|z|	Coef.	Std. Err.	z	P>|z|
4+ children	0.043	0.088	0.50	0.620	0.235	0.088	2.68	0.007
Benefit recipient$_{t-1}$								
No benefits				*reference category*				
1-6 months	0.195	0.073	2.67	0.008	0.062	0.068	0.91	0.363
≥7 months	-0.122	0.102	-1.20	0.232	0.051	0.087	0.58	0.560
ACC$_{t-1}$								
No ACC				*reference category*				
1-3 months	0.404	0.077	5.23	0.000	0.421	0.076	5.54	0.000
>3 months	-0.278	0.131	-2.12	0.034	-0.577	0.121	-4.78	0.000
Ethnicity								
NZ European				*reference category*				
Māori	0.359	0.063	5.74	0.000	0.297	0.064	4.61	0.000
Pacific peoples	-0.166	0.117	-1.42	0.154	-0.088	0.116	-0.75	0.451
_cons	-1.942	0.096	-20.25	0.000	-2.110	0.095	-22.24	0.000
$\sigma^2_{\varepsilon_2}$	0.810	0.061	13.32	0.000				
$\sigma^2_{\varepsilon_3}$	0.934	0.067	13.99	0.000				
ρ_ε	0.843	0.032	26.03	0.000				
Log likelihood				-28461.16				

Notes: Data sourced from IDI (2019). Authors' calculations. Time period = 2005 to 2015. *N*= 67,251.

Disclaimer

The results in this paper are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI), managed by Statistics New Zealand. The opinions, findings, recommendations, and conclusions expressed in this paper are those of the authors, not Statistics NZ.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes. Any person who has had access to the unit record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this paper have been confidentialised to protect these groups from identification. Careful consideration has been given to the privacy, security, and confidentiality issues associated with using administrative and survey data in the IDI.

Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.