Persistence of Low Pay Employment

Gail Pacheco, Alexander Plum

PLEASE DO NOT CITE WITHOUT AUTHORS' PERMISSION – THESE ARE PRELIMINARY FINDINGS

NEW ZEALAND
WORK RESEARCH INSTITUTE

SNZ Disclaimer

- Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975.
- The results presented in this study are the work of the authors, not of Statistics NZ.

Motivation

<u>Background:</u>

- Intensive discussion on inequality (e.g. OECD 2015, IMF 2017)
- Numerous studies on the effect of low pay employment on labour market prospects:
  - *stepping-stone* towards higher-paid jobs (e.g. Uhlendorff 2006)
  - *no-pay – low-pay cycle* (e.g. Stewart 2007)
- Studies provide evidence for state dependence in low pay:

$$P(\text{Low pay}_t | \text{Low pay}_{t-1}) \geq P(\text{Low pay}_t | \text{Higher pay}_{t-1})$$

Motivation

Aim of this study:

- Discussing the prevailing identification strategy which is based on annual labour market information
- Comparing the results with a model that uses a large administrative dataset with monthly earning information and accounts for the intensity of the low pay attachment
- Please note that results are **preliminary**

Motivation

Findings (preliminary):

1) Annual share of individuals affected by low pay is underestimated
2) Level of low pay attachment varies across individuals
3) Intensity of low pay attachment over time is highly correlated

$$\Downarrow$$

conventional identification strategy *under*- and *over*estimates the persistence in low pay substantially

Literature Review

United Kingdom (BHPS, Understanding Society):

- Stewart & Swaffield (1999): 'considerable persistence in low pay' [p. 40]
- Cai et al. (2017): 'those employees who are on low pay are more likely to be found on low pay in the future, compared with those who are (…) unemployed or on higher pay' [p. 27]

Italy (Survey on Households Income and Wealth):

- Cappellari (2007): 'considerable state dependence: the experience of low pay raises the probability of subsequent low pay episodes' [p. 465].

Germany (GSOEP):

- Uhlendorff (2006): 'strong true state dependence in low pay' [p. 18]

Europe (ECHP):

- Clark & Kanellopoulos (2013): 'positive, statistically significant state dependence in every single country' [p. 122]

Australia (HILDA):

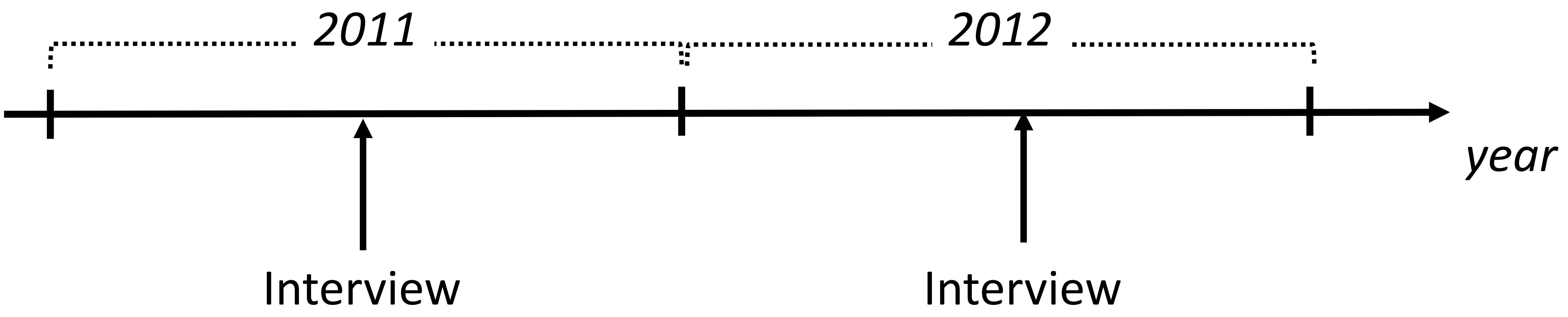
- Fok et al (2015): 'Consistent with the previous literature, the results clearly indicate that there is state dependence in (…) low-paid employment' [p. 885]

Literature Review

Table 1: Low pay persistence of related studies

Study	$\mathrm{P}(\mathrm{Lp}_t \mid \mathrm{Hp}_{t-1})$	$\mathrm{P}(\mathrm{Lp}_t \mid \mathrm{Lp}_{t-1})$
Uhlendorff (2006, Germany)	0.024 – 0.038	0.049 – 0.077
Mosthaf (2014, Germany)	0.033 – 0.007	0.091 – 0.168
Clark & Kanellopoulos (2013)		UK: 0.071
		Germany: 0.064
		Italy: 0.045
Cai et al. (2017, UK)	0.160	0.272
Fok et al. (2015, Australia)		0.123

Conceptual Framework



- Conventional approach: Identification of low pay employed with respect to the **time point of the interview**
- However: wages **not necessarily constant** over the year (job changes, promotion)
- However: Inland Revenue (IR) provides information of wages and salaries on the **monthly level**
- Possibility to derive **level of attachment** to the low pay sector

Conceptual Framework

Looking at earning dynamics (e.g. Baker & Solon 2003, Cappellarie & Jenkins 2014):

$$Y_{ik_m} = \mu_k + y_{ik_m}$$

With individual $i = 1, \dots, N$, year $k = 1, \dots, K$ and month $m = 1, \dots, M$

$$y_{ik_m} = \alpha_i + \nu_{ik_m}$$

with $\alpha_i \sim N(0, \sigma_\alpha^2)$ and $\nu_{ik_m} \sim N(0, \sigma_\nu^2)$

$$LP_{ik_m} = \mathbf{1}\left(Y_{ik_m} \leq \tau\right)$$

Looking at the annual level

$$LP_{ik} = \mathbf{1}\left(\sum_{m=1}^{M} LP_{ik_m} > 0\right)$$

Conceptual Framework

Looking at the annual level:

$$LP_{ik} = \mathbf{1}\left(\sum_{m=1}^{M} LP_{ik_m} > 0\right)$$

Share of individuals experiencing low pay employment:

$$LP_k = \frac{\sum_{i=1}^{N} LP_{ik}}{N \times M} \geq \frac{LP_{ik_m}}{N} \qquad \text{if } \sigma_\nu^2 > 0$$

The share of month an individual was low paid employed (low pay attachment):

$$LP_{ik}^{\mathrm{S}} = \frac{\sum_{m=1}^{M} LP_{ik_m}}{M} \geq LP_{ik_m} \qquad \text{if } \sigma_\nu^2 > 0$$

Conceptual Framework

Correlation of low pay attachment over time:

$$corr\left(LP_{ik}^{S}, LP_{ik+1}^{S}\right) \geq corr\left(LP_{ik_m}, LP_{ik+1_m}\right) \qquad \text{if } \sigma_{\nu}^{2} > 0$$

$$Y_{ik_m} = \mu_k + \alpha_i + \nu_{ik_m}$$

$$Y_{ik} = M \times \mu_k + M \times \alpha_i + \sum_{m=1}^{M} \nu_{ik_m}$$

$$\bar{Y}_{ik} = \mu_k + \alpha_i + \frac{\sum_{m=1}^{M} \nu_{ik_m}}{M}$$

$$\xrightarrow[M \to \infty]{} 0$$

Simulation

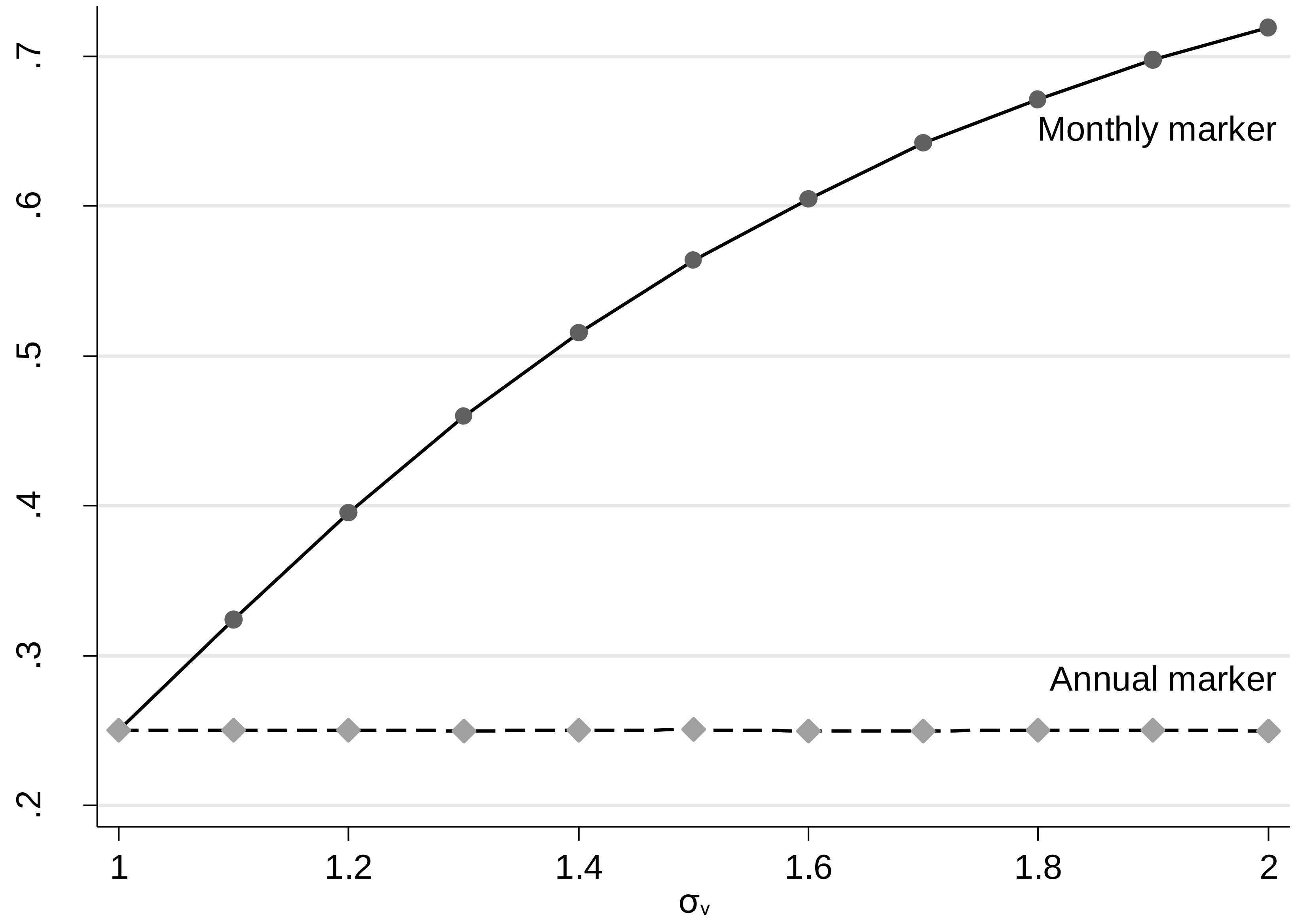
Design:

- 5000 individuals, each employed for 12 month
- $\mu = \log(2000)$

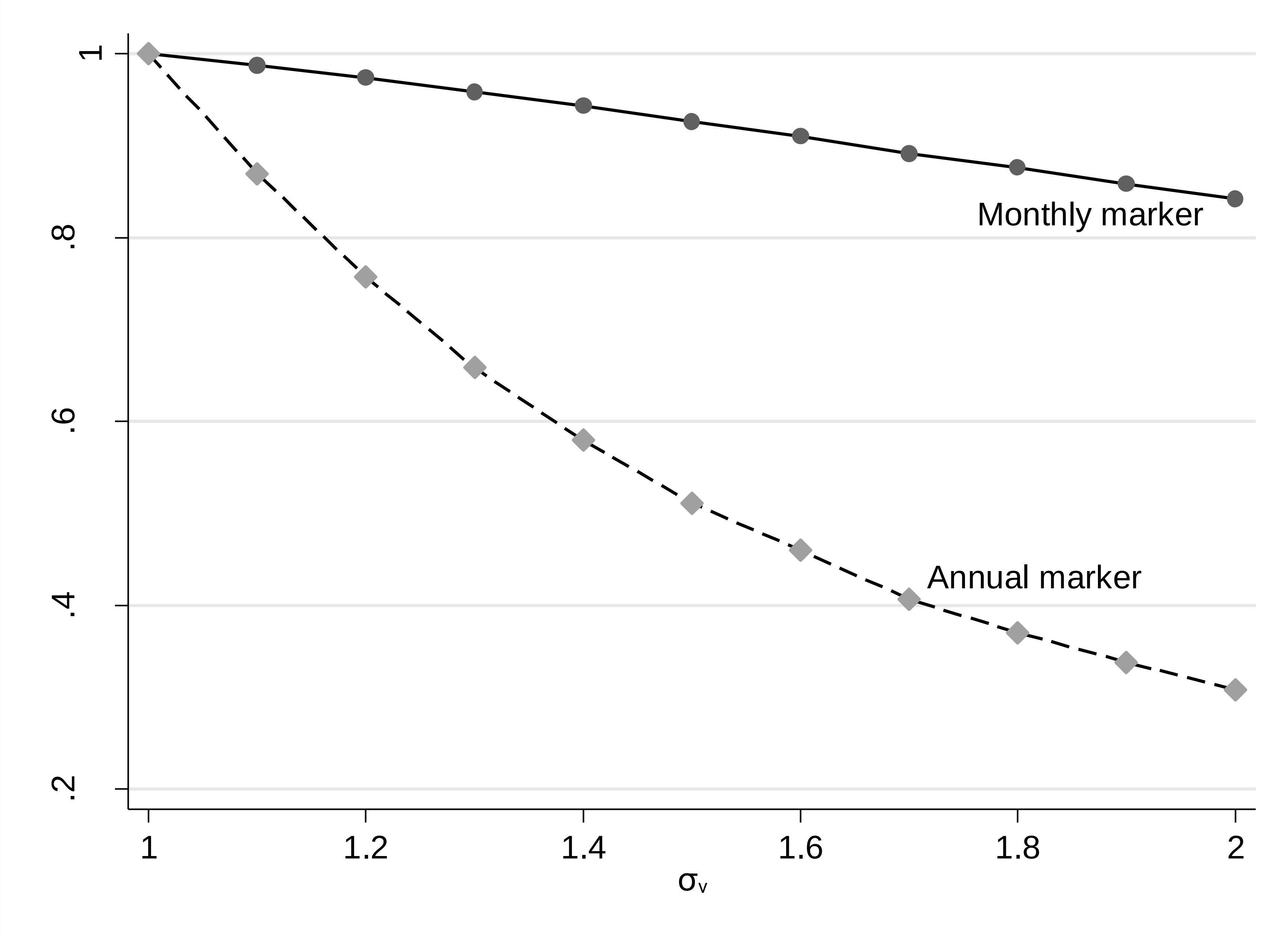
- $\sigma_\alpha = \log(2)$
- $\sigma_v = \{\log(1), \log(1.1) \dots \log(2)\}$

- $\tau = 25th$ percentile
- 100 replications

Simulation



Simulation



Descriptive Statistics

<u>Statistics New Zealand's Integrated Data Infrastructure (IDI):</u>

- IDI links longitudinal microdata about individuals, households etc. from various sources
- Backbone is the Central Linking Concordance (CLC) which contains a list of all individuals with some characteristics (e.g. sex, date of birth)

<u>Inland Revenue tax data (IR):</u>

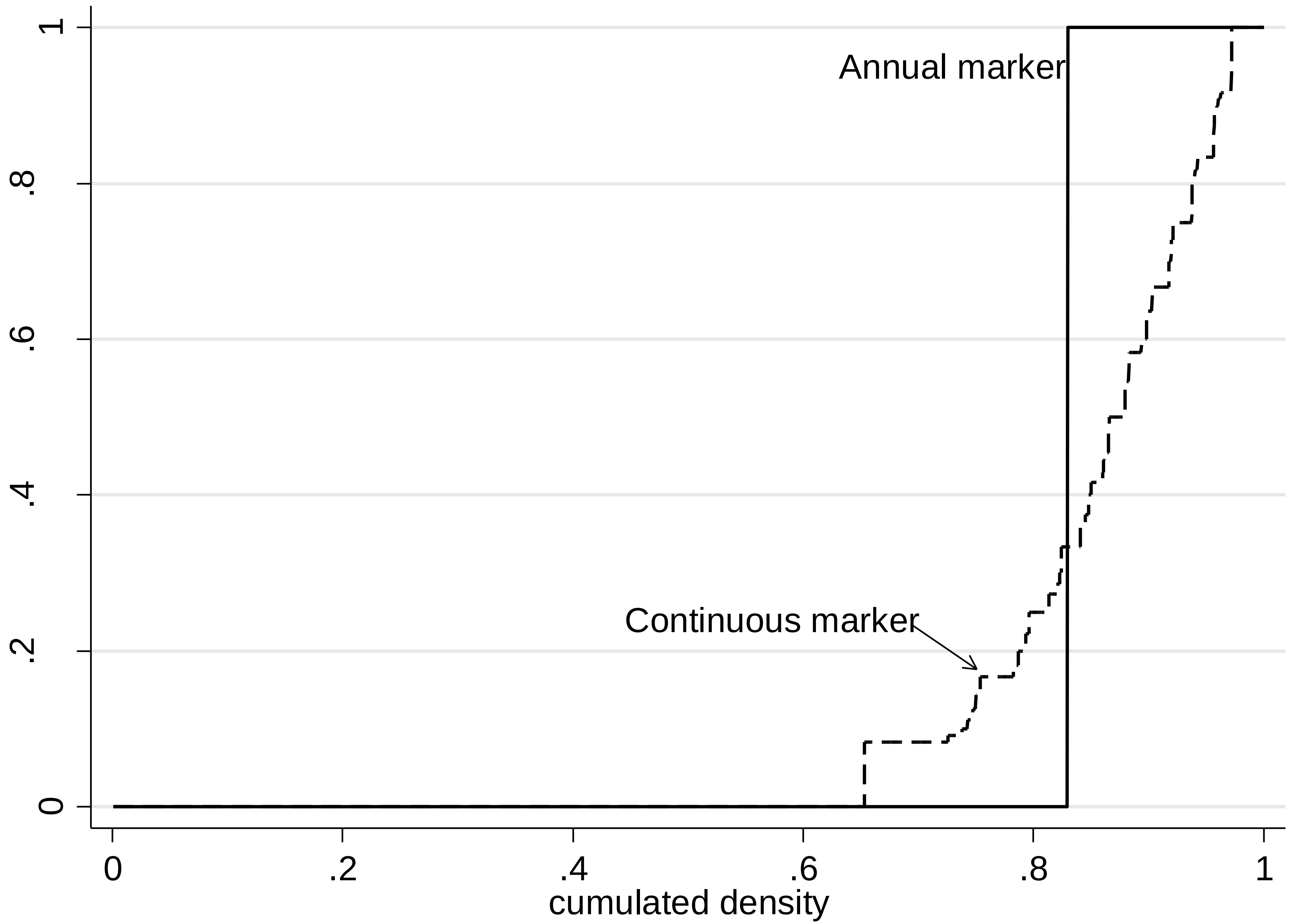
- Information on person tax data from Inland Revenue
- Data are provided from 1 April 1999 onwards and the geographic coverage refers to all New Zealand
- Data are collected and supplied *monthly* to the IDI
- For our analysis we use the gross earnings before tax that come from wages and salaries

Descriptive Statistics

<u>Data restrictions:</u>

- Restrict to men 25-55 (OECD: 95% FT employed)
- Drop those with wages below $30h \times 4.2\text{weeks} \times MW_{\text{year}}$
- Time frame 2000-2016
- Employed at least 6 months per year (5 consecutive years)
- Age group adjusted monthly low pay threshold (OECD, percentile)
- Using a random subsample of $N = 39{,}552$ observations

Descriptive Statistics



Source: IDI (2018) and own calculations. N= 39,552

Descriptive Statistics

Table 3: Comparing marker

		Group $marker_t$				
		No Lp_t	Lp $<25\%_t$	Lp $25\text{-}50\%_t$	Lp $>50\%_t$	
Annual $marker_t$	No Lp_t					81.82%
	Lp_t					18.18%
		64.81%	16.17%	6.84%	12.18%	

Source: IDI (2018) and own calculations. N= 39,552

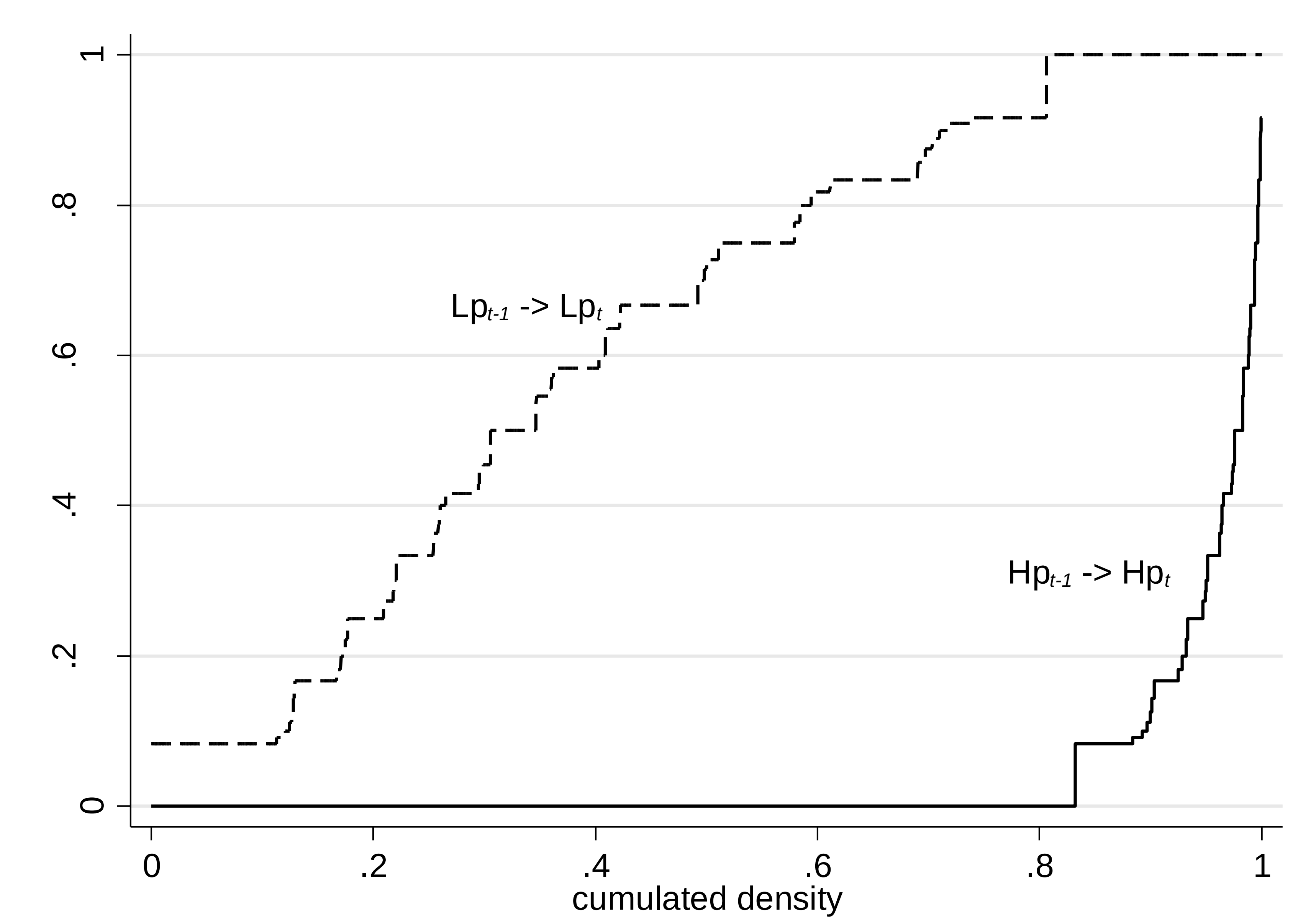
Descriptive Statistics

Table 3: Comparing marker

		Group marker$_t$				
		No Lp$_t$	Lp <25%$_t$	Lp 25-50%$_t$	Lp >50%$_t$	
Annual marker$_t$	No Lp$_t$	79.22%	12.89%	4.79%	3.11%	81.82%
	Lp$_t$	-	30.94%	16.06%	53.00%	18.18%
		64.81%	16.17%	6.84%	12.18%	

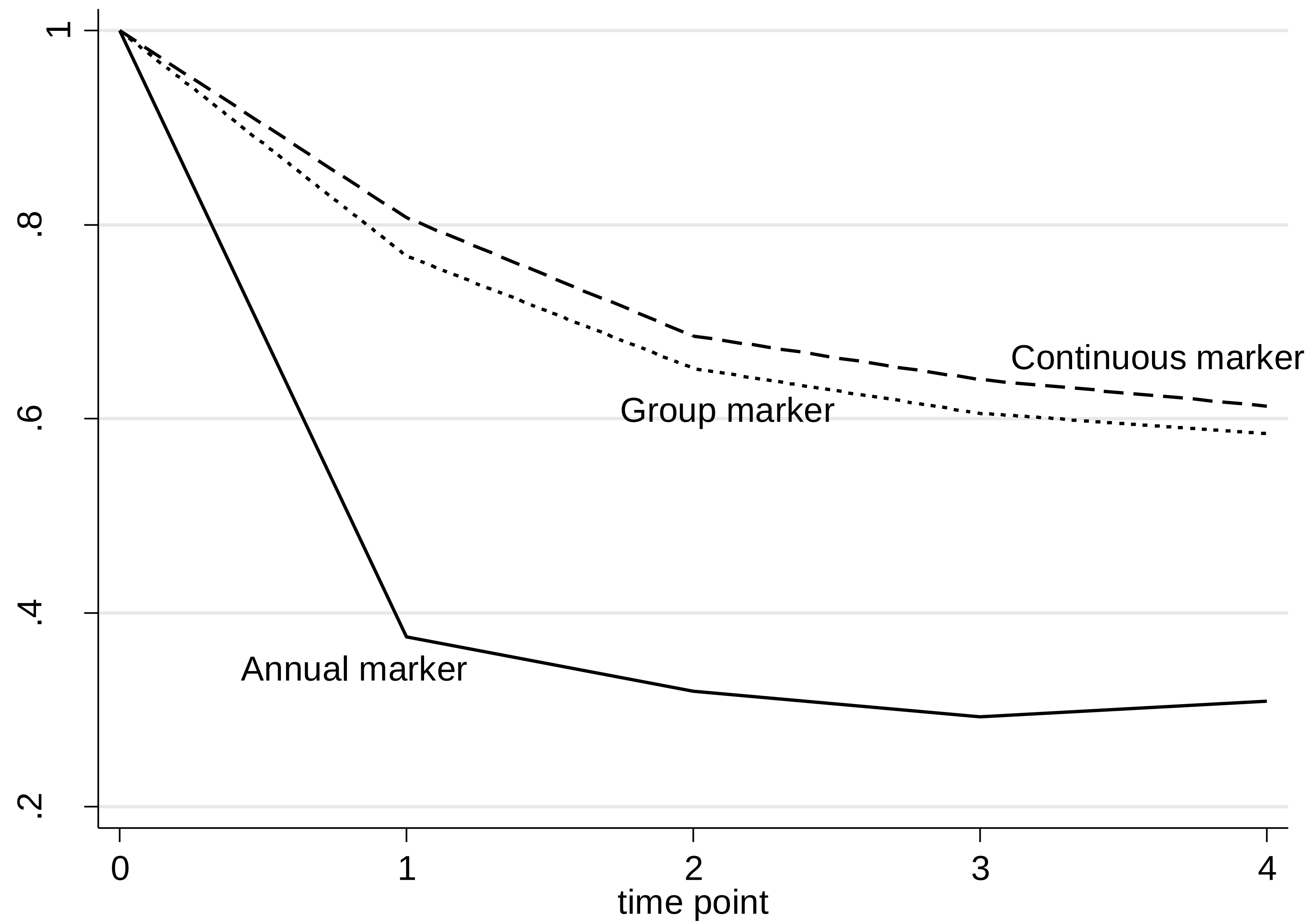
Source: IDI (2018) and own calculations. N= 39,552

Descriptive Statistics



Source: IDI (2018) and own calculations. N= 39,552

Descriptive Statistics



Source: IDI (2018) and own calculations. N= 39,552

Descriptive Statistics

Table 4: Transition matrix

	No Lp_t	Lp $<25\%_t$	Lp $25\text{-}50\%_t$	Lp $>50\%_t$	
No Lp_{t-1}	88.75%	9.48%	1.18%	0.6%	64.94%
Lp $<25\%_{t-1}$	43.69%	39.36%	12.08%	4.88%	16.3%
Lp $25\text{-}50\%_{t-1}$	13.61%	31.74%	31.42%	23.23%	6.84%
$Lp>50\%_{t-1}$	2.69%	8.4%	14.03%	74.88%	11.92%
	66.01%	15.74%	6.56%	11.7%	

Source: IDI (2018) and own calculations. N= 39,552

Econometric Model

<u>Basic concept:</u>

- First-order Markov process: lagged dependent variable has a genuine effect
- Controlling for unobserved heterogeneity (Heckman 1981a) and its correlation with the initial conditions (Heckman 1981b)
- Applying a multivariate random effects probit model which was also used in various other low pay studies (Stewart 2007, Buddelmeyer et al. 2010, Knabe & Plum 2013, Clark & Kanellopoulos 2013)

Econometric Model

The following binary outcome variables are defined as:

$$y_{it}^{(hp)} = \begin{cases} 1 & \text{no low−pay spells,} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{it}^{(lp\ 1)} = \begin{cases} 1 & \text{low−pay spells < } \textit{25} \text{ percent,} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{it}^{(lp\ 2)} = \begin{cases} 1 & \text{low−pay spells } \textit{25 − 50} \text{ percent,} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{it}^{(lp\ 3)} = \begin{cases} 1 & \text{low−pay spells > } \textit{50} \text{ percent,} \\ 0 & \text{otherwise} \end{cases}$$

Econometric Model

The observed binary outcome variables are:

$$y_{it}^{(hp)} = \mathbf{1}\left(\gamma_{11} y_{it-1}^{(lp\ 1)} + \gamma_{12} y_{it-1}^{(lp\ 2)} + \gamma_{13} y_{it-1}^{(lp\ 3)} + x_{1it}' \beta_1 + \alpha_{1i} + u_{1it} > 0\right)$$

and if $y_{it}^{(hp)} = 0$,

$$y_{it}^{(lp\ 1)} = \mathbf{1}\left(\gamma_{21} y_{it-1}^{(lp\ 1)} + \gamma_{22} y_{it-1}^{(lp\ 2)} + \gamma_{23} y_{it-1}^{(lp\ 3)} + x_{2it}' \beta_2 + \alpha_{2i} + u_{2it} > 0\right)$$

and if $y_{it}^{(hp)} = 0$ and $y_{it}^{(lp\ 1)} = 0$,

$$y_{it}^{(lp\ 2)} = \mathbf{1}\left(\gamma_{31} y_{it-1}^{(lp\ 1)} + \gamma_{32} y_{it-1}^{(lp\ 2)} + \gamma_{33} y_{it-1}^{(lp\ 3)} + x_{3it}' \beta_3 + \alpha_{3i} + u_{3it} > 0\right)$$

Econometric Model

To take care of the “initial conditions problem”, we follow the suggestion of Wooldridge (2005) by applying a conditional random-intercept model:

$$\alpha_{ji} = \pi_{j1} y_{i0}^{(lp\ 1)} + \gamma_{j2} y_{i0}^{(lp\ 2)} + \gamma_{j3} y_{i0}^{(lp\ 3)} + \bar{x}'_{jit}\delta_j + \kappa_{ji}$$

with $j \in \{1,2,3\}, u_{jit} \sim N(0,1), \kappa_{ji} \sim N\left(0, \sigma_{\kappa_j}^2\right)$.

$$MSL = \prod_{i=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{t=1}^{T_i} P_{it}(\kappa_1^r, \kappa_2^r, \kappa_3^r) \right\}$$

⇒ All written in *Mata*

Results

Table 5: Regression results

	No Lp_t	Lp $<25\%_t$	Lp $25\text{-}50\%_t$
No Lp_{t-1}	*reference category*		
Lp $<25\%_{t-1}$	-0.656	-0.473	0.143
	(0.027)	(0.042)	(0.083)
Lp $25\text{-}50\%_{t-1}$	-1.459	-1.137	-0.308
	(0.043)	(0.047)	(0.081)
$Lp{>}50\%_{t-1}$	-2.367	-2.098	-1.240
	(0.056)	(0.050)	(0.078)
Initial labour market position	✓	✓	✓
Exogenous regressors	✓	✓	✓
Random effects (uncorrelated)	✓	✓	✓

Source: IDI (2018) and own calculations. N= 35,874

Results

Table 6: Predicted probabilities

	No Lp_t	Lp <25%$_t$	Lp 25-50%$_t$	Lp >50%$_t$
No Lp_{t-1}	0.777	0.178	0.028	0.017
	(0.134)	(0.097)	(0.023)	(0.017)
Lp <25%$_{t-1}$	0.616	0.256	0.087	0.041
	(0.166)	(0.089)	(0.052)	(0.033)
Lp 25-50%$_{t-1}$	0.390	0.269	0.182	0.160
	(0.156)	(0.046)	(0.061)	(0.077)
Lp>50%$_{t-1}$	0.174	0.127	0.156	0.543
	(0.094)	(0.029)	(0.033)	(0.104)

Source: IDI (2018) and own calculations. N= 35,874

Results

Table 7: Predicted probabilities (annual data)

	No Lp_t	Lp_t
No Lp_{t-1}	0.865	0.135
	(0.079)	(0.079)
Lp_{t-1}	0.732	0.268
	(0.113)	(0.113)

Source: IDI (2018) and own calculations. N= 35,874

Conclusion

<u>Findings (preliminary):</u>

1) Annual share of individuals affected by low pay is underestimated
2) Level of low pay attachment varies across individuals
3) Intensity of low pay attachment over time is highly correlated

$$\Downarrow$$

conventional identification strategy *under*- and *over*estimates the persistence in low pay substantially

Thank you very much for your time

Questions?