Nowcasting current or yearly income?

Patryk Bronka, Diego Collado, **Matteo Richiardi**
CeMPA

InGRID-2 expert meeting on Nowcasting and mid-term projections through microsimulation models, 16 July 2021

CeMPA Centre for Microsimulation and Policy Analysis
iSER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH
University of Essex

Nowcasting: Main issues

1. **How to update variables** (in particular: employment state)

2. **How to use the updated variables** (in particular: yearly vs. monthly model)

Income concept for T-B models

- The EU-27 EUROMOD model is based on SILC data.
- UKMOD on the other hand, and the UK component of EUROMOD before that, are based on FRS.
- The two sources of data have a different income concept.
- While respondents are interviewed throughout the year in both FRS and EU-SILC data,
  - SILC respondents are asked about *income in previous calendar year* (that is, January to December before the interview), irrespective of the source, while
  - FRS respondents are asked about *current income*. With regard to income from benefits, FRS respondents provide figures which relate most commonly to the last week, the last two weeks or the last month. With regard to earnings, employees are asked about the latest salary payment, and the self-employed about annual profits from the most recent tax returns. For income from investment and employee non-cash income, respondents report their most recent annual or half-yearly income that they received from this source. All income is then annualised.

Implications: Objectives for Nowcasting

- EU-27: **Predict yearly income**

- UK: **Predict current income**

The “representative month” assumption

- “instant” simulations misrepresent policies that condition on yearly variables (e.g. most taxes),
- yearly simulations misrepresent policies that condition on current state
    - This is particularly problematic for Covid-19 shocks and policies.

- Ideal world: combine both approaches and reconstruct yearly incomes from monthly incomes.
- Experience with nowcasting (combining standard input data with higher frequency information) might help moving forward towards this goal.

Nowcasting in the public version of UKMOD

Employment shock	Employees	Self-employed
- extensive margin	Workers randomly put into unemployment based on aggregate ONS unemployment data	
- intensive margin	(Modelled jointly with income support)	Workers are randomly considered either unaffected (same number of hours worked) or affected (reduced number of hours worked), based on Understanding society Covid-19 study data
Income support	Workers randomly put to either full time or part-time furlough based on monthly HMRC aggregate data, partially disaggregated by sectors.	Workers randomly given a SEISS grant based on monthly HMRC aggregate data

Notes: ONS – Office for National Statistics, UK statistical agency; HMRC – HM Revenue & Customs, UK's tax, payments and customs authority.

- Employment state is updated based on the month of the interview

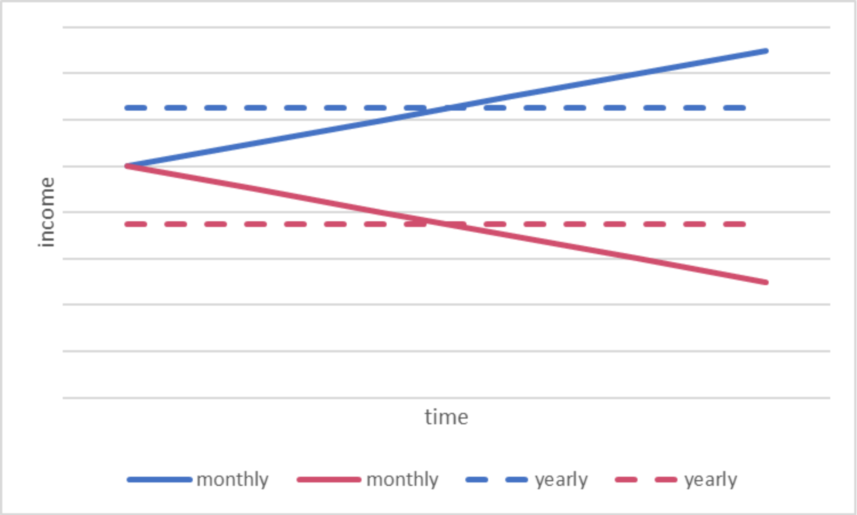
← mimics what the next FRS would observe

A bigger question

- Are simulations based on current income comparable to simulations based on yearly incomes?

Theoretical insights /1

- Taking a snapshot of income rather than measuring the inflow over a longer period of time has important consequences, as the instantaneous measure is by construction **more volatile** than the cumulated measure (which is basically a mean).



- **Current income is more unequally distributed than yearly income,** on average.
- Moreover, if individual income is subject to shocks, current income is even more unequally distributed than yearly income, on average.
- **The difference increases with the variance of the shocks**.

Theoretical insights /2

- Current income:

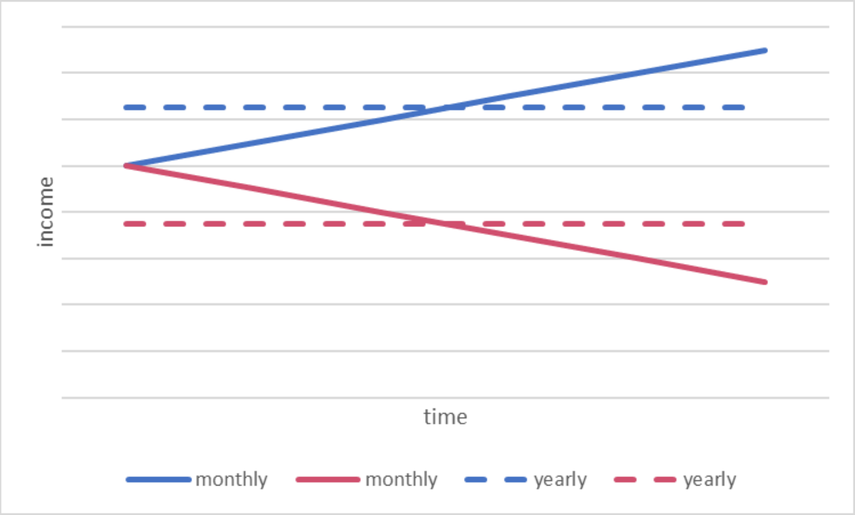
$y_{i,t} = \alpha_i + \beta_i t + u_{i,t}$ u i.i.d. with E(u)=0

$Var(y_t) = Var(\alpha) + Var(\beta)\, t^2 + 2Cov(\alpha,\beta)\, t + V(\varepsilon_t)$

- Average income:

$y_i = \alpha_i + \beta_i T$

$Var(y) = Var(\alpha) + Var(\beta)\, T^2 + 2Cov(\alpha,\beta)\, T$



Using variance as the measure of inequality

Does that matter?

- Using data from the British Household Panel Survey (BHPS), Böheim and Jenkins (2006) show that estimates of cross-sectional and longitudinal income distribution summary statistics, computed on current and annual income respectively, are remarkably similar.
- They explain this finding partly with the fact that some income sources classified as 'current income' refer to usual income rather than the most recent income, and partly with the fact that within-year income volatility appears to be low for most individuals.
- Jenkins (2010) concludes that the distinction between measures of current and annual income is unimportant relative to other issues, at least in BHPS data.

Bönheim, R. and Jenkins, Stephen P. (2006). A comparison of current and annual measures of income in the British Household Panel Survey. Journal of Official Statistics 22 (4), 733-758.

Jenkins, Stephen P. (2010). The British Household Panel Survey and its income data. ISER WP No. 2010-33.

Does that matter now?

- The old results of Jenkins et al. might not hold any more though:

1. Earnings volatility might have increased in recent years (gig economy, etc.) ← but a large literature shows roughly constant if not slightly declining trends, see Moffitt et al. (2020) and Ziliak et al. (2020) for the US.

2. **Covid-19** surely has produced a spike in volatility in 2020.

Data: Understanding Society Covid-19 study

- 8 waves are now available.
- The surveys started in April 2020 and took place monthly until July 2020. From September 2020 onwards, they have taken place every other month.
- Last wave is March 2021.
- The surveys complement the annual interviews of the Understanding Society study. The data can be linked to data on the same individuals from previous waves of the annual interviews.
- Employment module: contains information on employment state and government support (CJRS and SEISS grants), as well as hours worked and earnings.

Models

- We consider monthly transitions between 4 states:
  - Employment
  - Self-employment
  - Non-employment
  - Furlough
- For the self-employed, we additionally model the probability of receiving a grant from SEISS (Self-Employment Income Support Scheme).

Overview of transitions

We model the transitions between the states below using multinomial logit models.
Transition models control for a number of observable characteristics such as gender, age, education, household composition, industry, working hours, baseline pay, lagged pay, occupation, time.
*New working hours and unit wages predicted, conditional on baseline values and current occupation / industry. We allow for correlation in the error terms using seemingly unrelated regressions.
+New industry and occupation predicted, conditional on age, gender, education

From / To	Employment	Self-employment	Furlough	Non-employment
Employment	No changes / some changes* ✓	✓	✓	✓
Self-employment	* ✓	* / SEISS ✓		✓
Furlough	* ✓	* ✓	✓	✓
Non-employment	*+ ✓	*+ ✓		✓

Transitions from employment

- We estimate a multinomial logit model with 5 outcomes:
  - No changes to worker's situation
  - Some changes (monthly pay differs from the baseline by more than +/- 5%)
  - Furlough
  - Self-employment
  - Non-employment
- For the sample of those for whom the outcome is self-employment or some changes to the employment situation, we additionally model:
  - New working hours
  - New hourly wage

From / To	Employment	Self-employment	Furlough	Non-employment
Employment	No changes / some changes* ✓	✓	✓	✓
Self-employment	* ✓	* / SEISS ✓		✓
Furlough	* ✓	* ✓	✓	✓
Non-employment	*+ ✓	*+ ✓		✓

Multinomial logit

- Outcome depends on:
  - Age
  - Gender
  - Interaction of age and gender
  - Education (Low / Medium / High)
  - Household composition (Single adult no children; Single adult living with children; Multiple adults no children; Multiple adults living with children)
  - Industry (13 categories, SIC07 classification with some extra aggregation)
  - Lagged hours of work + interaction with gender
  - Quintile of monthly pay in the baseline (Jan – Feb wave)
  - Lagged monthly pay
  - Occupation (9 categories, SOC 2010 aggregated up to 1 –digit)
  - Lagged furlough status + interaction with month
  - Month

Hours and wages

- Modelled jointly using the seemingly unrelated regressions, to allow the errors associated with the dependent variables to be correlated.
- Wages depend on gender, age, interaction of gender and age, household composition, baseline hours of work, baseline hourly pay, industry, occupation, education, month
- Hours use the same set of independent variables + hourly pay (contemporaneous)

Transitions from furlough

- We estimate a multinomial logit model with 4 outcomes:
  - Employment
  - Furlough
  - Self-employment
  - Non-employment
- For the sample of those for whom the outcome is self-employment or employment, we additionally model:
  - New working hours
  - New hourly wage

From / To	Employment	Self-employment	Furlough	Non-employment
Employment	No changes / some changes* ✓	✓	✓	✓
Self-employment	* ✓	* / SEISS ✓		✓
Furlough	* ✓	* ✓	✓	✓
Non-employment	*+ ✓	*+ ✓		✓

Multinomial logit

- Outcome depends on:
  - Age
  - Gender
  - Interaction of age and gender
  - Education (Low / Medium / High)
  - Household composition (Single adult no children; Single adult living with children; Multiple adults no children; Multiple adults living with children)
  - Industry (13 categories, SIC07 classification with some extra aggregation)
  - Lagged hours of work + interaction with gender
  - Quintile of monthly pay in the baseline (Jan – Feb wave)
  - Lagged monthly pay
  - Occupation (9 categories, SOC 2010 aggregated up to 1 –digit)
  - Month
- Hours and wages: same set of covariates as for transitions from employment

Transitions from self-employment

From / To	Employment	Self-employment	Furlough	Non-employment
Employment	No changes / some changes* ✓	✓	✓	✓
Self-employment	* ✓	* / SEISS ✓		✓
Furlough	* ✓	* ✓	✓	✓
Non-employment	*+ ✓	*+ ✓		✓

- Multinomial logit model with 3 outcomes:
  - Not-employed
  - Self-employed
  - Employee
- For the sample of those for whom the outcome is self-employment or employment, we additionally model:
  - New working hours
  - New hourly wage
- Additionally, for the self-employed whose monthly pay is reduced in comparison to the baseline, and below the level of eligibility for SEISS, a logit model to predict the probability of SEISS grant.

Multinomial logit

- Outcome depends on:
  - Age
  - Gender
  - Interaction of age and gender
  - Education (Low / Medium / High)
  - Household composition (Single adult no children; Single adult living with children; Multiple adults no children; Multiple adults living with children)
  - Industry (13 categories, SIC07 classification with some extra aggregation)
  - Lagged hours of work + interaction with gender
  - Quintile of monthly pay in the baseline (Jan – Feb wave)
  - Lagged monthly pay
  - Occupation (9 categories, SOC 2010 aggregated up to 1 –digit)
  - Month

Hours and wages

- Same set of covariates as for transitions from employment
- Modelled jointly using the seemingly unrelated regressions, to allow the errors associated with the dependent variables to be correlated.
- Wages depend on gender, age, interaction of gender and age, household composition, baseline hours of work, baseline hourly pay, industry, occupation, education, month
- Hours use the same set of independent variables + hourly pay (contemporaneous)

SEISS grant

- Logit model. Depends on:
  - Age
  - Gender
  - Interaction of age and gender
  - Education (Low / Medium / High)
  - Household composition (Single adult no children; Single adult living with children; Multiple adults no children; Multiple adults living with children)
  - Industry (13 categories, SIC07 classification with some extra aggregation)
  - Lagged hours of work + interaction with gender
  - Quintile of monthly pay in the baseline (Jan – Feb wave)
  - Lagged monthly pay
  - Occupation (9 categories, SOC 2010 aggregated up to 1 –digit)
  - Month
  - Lagged SEISS

Transitions from non-employment

- Multinomial logit model with 3 outcomes:
  - Not-employed
  - Self-employed
  - Employee

From / To	Employment	Self-employment	Furlough	Non-employment
Employment	No changes / some changes* ✓	✓	✓	✓
Self-employment	* ✓	* / SEISS ✓		✓
Furlough	* ✓	* ✓	✓	✓
Non-employment	*+ ✓	*+ ✓		✓

- For the sample of those for whom the outcome is self-employment or employment, we additionally model:
  - New working hours
  - New hourly wage
- Additionally, we model the new industry and occupation

Multinomial logit

- Outcome depends on:
  - Age
  - Gender
  - Interaction of age and gender
  - Education (Low / Medium / High)
  - Household composition (Single adult no children; Single adult living with children; Multiple adults no children; Multiple adults living with children)
  - Lagged furlough state
  - Month

Hours and wages

- Same set of covariates as for transitions from employment and self-employment for wages
- For the hours, only the following:
  - Age
  - Gender
  - Interaction of age and gender
  - Household composition
  - Industry
  - Occupation
  - Education
  - Month

Industry and occupation

- Multinomial logit models with 13 outcomes for the industry model, and 9 outcomes for the occupation model.
- Very simple specification, estimated on the baseline (ideally would want to estimate it on those moving from non-employment to employment in each month, but not enough data)
  - Age
  - Gender
  - Education

Results: transition matrices Monthly

Jan => April

	To					
		Employee	Furlough	Self-emp	Not-emp	Total
From	Employee	10557	5024	4	377	15962
		66.14	**31.47**	**0.03**	**2.36**	**100.00**
		97.44	100.00	0.17	9.69	72.28
	Self-emp	61	0	2298	111	2470
		2.47	**0.00**	**93.04**	**4.49**	**100.00**
		0.56	0.00	98.33	2.85	11.18
	Not-emp	216	0	35	3402	3653
		5.91	**0.00**	**0.96**	**93.13**	**100.00**
		1.99	0.00	1.50	87.46	16.54
	Total	10834	5024	2337	3890	22085
		49.06	22.75	10.58	17.61	100.00
		100.00	100.00	100.00	100.00	100.00

Jan => April, weighted

	To					
	state_Jan Feb	Employee	Furlough	Self-emp	Not-emp	Total
From	Employee	17506185	8696246	6186	654242	26862859
	Self-emp	107829	0	3685213	183428	3976470
	Not-emp	346555	0	52880	5340817	5740252
	Total	17960569	8696246	3744279	6178487	36579581

Results: transition matrices Monthly

April => May

		To					
		Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	9514	1074	2	0	244	10834
		87.82	**9.91**	**0.02**	**0.00**	**2.25**	**100.00**
		88.54	21.37	0.24	0.00	6.22	49.06
	Furlough	1054	3952	2	0	16	5024
		20.98	**78.66**	**0.04**	**0.00**	**0.32**	**100.00**
		9.81	78.63	0.24	0.00	0.41	22.75
	Self-emp	7	0	801	1490	39	2337
		0.30	**0.00**	**34.27**	**63.76**	**1.67**	**100.00**
		0.07	0.00	97.56	95.09	0.99	10.58
	Not-emp	171	0	16	77	3626	3890
		4.40	**0.00**	**0.41**	**1.98**	**93.21**	**100.00**
		1.59	0.00	1.95	4.91	92.38	17.61
	Total	10746	5026	821	1567	3925	22085
		48.66	22.76	3.72	7.10	17.77	100.00
		100.00	100.00	100.00	100.00	100.00	100.00

April => May, weighted

		To					
	state_April	Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	15684326	1882668	2046	0	391529	17960569
	Furlough	1850488	6814756	2642	0	28360	8696246
	Self-emp	11268	0	1253085	2421320	58606	3744279
	Not-emp	289362	0	26969	121085	5741071	6178487
	Total	17835444	8697424	1284742	2542405	6219566	36579581

Results: transition matrices Monthly

May => June

		To					
		Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	8390	2121	29	0	206	10746
		78.08	**19.74**	**0.27**	**0.00**	**1.92**	**100.00**
		82.68	39.12	2.95	0.00	5.11	48.66
	Furlough	1526	3301	99	0	100	5026
		30.36	**65.68**	**1.97**	**0.00**	**1.99**	**100.00**
		15.04	60.88	10.07	0.00	2.48	22.76
	Self-emp	16	0	786	1	18	821
		1.95	**0.00**	**95.74**	**0.12**	**2.19**	**100.00**
		0.16	0.00	79.96	0.07	0.45	3.72
	SEISS	46	0	0	1496	25	1567
		2.94	**0.00**	**0.00**	**95.47**	**1.60**	**100.00**
		0.45	0.00	0.00	99.93	0.62	7.10
	Not-emp	170	0	69	0	3686	3925
		4.33	**0.00**	**1.76**	**0.00**	**93.91**	**100.00**
		1.68	0.00	7.02	0.00	91.35	17.77
	Total	10148	5422	983	1497	4035	22085
		45.95	24.55	4.45	6.78	18.27	100.00
		100.00	100.00	100.00	100.00	100.00	100.00

May => June, weighted

		To					
	state_May	Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	13783428	3662622	52713	0	336681	17835444
	Furlough	2664211	5712532	147683	0	172998	8697424
	Self-emp	23939	0	1229653	2482	28668	1284742
	SEISS	72180	0	0	2421929	48296	2542405
	Not-emp	284144	0	109286	0	5826136	6219566
	Total	16827902	9375154	1539335	2424411	6412779	36579581

Results: transition matrices
Single transition

Jan => June

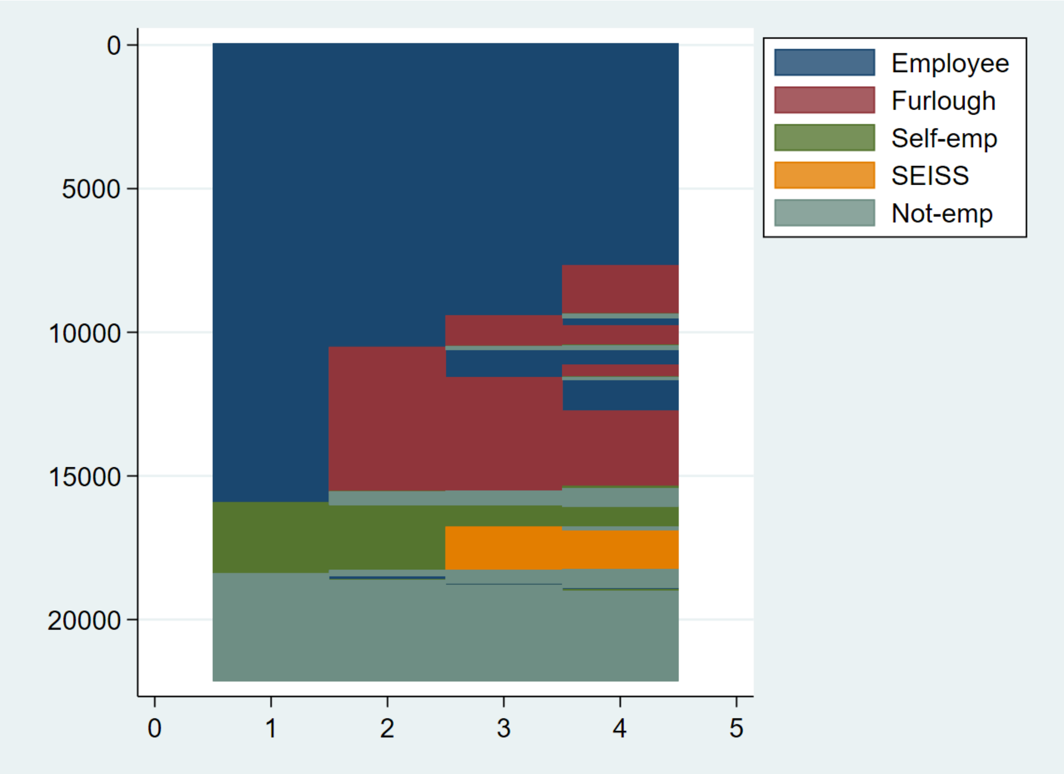
		To					
		Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	9893	5394	144	0	531	15962
		61.98	**33.79**	**0.90**	**0.00**	**3.33**	**100.00**
		97.49	99.48	14.65	0.00	13.16	72.28
	Self-emp	103	28	776	1497	66	2470
		4.17	**1.13**	**31.42**	**60.61**	**2.67**	**100.00**
		1.01	0.52	78.94	100.00	1.64	11.18
	Not-emp	152	0	63	0	3438	3653
		4.16	**0.00**	**1.72**	**0.00**	**94.11**	**100.00**
		1.50	0.00	6.41	0.00	85.20	16.54
	Total	10148	5422	983	1497	4035	22085
		45.95	24.55	4.45	6.78	18.27	100.00
		100.00	100.00	100.00	100.00	100.00	100.00

Jan => June, weighted

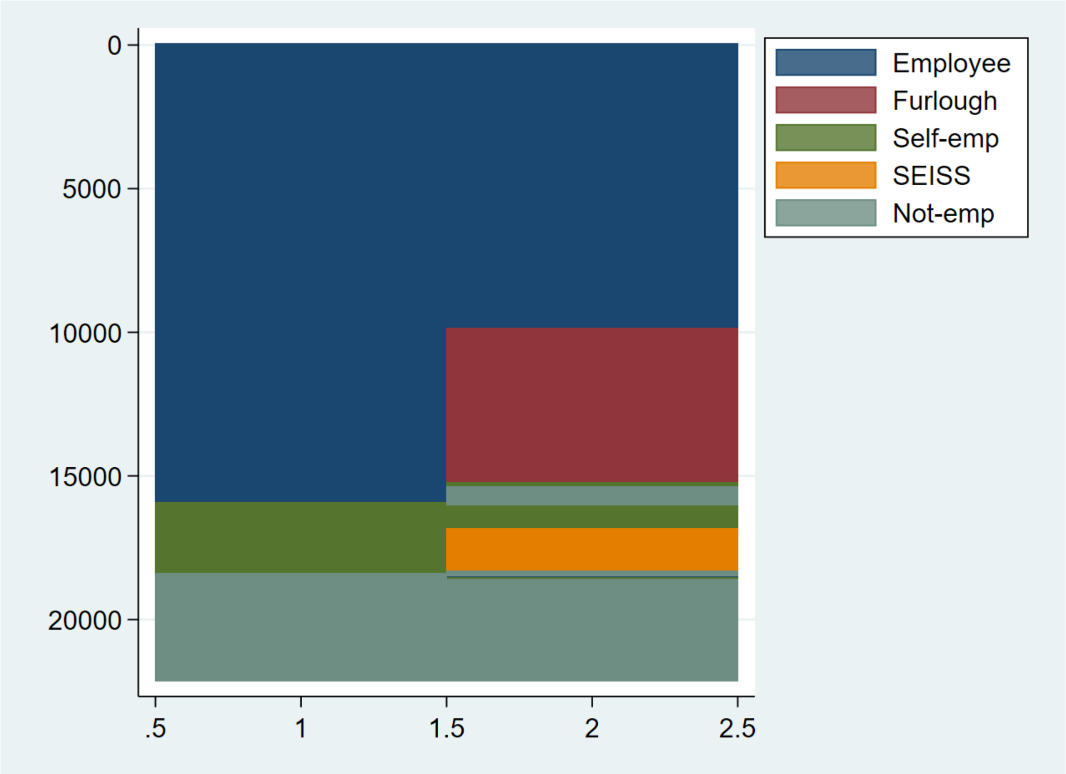
		To					
	state_Jan Feb	Employee	Furlough	Self-emp	SEISS	Not-emp	Total
From	Employee	16411710	9325546	225090	0	900513	26862859
	Self-emp	169321	49608	1214773	2424411	118357	3976470
	Not-emp	246871	0	99472	0	5393909	5740252
	Total	16827902	9375154	1539335	2424411	6412779	36579581

Monthly vs single transition

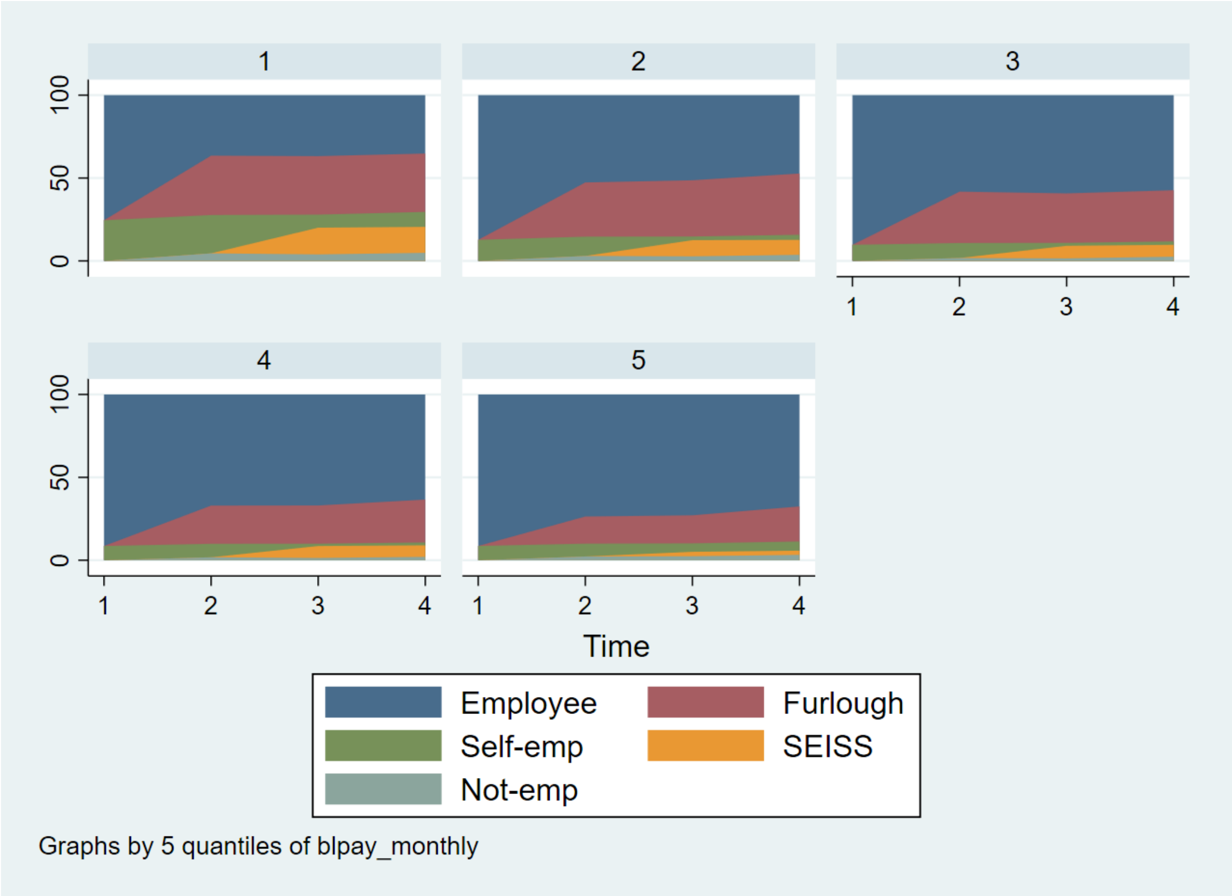
Monthly



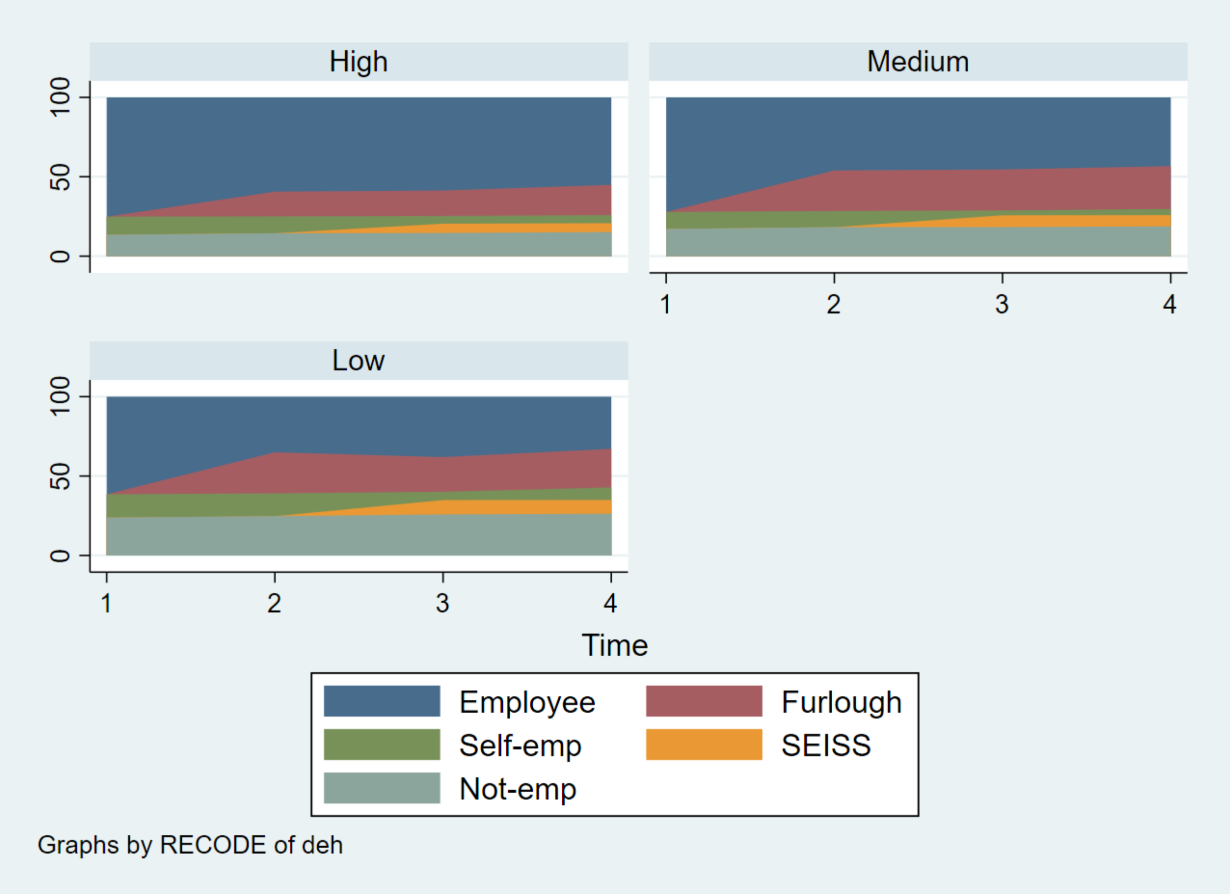
Single



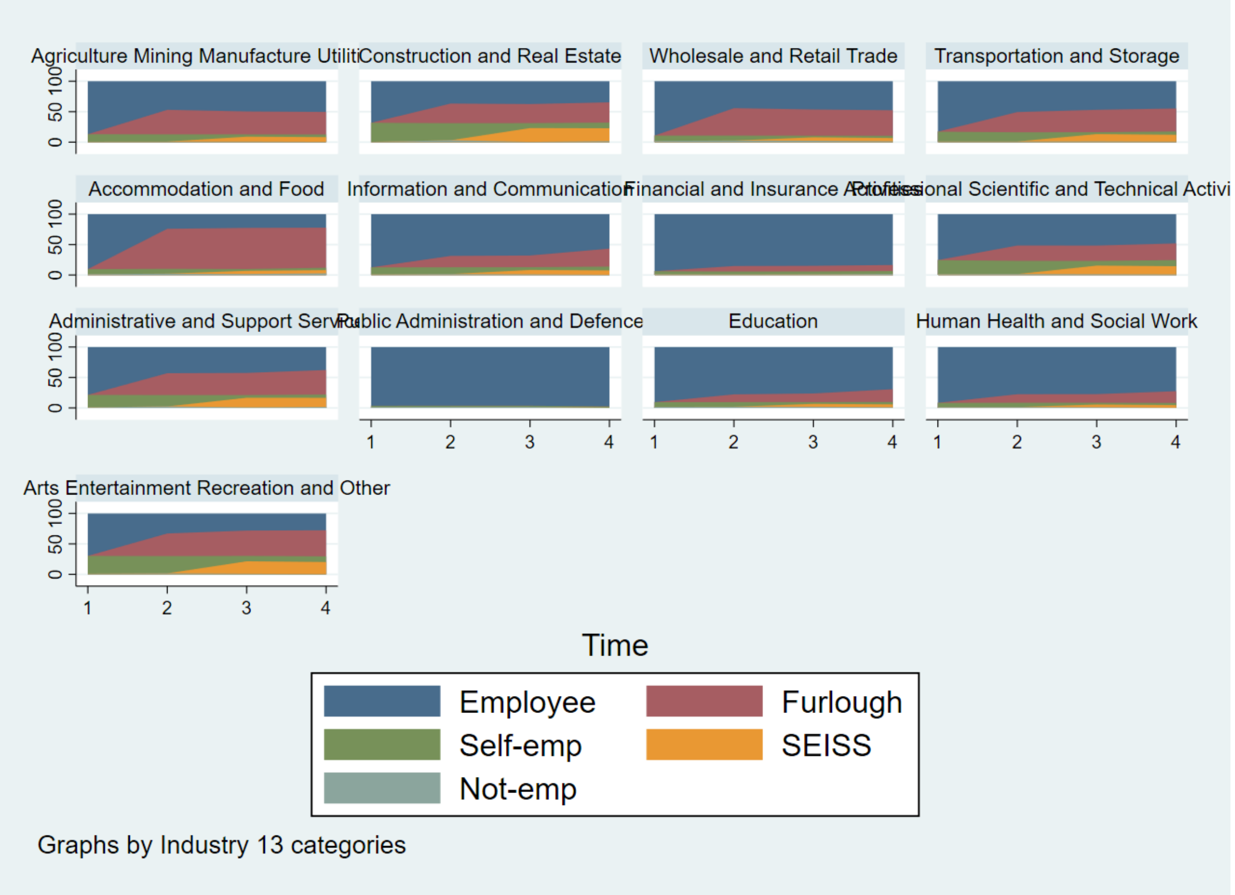
Monthly transitions by: baseline pay quintile



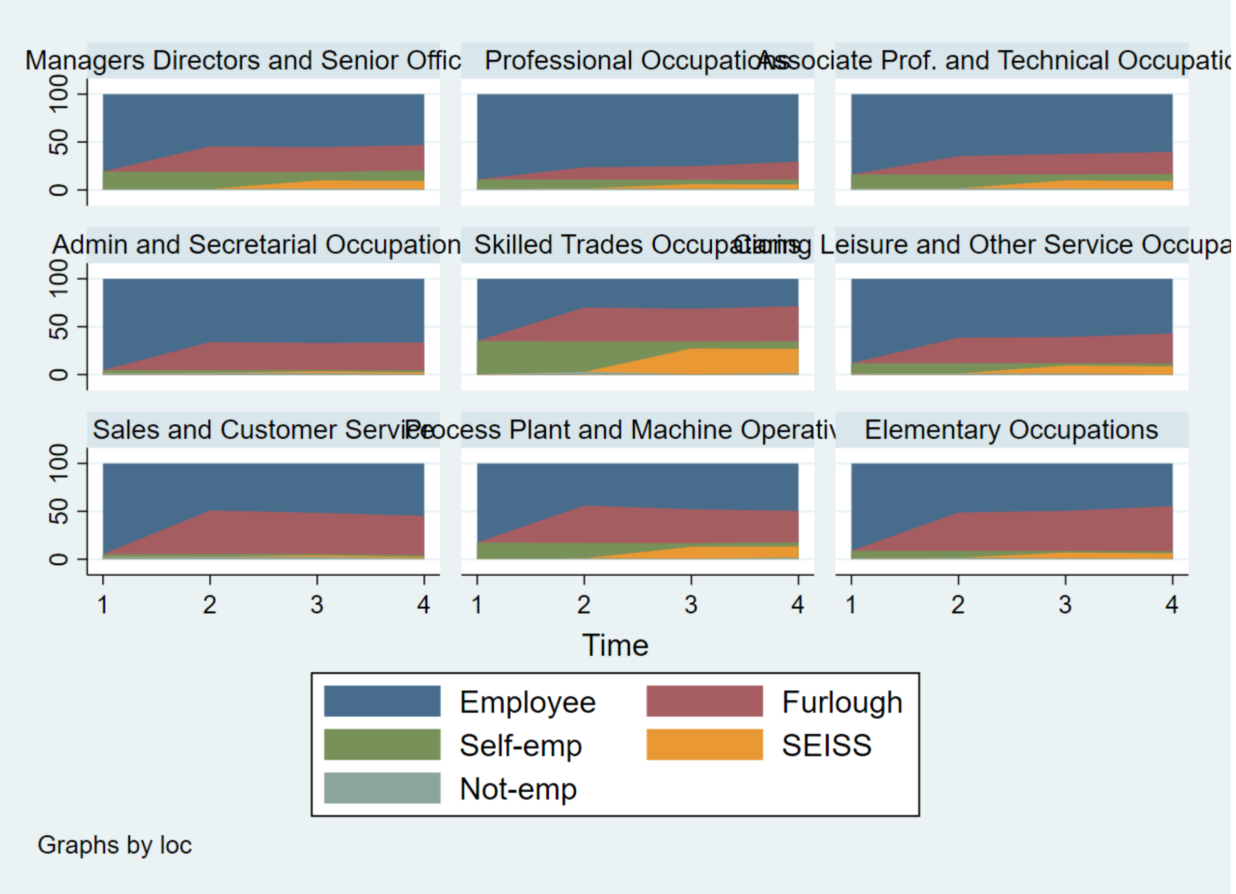
Monthly transitions by: education



Monthly transitions by: industry

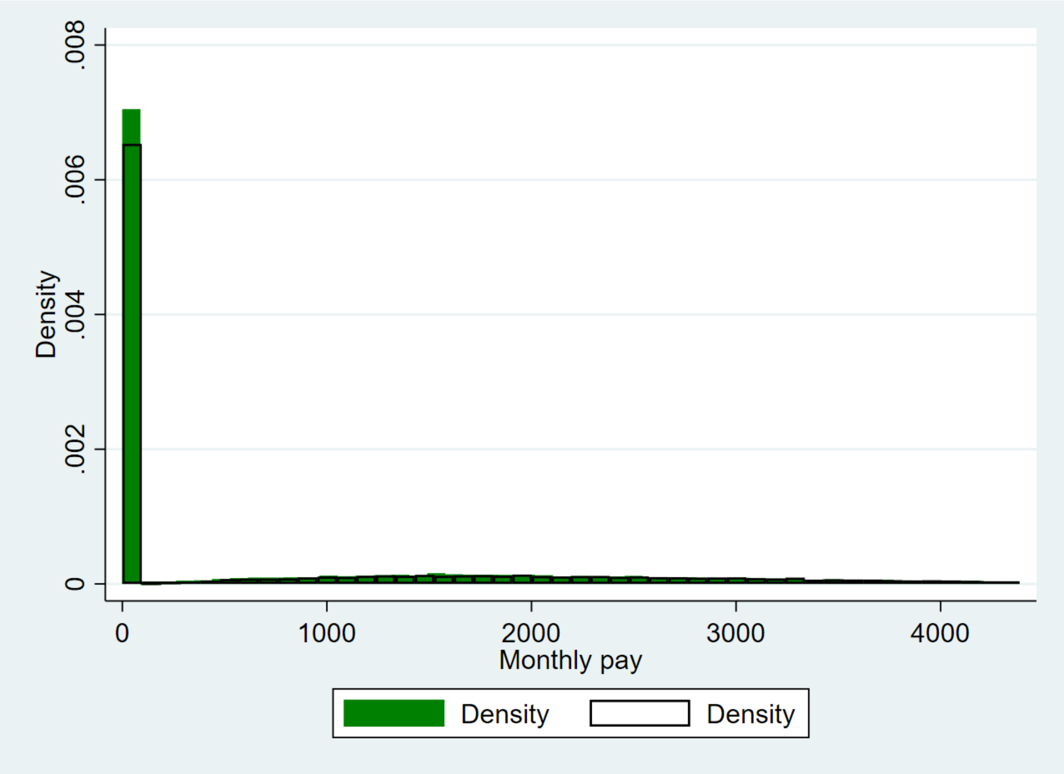


Monthly transitions by: occupation

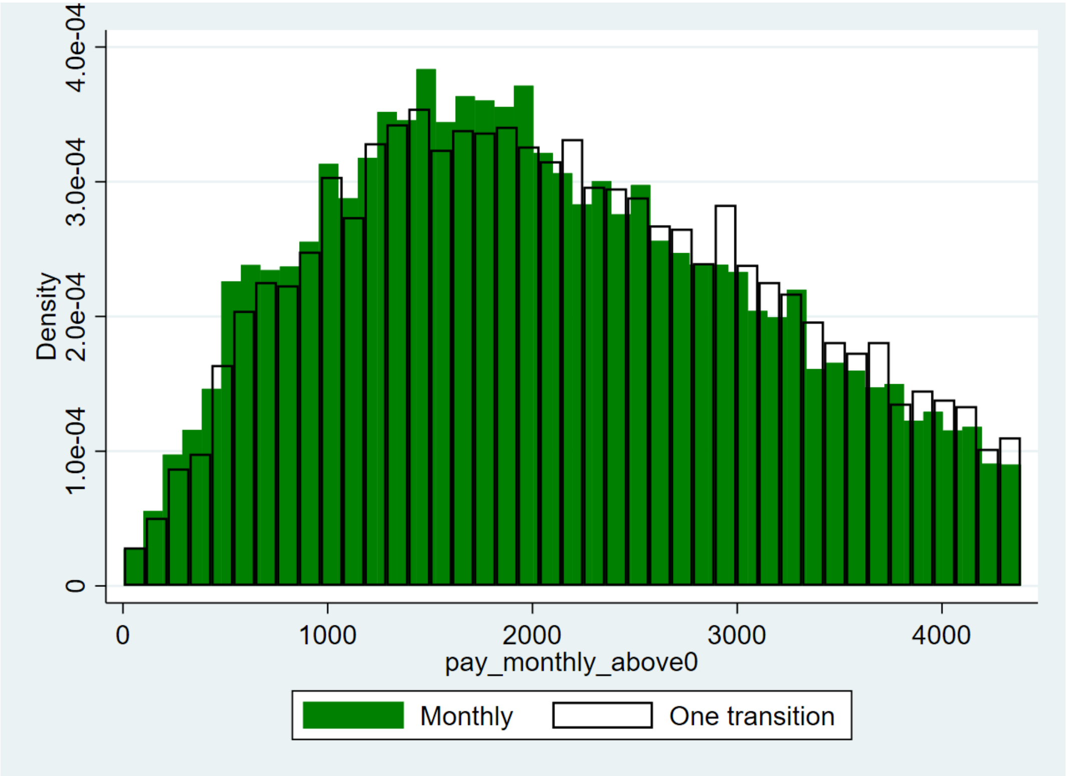


Market income distribution

With zero incomes

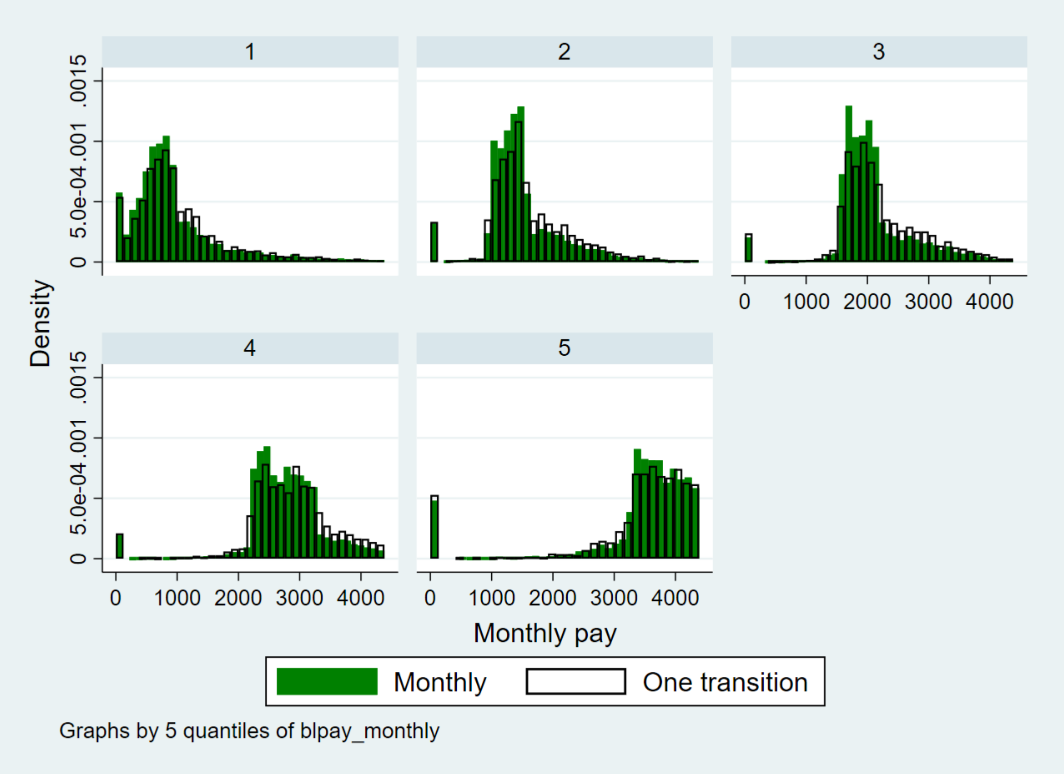


Positive incomes only

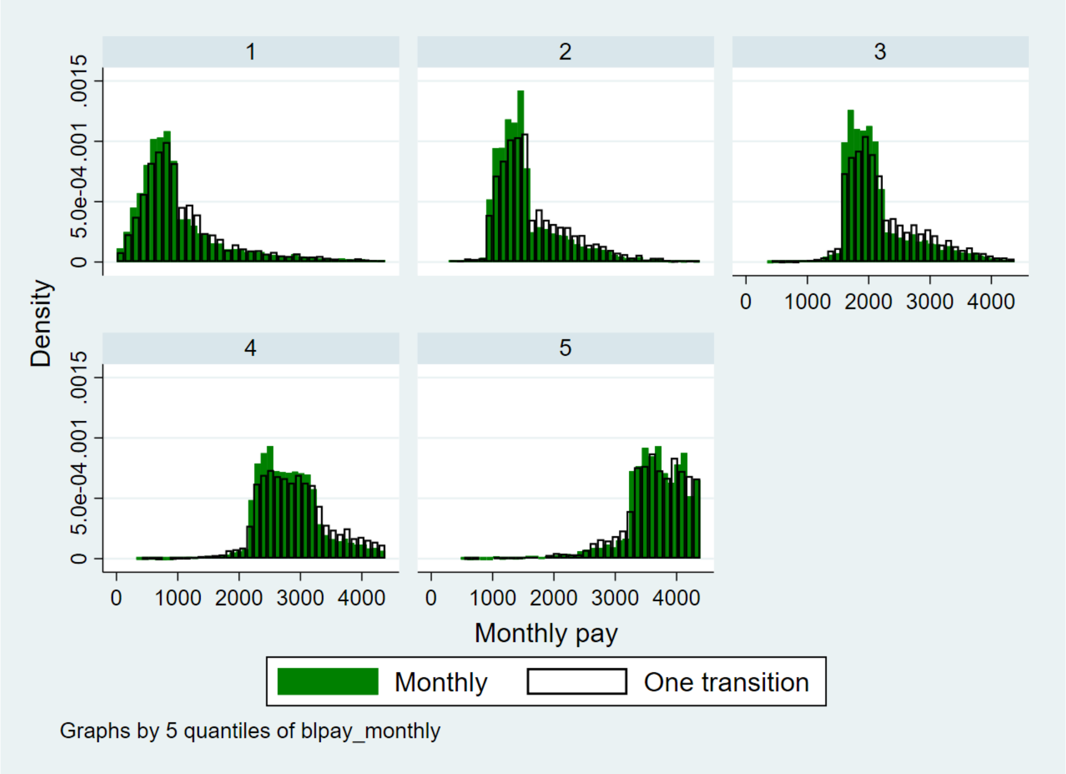


Market income distribution, by baseline pay

With zero incomes



Positive incomes only



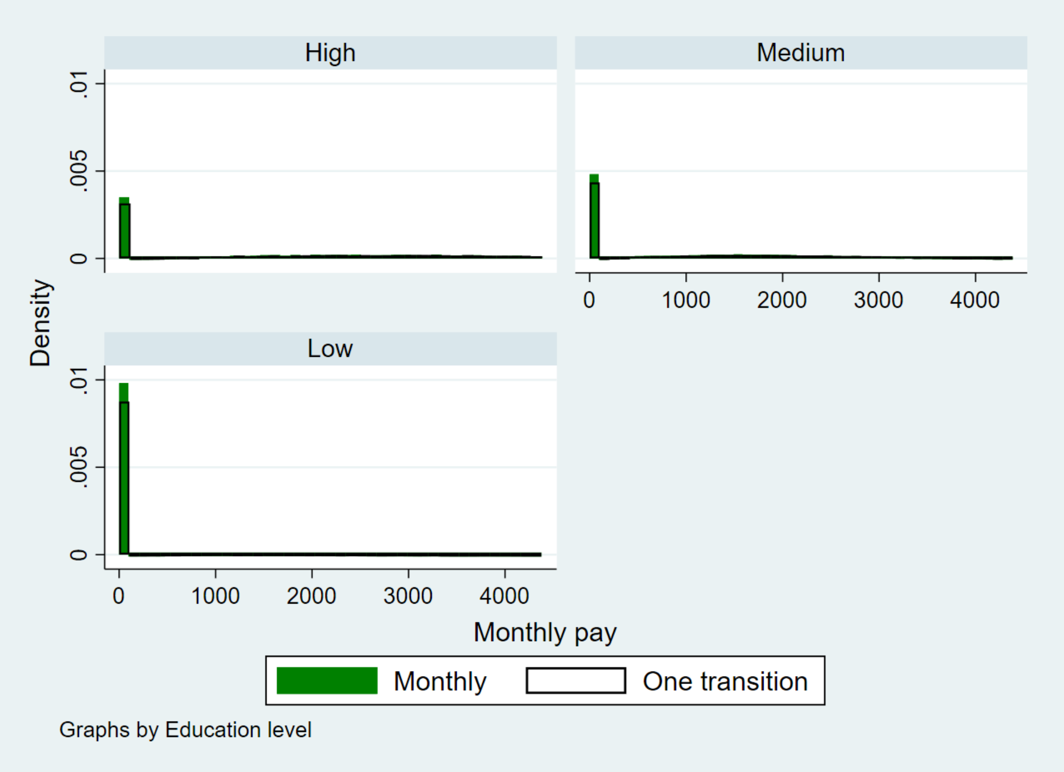
CeMPA Centre for Microsimulation and Policy Analysis

iSER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

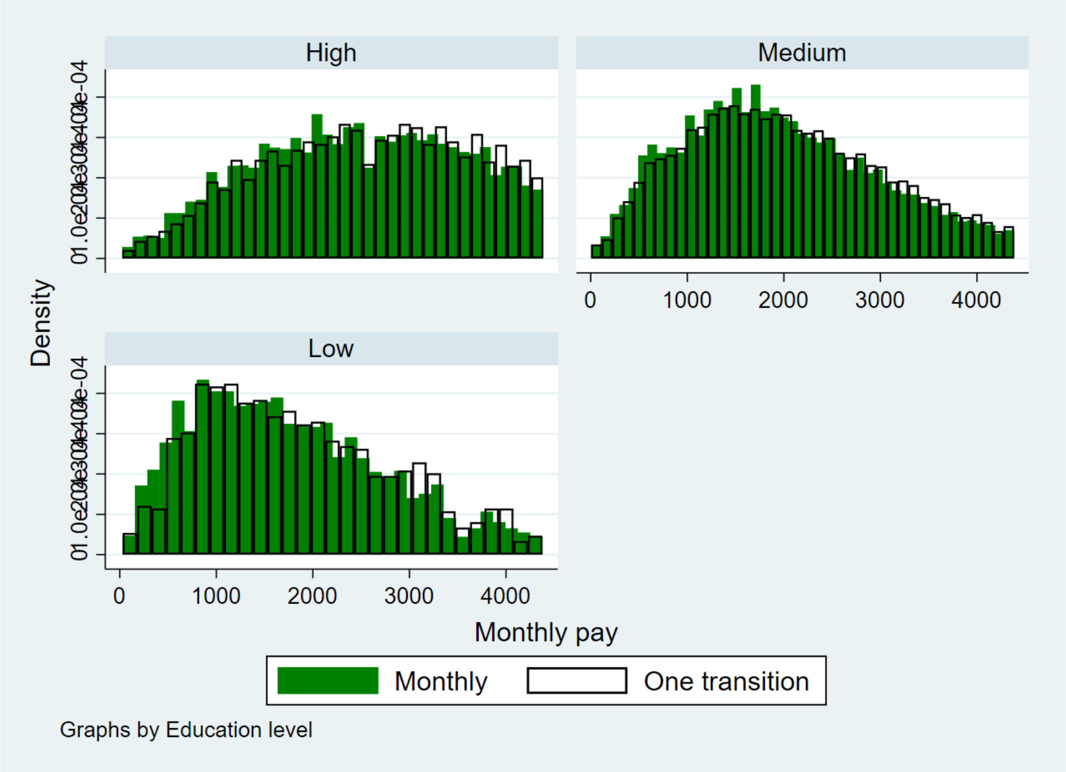
University of Essex

Market income distribution, by education

With zero incomes



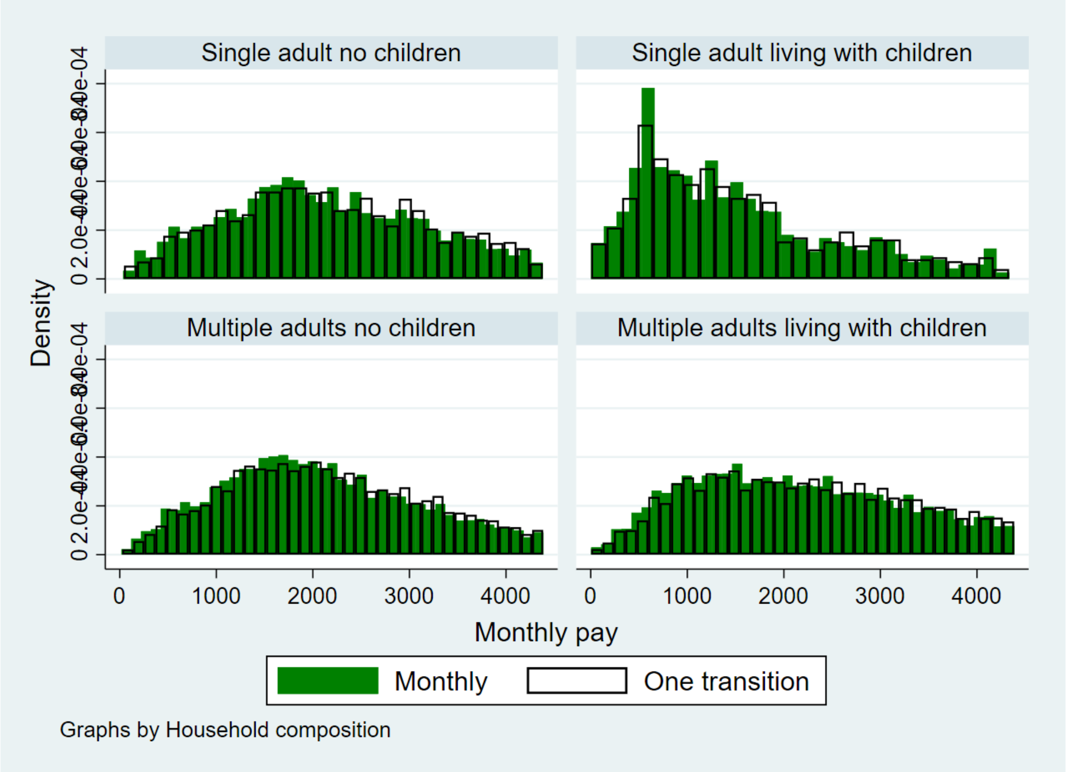
Positive incomes only



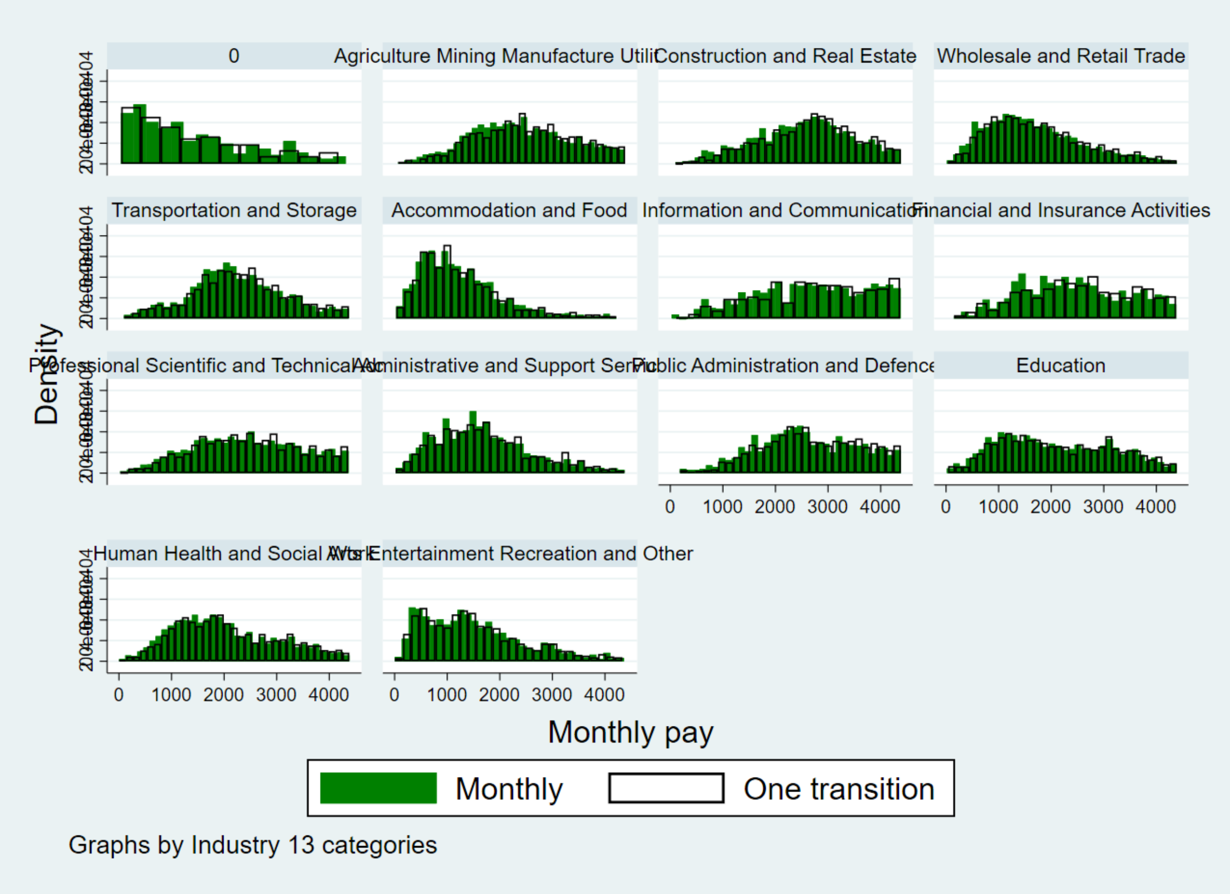
Market income distribution, by household composition

With zero incomes

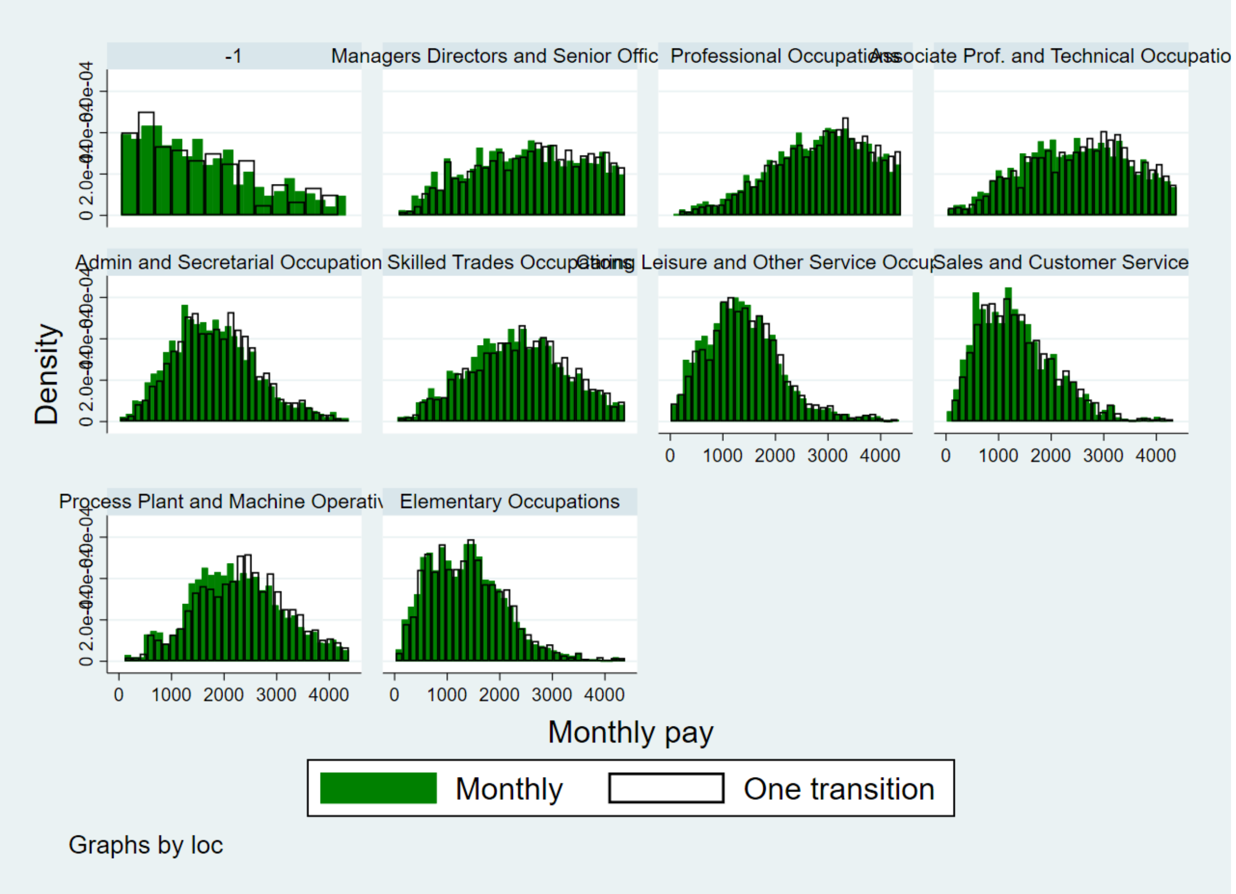
Positive incomes only



Market income distribution, by industry, positive incomes only



Market income distribution, by occupation, positive incomes only



(Preliminary) Conclusions

1. Simulations with *representative months* are misleading (especially at times of Covid)
2. Using current income rather than yearly income has implications for the measurement of inequality and poverty (especially at times of Covid).

Thanks !

matteo.richiardi@essex.ac.uk