# Estimating intergenerational income mobility in New Zealand

Leon Iusitini Doctoral candidate Ieon.iusitini@aut.ac.nz

> Supervisors: Prof Gail Pacheco Dr Michael Fletcher Dr Alexander Plum

## **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 author, not Statistics New Zealand.

## Aim and rationale

- Estimate degree of intergenerational income mobility (persistence) between parents and their offspring in New Zealand
- This presentation:
  - *Quantifies* persistence/mobility (doesn't attempt to explain/decompose it)
  - Presents results for sons and fathers only
- Intergenerational mobility is an (imperfect) indicator of equality of opportunity

## Literature

- Vast literature, but consensus on two findings:
  - 1. Measurement error now better recognised and matters a lot to estimates of intergenerational mobility
    - Attenuation bias from transitory shocks (Solon, 1992; Mazumder, 2005)
    - 'Lifecycle bias' from heterogeneous income profiles (Jenkins, 1987; Haider & Solon, 2006; Nybom & Stuhler, 2016)
  - 2. Intergenerational mobility higher in Nordic countries (Denmark, Finland, Norway, Sweden) and Canada than UK and US (Björklund & Jantti, 2009; Bratsberg et al., 2007; Corak et al., 2014; Solon, 2002)
- Only two New Zealand studies (Gibbons, 2010; Andrews & Leigh, 2008)
- I estimate intergenerational income mobility with census microdata linked over time
- I construct a proxy for lifetime average income for sons and fathers

## **Measurement of intergenerational mobility**

 Standard approach to measuring association between offspring's and parents' incomes is to apply OLS to

$$Y_{ij}^{\text{son}} = \alpha + \beta Y_i^{\text{father}} + Age^{\text{son}} + Age^{2 \text{ son}} + Age^{\text{father}} + Age^{2 \text{ father}} + \varepsilon_i$$

- $Y_{ij}^{\text{son}}$  = log of lifetime income of son *j* in family *i*
- $Y_i^{\text{father}} = \log \text{ of lifetime income of father in family } i$
- $\varepsilon_i$  = error term capturing factors  $\perp$  to  $Y_i^{\text{father}}$

•  $\beta$  = intergenerational income elasticity (IGE) = 'regression to the (geometric) mean'

## **New Zealand Longitudinal Census**

- NZLC links individuals' census records over time
- Each census linked backwards to previous census to create six pairs:

2013-2006, 2006-2001, 2001-1996, 1996-1991, 1991-1986, 1986-1981

- Individuals can be linked across up to seven censuses (32 year time-span)
- Person is linkable if at earlier census they had been born, filled out a census form, and resided in New Zealand
- Linking was largely deterministic (70%) based on sex, date of birth, area of usual residence (country of birth, Māori descent), not names
- A further 3% linked probabilistically
- 72% average link rate between census pairs, 32% across 2006-1981

## **New Zealand Longitudinal Census**

- People who change address frequently are less likely to be linked, those overseas during a census cannot be linked back over period of their absence
- Groups less likely to be linked:
  - Young adults (especially those in their 20s)
  - Males
  - Māori, Pacific, Asian
  - People living in more socio-economically deprived areas
- Potential for selection bias

## **New Zealand Longitudinal Census**

• 1981 census defined a family as:

"...a husband and wife with or without never married children of any age or a lone parent with one or more never married children, living in a private household. The term 'children' includes step children and adopted children, but not foster children....A family is not necessarily all related people in a household, but only those related by blood, marriage, or adoption, who normally live together as a single family unit and who are present on census night"

- Fathers linked to sons by variables 'Family ID number' and 'Role in family'
- I use the father enumerated in 1981 and track his income over time, regardless of any changes in the son's father/father figure over time
- Fathers may have multiple sons in the 1981 census, hence there are brothers in my sample
- In the NZLC, there are 209,607 sons aged 0 to 14 on census night 1981 who were born in New Zealand and had a father enumerated in the family

## **Full son-father sample**

• NZLC linkage varies among these 209,607 sons:



## **Full son-father sample**

• NZLC linkage varies among these 209,607 sons:



## Income data in NZLC

- All censuses from 1981 through 2013 collected data on 'total personal income':
  - gross annual income from all sources over previous 12 months
  - self-reported
  - banded (interval-censored, top-coded, bottom-coded)
- Classifications (bands) changed over time
- I assign each band the *median* of the band (calculated by Statistics New Zealand from more granular data, mostly HES), *except for 1981 where mid-points were used*
- Then deflate the medians to 2012 Q3 dollars using the CPI
- Then recode zero incomes to \$1

## **Proxying for lifetime average income**

- I use Mazumder's (2016) method of taking a time average centred at an age at which current income is known to be representative of lifetime average income
- No New Zealand studies estimating this age, but studies in other countries have found:

Country	Age window	Study
USA	"between early thirties and mid-forties"	Haider and Solon (2006)
Norway	"late thirties and early forties"	Nilsen et al. (2012)
Germany	30-40	Brenner (2010)
Sweden	34-40	Böhlmark and Lindquist (2006)
Sweden	33	Nybom and Stuhler (2016)

 My proxy for lifetime income will take a multiyear average of income (else, for sons, a single year observation) centred at approx. age 35 years

# **Proxying for sons' lifetime average income**

		Census year						
Year son born		1981	1986	1991	1996	2001	2006	2013
1967	Son's age	14	19	24	29	34	39	46

Lifetime income proxy:	Income observations used:						
	1991	1996	2001	2006	2013		
	24	29	34	39	46	35.0 years	
			else				
		1996	2001	2006	Ι		
		29	34	39		34.0 years	
Multiyear average			else				
	1991	1996	2001	2006	2013		
	24				46	35.0 years	
			else				
		1996	2001	2006			
		29		39		34.0 years	
else			else				
<b>.</b>			2001	]			

34

34.0 years

Single year observation

# **Proxying for sons' lifetime average income**

		Census year						
Year son born		1981	1986	1991	1996	2001	2006	2013
1981		0	5	10	15	20	25	32
1980		1	6	11	16	21	26	33
1979		2	7	12	17	22	27	34
1978		3	8	13	18	23	28	35
1977		4	9	14	19	24	29	36
1976		5	10	15	20	25	30	37
1975		6	11	16	21	26	31	38
1974	Son's age	7	12	17	22	27	32	39
1973		8	13	18	23	28	33	40
1972		9	14	19	24	29	34	41
1971		10	15	20	25	30	35	42
1970		11	16	21	26	31	36	43
1969		12	17	22	27	32	37	44
1968		13	18	23	28	33	38	45
1967		14	19	24	29	34	39	46

# **Proxying for sons' lifetime average income**

		Census year						
Year son born		1981	1986	1991	1996	2001	2006	2013
1981		0	5	10	15	20	25	32
1980		1	6	11	16	21	26	33
1979		2	7	12	17	22	27	34
1978		3	8	13	18	23	28	35
1977		4	9	14	19	24	29	36
1976		5	10	15	20	25	30	37
1975		6	11	16	21	26	31	38
1974	Son's age	7	12	17	22	27	32	39
1973		8	13	18	23	28	33	40
1972		9	14	19	24	29	34	41
1971		10	15	20	25	30	35	42
1970		11	16	21	26	31	36	43
1969		12	17	22	27	32	37	44
1968		13	18	23	28	33	38	45
1967		14	19	24	29	34	39	46

# **Proxying for fathers' lifetime average income**

		Census year						
Year father born		1981	1986	1991	1996	2001	2006	2013
1966		15	20	25	30	35	40	47
1965		16	21	26	31	36	41	48
1964		17	22	27	32	37	42	49
1963		18	23	28	33	38	43	50
1962		19	24	29	34	39	44	51
1961		20	25	30	35	40	45	52
1960		21	26	31	36	41	46	53
1959		22	27	32	37	42	47	54
1958		23	28	33	38	43	48	55
1957	Father's age	24	29	34	39	44	49	56
1956		25	30	35	40	45	50	57
1955		26	31	36	41	46	51	58
1954		27	32	37	42	47	52	59
1953		28	33	38	43	48	53	60
1952		29	34	39	44	49	54	61
1951		30	35	40	45	50	55	62
1950		31	36	41	46	51	56	63
1949		32	37	42	47	52	57	64
1948		33	38	43	48	53	58	65

## **Selected son-father sample**



- For whom a proxy could be calculated
- Proxy was positive
- Had a father for whom a positive proxy could also be calculated

#### **Sample representativeness**

■ Full son-father sample (n=209,607)

□ Selected son-father sample (n=4,617)



## Sons' and fathers' lifetime average incomes



Lifetime income proxy (in 2012 Q3 dollars)

## **Estimating the IGE**

- Take the log of sons' and fathers' lifetime income proxies
- Using OLS, regress sons' log lifetime income on fathers' log lifetime income
- Controls for:
  - Son's age at 1981 census
  - Son's age squared
  - Father's age at 1981 census
  - Father's age squared
- Standard errors adjusted for clustering within families (selected sample contains brothers)

## **Regression results**

Dependent variable: Son's log lifetime income

Father's log lifetime income	0.239 ***
	(0.038)
Son's age at 1981 census	0.053 **
	(0.019)
Son's age squared	-0.003 *
	(0.001)
Father's age at 1981 census	0.221
C C	(0.116)
Father's age squared	-0.004
0	(0.002)
Constant	4.826 **
	(1.727)
R-squared	0.0122
Number of observations	4,617

Robust standard errors (adjusted for 4,416 clusters in families) in parentheses \*\*\*p<0.001 \*\*p<0.01 \*p<0.05

#### **Robustness checks**



#### **Cross-national comparisons of son-father IGE**



## **Conclusions and next steps**

- Linked census data has considerable limitations (self-report, banded income data, attrition, etc.)
- Tentatively, intergenerational income persistence appears to be at the lower end of the cross-national spectrum
- Next steps:
  - Repeat for daughter-mother, daughter-father, son-mother dyads
  - Estimate IGE for different subgroups (where sample size permits)
  - Decompose IGE into pathways through which parental income is 'transmitted' to offspring