Effects of Taxes and Safety-Net Pensions on Life-cycle Labor Supply, Savings and Human Capital: The Case of Australia

Fedor Iskhakov,
Australian National University and CEPAR

Michael Keane,
University of New South Wales and CEPAR

DSE Summer School - Bonn
August 17, 2021

Overview

This paper uses a dynamic structural model to assess the impact of the Australian social security system on labor supply, savings, human capital investment and consumer welfare.

We use the model to do experiments to see if the system could be improved.

In this presentation I will:

- Give the background to the substantive problem
- Explain how the paper contributes to the literature methodologically
- Present the model and results

The next presentation will discuss the method for solving the model in detail.

The Australian Social Security System

The Australian social security system is ranked among the best in the world by Mercer, the OECD, IMF etc.

Two components:

1. "Superannuation"
    - Defined contribution pension plan
    - Mandatory employer contributions to private accounts (9.5%)
    - Workers choose among investment options
    - Accessible from age 65 (Age 60 if retired)
    - Avoids fiscal burden on government
2. Age Pension
    - Provides safety net at ages 65 +
    - Benefits do not depend on work history (unlike SS in US)
    - Pure means-tested transfer ("welfare") program

Age Pension Benefit Rule

Estimate Effective Income and Asset Taper Rates

$$\text{benefit}_{\max} = \underset{(183.96)}{10,759.73} + \underset{(173.52)}{1,846.92}(\text{when year} \geq 2010),$$

$$\text{pension} = \max\left\{\text{benefit}_{\max} - \max\left[\max\{\underset{(0.020)}{0.27794}\,\text{income}, \underset{(0.0004)}{0.00499}(\text{wealth} - \underset{(46,895.27)}{117,082.60})\}, 0\right]\right\}$$

Age Pension Benefits

We estimate effective taper rates of only:

- 27.8% on Income
- 0.5% on Assets

The low taper rates lead to very poor targeting (75% get some benefits). This means the Age Pension is a large program:

- Income Taxes = $ 180 bil. (2014)
- Age Pension = $ 50 bil. (2014)

Goal: Use structural model to assess:

- Effects of Age Pension on:
  - Labor supply
  - Asset and human capital accumulation
- Effects of changes in Age Pension rules designed to improve targeting of benefits

Literature

This paper extends the literature on dynamic life-cycle labor supply models with human capital:

Keane and Wolpin (IER, 2001) was the first paper to structurally estimate a dynamic life-cycle model with both human capital and saving. Imai and Keane (IER, 2004) followed (focusing on how HC affects labor supply elasticities).

In the 2001 paper labor supply was discrete: (FT, PT, 0), while in the 2004 paper it was continuous.

Keane and Wasi (EJ, 2016) was first paper with assets, HC and mixed discrete/continuous hours.

The present paper develops an arguably more realistic model where hours are bunched at several discrete levels. And a better algorithm to handle mixed discrete/continuous choices (hours/savings).

Literature

Means-tested transfers have potentially important effects on asset and human capital accumulation.

Yet there are very few papers estimating dynamic models with means-tested transfers:

- Keane and Wolpin (IER, 2010) - which is basically the "Career Decisions of Young Women"
- Blundell, Costa-Dias, Meghir, Shaw (ECMA, 2016)

These papers do not focus on the targeting issue we emphasize here.

Our Life-Cycle Labor Supply Model

1. Discrete time = Age from 19 to 100 (stochastic survival)
2. Annual decisions on:
   - Consumption/Saving (continuous choice)
   - Hours chosen from [0, 24, 40, 45, 50, 60] per week (discrete choice)
   - Previous Life-cycle labor supply models have not accounted for bunching of hours
3. Human capital accumulation
   - Learning-by-doing
4. We model Age Pension, Superannuation and Tax Rules
5. Observed and unobserved heterogeneity
   - Education $\rightarrow$ Shifts human capital production function
   - Unobserved types $\rightarrow$ Shifts skill endowment and tastes for leisure

Our Life-cycle labor Supply Model

Hours of labor supply $h_t \in H$ (choice variable)

Human capital: $K_t = f\left(\sum_{\tau=1}^{t-1} h_\tau, \text{age}, \text{education}, \text{type}\right)$

Wage: $wage_{t+1} = K_t \cdot R_t \cdot \epsilon_{t+1}^{wage}$,

- $R_t = 1$ is rental rate on human capital,
- Wage draw: $\epsilon_t^{wage} \sim lnN(0, \sigma_t^{wage})$
- Timing: h_t chosen based on K_t, wage draw revealed at $t+1$

M_t = Consumable wealth in the beginning of the period

Consumption $c_t \leq M_t + a_0$ (credit constraint)

Intertemporal budget constraint
$M_{t+1} = \left(M_t - c_t\right)(1+r) + h_t \cdot wage_{t+1} - Tax_{t+1} + transfers_{t+1}$

Our Life-cycle Labor Supply Model

Intertemporal budget constraint (Details on Transfers)

$$M_{t+1} = (M_t - c_t)(1+r) + h_t \cdot wage_{t+1} - Tax_{t+1} + transfers_{t+1}$$

$$\begin{aligned} M_{t+1} = (M_t - c_t)(1+r) + h_t \cdot wage_{t+1} - Tax_{t+1} \\ + pens_{t+1} \cdot \mathbb{1}\{t+1 \geq 65\} \\ + super_{t+1} \cdot \mathbb{1}\{t+1 = 65\} \\ + tr_{t+1} \cdot \mathbb{1}\{t+1 \leq 22\} \end{aligned}$$

where:

- $pens_{t+1}$ denotes Age Pension benefits,
- $super_{t+1}$ denotes the superannuation payment
- tr_{t+1} denotes transfers from parents to youth

The Pension and Super rules are estimated from data (see below)

Our Life-cycle Labor Supply Model

- Human Capital Production Function
- Let $\mathcal{E}_t$ denote the ratio of total work time to maximum work time up through t-1, i.e. "normalized" work experience, $0 \leq \mathcal{E}_t \leq 1$

$$\mathcal{E}_t = \frac{1}{t \cdot h_{max}} \sum_{\tau=1}^{t-1} h_\tau$$

$$K_t = \exp\left(\eta_{0,edu} + \eta_{0,type} + \eta_{1,edu} \cdot t\mathcal{E}_t + \eta_{2,edu} \cdot (t\mathcal{E}_t)^2 + \eta_3 t + \eta_4 t^2\right)$$

where $t \cdot \mathcal{E}_t$ is total work experience.

- Heterogeneity: education and type specific intercepts in wage function

Our Life-Cycle Labor Supply Model

- Preferences for Consumption and Bequests

$$u(c_t) = \frac{c_t^{1-\zeta} - 1}{1-\zeta}$$

$$w(B_t) = b_{scale} \cdot \frac{(B_t + a_0)^{1-\xi} - a_0^{1-\xi}}{1-\xi}$$

- $B_t = M_t - c_t$ is bequeathed wealth (if person dies at age t)
- $\zeta > 0$, $\xi > 0$, $b_{scale} > 0$ are parameters to be estimated
- a_0 = credit constraint (maximum amount of borrowing)

Our Life-cycle Labor Supply Model

- Preferences: Disutility of Work Hours

$$v_t(h_t) = \mathbb{1}\{h_t > 0\} \cdot \kappa_{type}(\tau_{uh}) \cdot \kappa_{age}(t) \cdot \gamma(h_t)$$

$\gamma = (\gamma^{(1)}, \ldots, \gamma^{(5)})$ disutilities of the five discrete hours levels

Type: high ($\kappa_{type} = 1$) or low ($\kappa_{type} = \kappa_1 > 1$)

Age effects:

$$\kappa_{age}(t) = 1 + \kappa_2 (t - 40)^2 \cdot \mathbb{1}\{t > 40\} + \kappa_3 (t - 25) \cdot \mathbb{1}\{t < 25\}$$

Age effects may proxy for declining health at older ages

Our Life-Cycle Labor Supply Model

- State vector $X_t = (M_t, \mathcal{E}_t, \text{education}, \text{type})$
- Bellman Equation

$$V_t(X_t) = \max_{\substack{0 \le c_t \le M_t + a_0, \\ h_t \in H_t}} \Big\{ u(c_t) - v_t(h_t, \tau_{uh}) \\ + \delta_t \beta(\tau_{edu}) E\left[V_{t+1}(X_{t+1}) \,|X_t, c_t, h_t \right] \\ + (1-\delta_t) w(M_t - c_t) \Big\},$$

Note: c_t continuous, h_t discrete

$\tau = (\tau_{uh}, \tau_{edu})$ types for education and taste of work

H_t choice set in period t

$\beta(\tau_{edu})$ discount factor dependent on education

δ_t survival probability

HILDA Data

Household, Income and Labor Dynamics in Australia survey (HILDA)

- The primary source of data is the Household, Income and Labor Dynamics in Australia Survey (HILDA).
- Annual waves 2001-2016, Australian national representative sample
- Data on income, wages and labor supply (each year)
- Data on wealth in particular years
- First wave administered to 19,914 people

Structural estimation sample:

- Single and married men between age 19 and 89
  - 10,133 individuals, unbalanced panel of 81,197 observations
  - Born 1916 - 1997

HILDA Data

Description of the estimation sample

	College graduates	High school graduates	High school dropouts	All
Number of individuals	2,391	5,254	2,488	10,133
Number of data points	20,207	41,965	19,025	81,197
Average num. of obs. per individual	8.45	7.99	7.65	8.01
Num. of obs. age 16-30	2,828	7,458	2,110	12,396
Num. of obs. age 31-40	5,135	9,042	2,555	16,732
Num. of obs. age 41-50	4,957	9,405	3,633	17,995
Num. of obs. age 51-60	3,873	7,405	3,532	14,810
Num. of obs. age 61-70	2,135	4,735	3,127	9,997
Num. of obs. age 71-89	1,279	3,920	4,068	9,267
Non-missing wage	75.03%	65.77%	45.71%	63.38%
Non-missing wealth	22.68%	22.59%	22.14%	22.51%
Non-missing super	20.01%	17.66%	12.94%	17.14%

Putting Institutional Settings in the Model

Approximate Pension, Super and Tax Rules

- We approximate the rules as functions of variables in our model
- We fit the approximate rules using the HILDA data

Age Pension Benefit Rule, 2001-2016

- We use the same equation we presented in the Intro:

$$\text{benefit}_{\text{max}} = \underset{(183.96)}{10,759.73} + \underset{(173.52)}{1,846.92}(\text{when year} \geq 2010),$$

$$\text{pension} = \max\left\{\text{benefit}_{\text{max}} - \max\left[\max\{\underset{(0.020)}{0.27794}\,\text{income}, \underset{(0.0004)}{0.00499}(\text{wealth} - \underset{(46,895.27)}{117,082.60})\}, 0\right]\right\}$$

Putting Institutional Settings in the Model

Superannuation:

- Superannuation is a function of earnings throughout one's career
  $\rightarrow$ Human capital at age 65 is a good proxy for lifetime earnings
  $\rightarrow$ Both depend on skill endowment and lifetime hours

- We still need to simplify the rules:

- Disregard the details of retirement income products (e.g. annuities)
  $\rightarrow$ Assume super is paid as lump sum at age 65

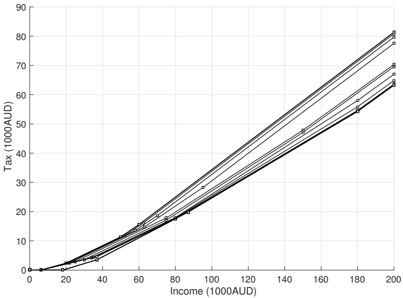
$$super_t = \rho_0 + \rho_1(\tau_{edu}) \cdot K_t, t = 65$$

- Not an unrealistic assumption:
  - Market for annuities is very thin,
  - Most people take lump sum payout.

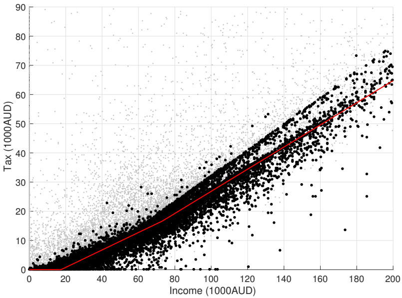
Putting Institutional Settings in the Model

Income Tax Rule, 2001-2016

Tax rules



Tax data



$$\text{tax} = \begin{cases} 0, & \text{if income} < \text{thld}_1 = \underset{(0.10016)}{17.39184}, \\ \underset{(0.00096)}{0.29907} \cdot (\text{income} - \text{thld}_1), & \text{if } \text{thld}_1 \leq \text{income} < \text{thld}_2, \\ \underset{(0.00556)}{0.37930} \cdot (\text{income} - \text{thld}_2) + \underset{(0.00096)}{0.29907} \cdot \text{thld}_1, & \text{if income} \geq \text{thld}_2 = \underset{(0.59292)}{73.17661}, \end{cases}$$

HILDA Data

Mapping Observed Hours into 6 Discrete Levels

K-medians cluster analysis | Correspondence to HILDA

h_t	Nobs	annual	week	Empl FT	Empl PT	Unemp	OLF
0	26,411	0	0	353	1,877	2,216	21,960
1	6,711	1200	24	1,303	5,408	0	0
2	23,387	2000	40	23,212	175	0	0
3	7,622	2250	45	7,622	0	0	0
4	12,115	2500	50	12,115	0	0	0
5	8,368	3000	60	8,368	0	0	0

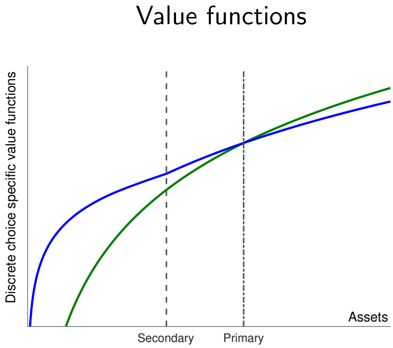
Model solution: DC-EGM

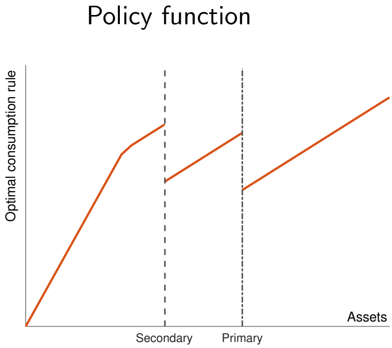
- Iskhakov, Jørgensen, Rust, Schjerning (2017) *QE*
  The Endogenous Grid Method For Discrete-Continuous Dynamic Choice Models With (Or Without) Taste Shocks
- Carroll (2006) *Economics Letters*
  The method of endogenous gridpoints for solving dynamic stochastic optimization problems.

Main idea of the endogenous grids

- Instead of searching for optimal decision in each point of the state space (traditional approaches)
- Look for the state variable (level of assets) where arbitrary chosen decision (consumption $\rightarrow$ savings) would be optimal (EGM)

Kinks and discontinuities with discrete-continuous choice





Primary kinks

1. The choice-specific (work/not work) value functions intersect (due to trade-off between income and disutility of work)

   $\Downarrow$

2. The upper envelope of the value functions has a kink (this is what we call a primary kink)

   $\Downarrow$

3. Discrete choice policy is to work on the left of the kink, and to retire on the right of the kink

   $\Downarrow$

4. Working and retiring have different corresponding optimal consumption policies

   $\Downarrow$

5. Combined consumption policy has a discontinuity

Secondary kinks

1. Value function in $t+1$ has a primary kink (because d-specific value functions intersect in $t+1$)

   $\Downarrow$

2. In the non-concave region around a primary kink in $t+1$ the maximand in the Bellman equation has multiple local optima

   $\Downarrow$

3. The Euler equation for the corresponding values of wealth has multiple solutions, all solutions are found in EGM

   $\Downarrow$

4. "Suboptimal" endogenous points have to be dropped: find the point where global maximum shifts from one solution to the other

   $\Downarrow$

5. Optimal consumption rule in period t has a discontinuity, the value function has a corresponding secondary kink

Adding extreme value shocks

Properties of the full solution

1. Value functions are non-concave and have kinks
2. Consumption functions have discontinuities
3. Discontinuities/kinks propagate through time and accumulate

Extreme value distributed taste shocks

- Smooth out primary kinks
- Extreme value distribution $\rightarrow$ closed form expectations for choice probabilities and expectation of the max (logsum)
- Two interchangeable interpretations
    - Structural: unobserved state variables
    - Logit smoothing: to streamline the solution
- Prevent propagation of kinks and discontinuities
- No complete smoothing in general: secondary kinks may persist

Estimation: Method of Simulated Moments

McFadden (1989) *Econometrica*
A method of simulated moments for estimation of discrete response models without numerical integration

- Method of simulated moment estimator
- Diagonal weighting matrix
- Logit smoothed simulator for better numerical performance
- POUNDerS derivative free trust region minimization algorithm
- "Practical optimization using no derivatives for sums of squares"
- POUNDerS is tailored to problems of minimizing a (weighted) sum of squared residuals. It works by approximating each moment condition as a quadratic function of the model parameters. It is derivative free, which is useful because computing derivatives is computationally expensive. But it relies on existence of derivatives for the quality of the approximation. The algorithm terminates if it succeeds in setting the (approximate) derivatives of the moments with respect to the parameters to zero (within tolerance).

Choice of moments to match

	High school		Dropouts		College	
Moments	Ages	N	Ages	N	Ages	N
Work status by age	19 - 86	67	19 - 88	70	23 - 89	64
hours when working	19 - 70	48	19 - 70	48	23 - 70	44
wage when working	19 - 70	48	19 - 70	48	23 - 70	44
variance of wage	19 - 70	10	19 - 70	10	23 - 70	10
skewness of earnings	19 - 85	13	19 - 85	13	23 - 85	13
hours = 20	19 - 86	67	19 - 86	68	23 - 89	64
hours = 40	19 - 82	61	19 - 84	64	23 - 79	57
hours = 45	19 - 77	55	19 - 83	56	23 - 76	51
hours = 50	19 - 76	58	19 - 88	66	23 - 77	53
wealth	25 - 85	13	25 - 85	13	25 - 85	13
work to work	19 - 70	48	19 - 70	48	23 - 70	44
nowork to nowork	19 - 70	48	19 - 70	48	23 - 70	44
super	65	1	65	1	65	1
Total		537		553		502

Estimates of the preference parameters

Parameter	Description	Estimate	Std.Err.
ζ	CRRA coefficient in consumption	0.79488	0.07327
γ_1	Disutility of working 1000 hours (20 per week)	1.4139	0.38508
γ_2	Disutility of working 2000 hours (40 per week)	2.0088	0.59712
γ_3	Disutility of working 2250 hours (45 per week)	2.9213	0.78915
γ_4	Disutility of working 2500 hours (50 per week)	2.8639	0.80946
γ_5	Disutility of working 3000 hours (60 per week)	3.8775	1.05032
κ_1	Correction coefficient for low type with disutility of work	0.50321	0.17973
κ_2	Quadratic coefficient on age for older workers	0.00008	0.00004
κ_3	Linear coefficient on age for young workers	0.05083	0.01554
ξ	CRRA coefficient in utility of bequest	0.48834	0.34766
b_{scale}	Scale multiplicator of the utility of bequest	0.68659	1.42044
β_{cg}	Discount factor, college	0.96963	0.00238
β_{hs}	Discount factor, highschool	0.96732	0.00189
β_{dr}	Discount factor, dropouts	0.96806	0.00138
λ	Scale of EV taste shocks	0.29950	0.08825

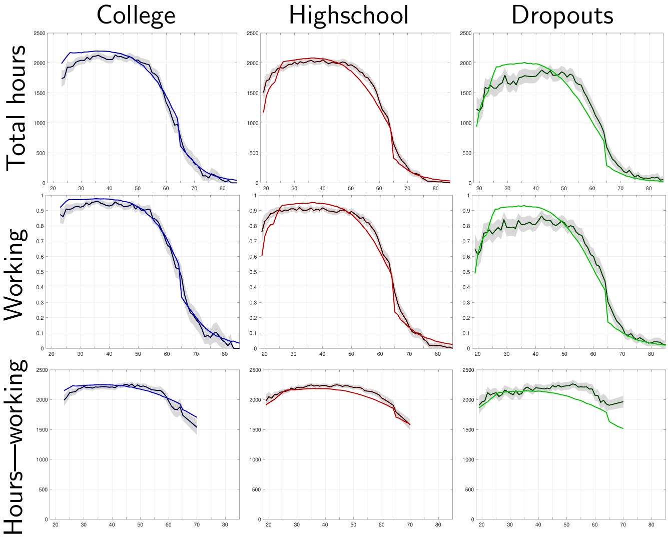
Human capital accumulation process

Parameter	Description	Estimate	Std.Err.
$\eta_{0,cg}$	Constant for college	2.78766	0.41169
$\eta_{0,hs}$	Constant for high school	2.56761	0.36634
$\eta_{0,dr}$	Constant for dropouts	2.45647	0.33269
$\eta_{0,high}$	Constant for high type	0.39311	0.41893
$\eta_{1,cg}$	Work experience for college	0.03041	0.00796
$\eta_{1,hs}$	Work experience for high school	0.02164	0.00768
$\eta_{1,dr}$	Work experience for dropout	0.01974	0.00682
$\eta_{2,cg}$	Work experience square for college	−0.00017	0.00021
$\eta_{2,hs}$	Work experience square for high school	−0.00002	0.00018
$\eta_{2,dr}$	Work experience square for dropout	0.00000	0.00010
η_3	Age (time index)	0.02676	0.00280
η_4	Age (time index) square	−0.00076	0.00004

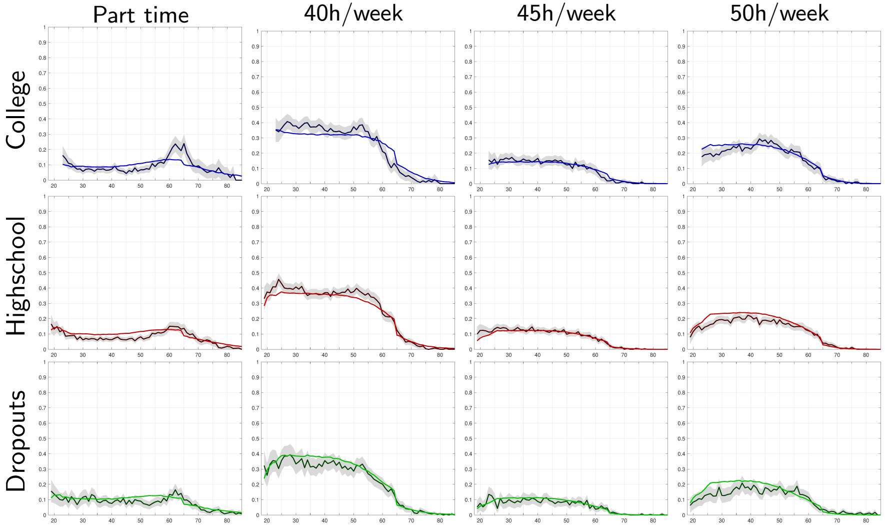
Estimates of other structural parameters

Parameter	Description	Estimate	Std.Err.
ς_0	St.dev. in shock distribution: constant	0.24485	0.24055
ς_1	St.dev. in shock distribution: age	0.00421	0.00935
tr	Transfer from parents	5.51308	1.43804
ρ_{cg}	Superannuation: human capital — college	6.30347	2.58472
ρ_{hs}	Superannuation: human capital — high school	5.43473	3.30737
ρ_{dr}	Superannuation: human capital — dropouts	6.47838	3.95647
ς_{w_0}	Initial wealth sigma	1.48960	6.69399
p_{cg}	High type proportion — college	0.90089	0.04952
p_{hs}	High type proportion — high school	0.80130	0.04366
p_{dr}	High type proportion — dropout	0.69306	0.04411

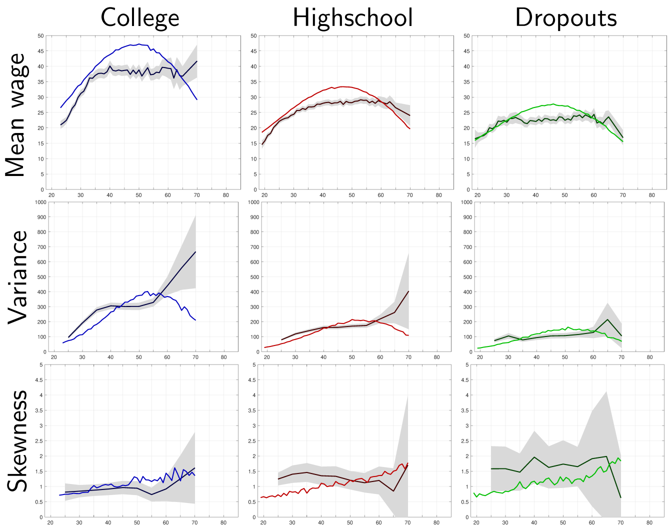
Goodness of fit: total hours and participation



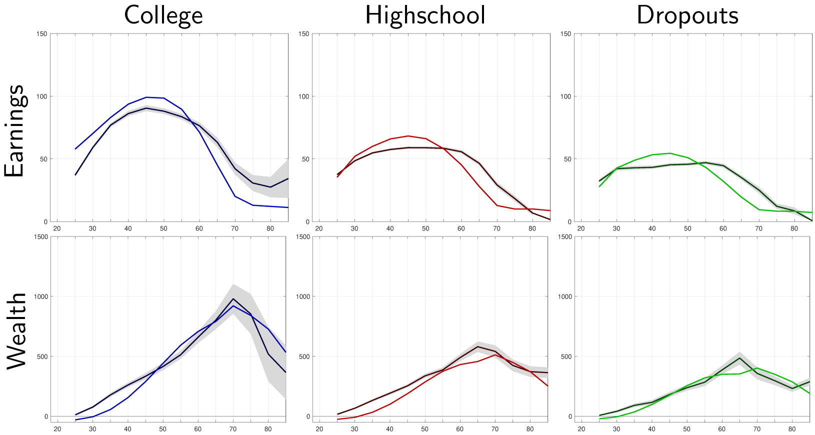
Goodness of fit: discrete level of hours



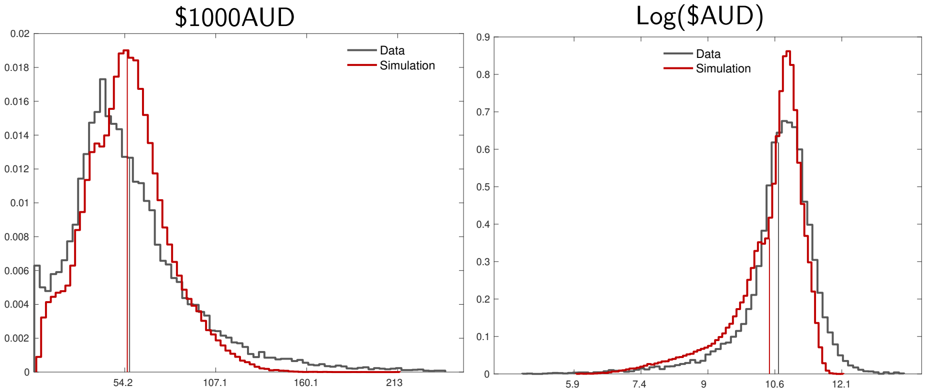
Goodness of fit: lifecycle wage distribution



Goodness of fit: earnings and wealth



Goodness of fit: overall income distribution



Policy simulations

Policies to be simulated:

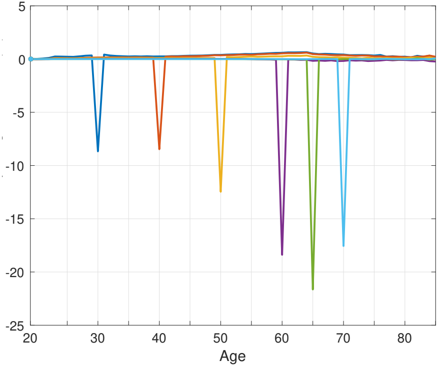
1. Transitory wage/tax changes
2. Permanent wage/tax changes
   - Anticipated and Unanticipated
3. Eliminate Age Pension
4. Improved Targeting of Age Pension
   - Change income and asset taper rates

We emphasize how labor supply elasticities and policy impacts differ by age and education

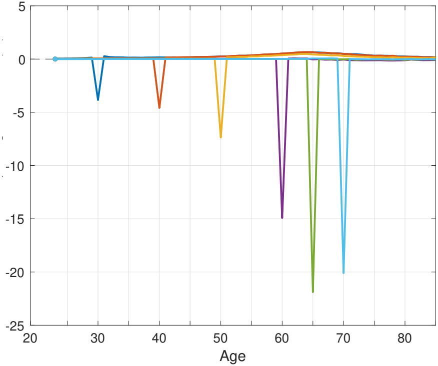
Frisch Elasticities

Transitory 10% wage decrease (tax increase) $\rightsquigarrow$ % change in hours

High school dropouts



College graduates



(Anticipated effects)

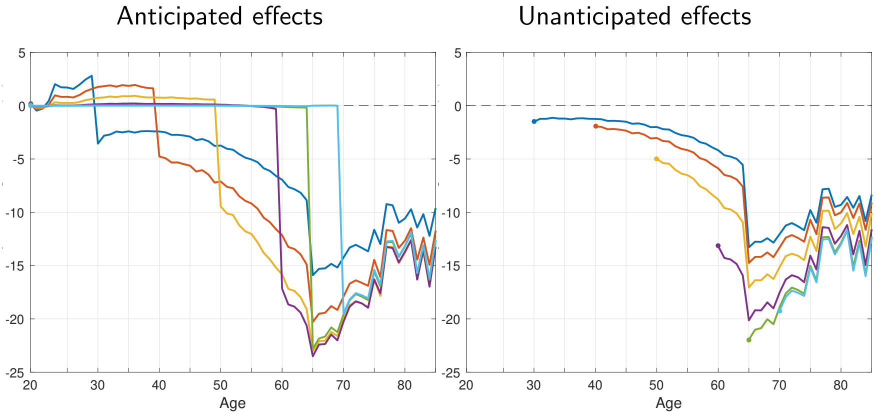
Frisch elasticities

- Frisch elasticities increase with age:
- For college grads it is 0.4 at age 30 and 2.2 at age 65.
- The increase is greater for the more educated
- Consistent with Imai-Keane (2004) and Keane-Wasi (2016)

Permanent 10% wage decrease ⇝ % change in hours

High school graduates

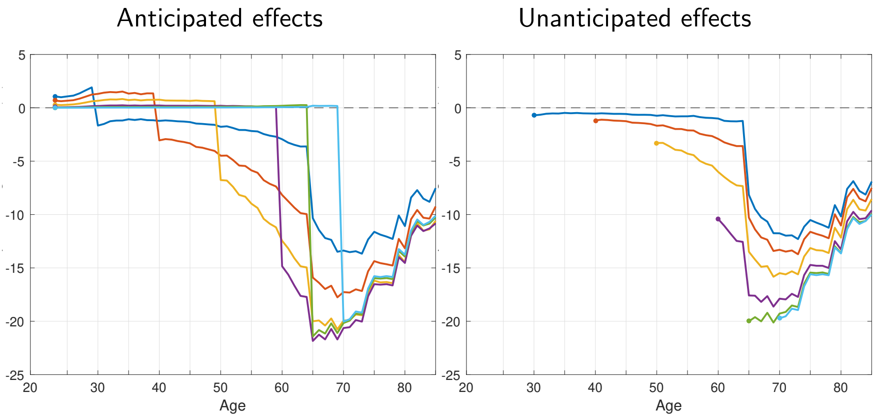
No compensation ⟶ Marshall effects



Permanent 10% wage decrease ⇝ % change in hours

College graduates

No compensation ⟶ Marshall effects



Permanent 10% wage decrease $\rightsquigarrow$ hours

- Larger decline in hours if policy is anticipated: labor supply is shifted towards the beginning of life cycle where wage is not yet decreased

- Effect is very different at different points of the life cycle
- Much larger hours decline if wage decrease occurs at older ages

- Elasticities smaller for college grads than HS grads at younger ages
- But catch up at older ages

- Key Point: Effect of HC on labor supply elasticities not changed by hours bunching

Effects of the age pension

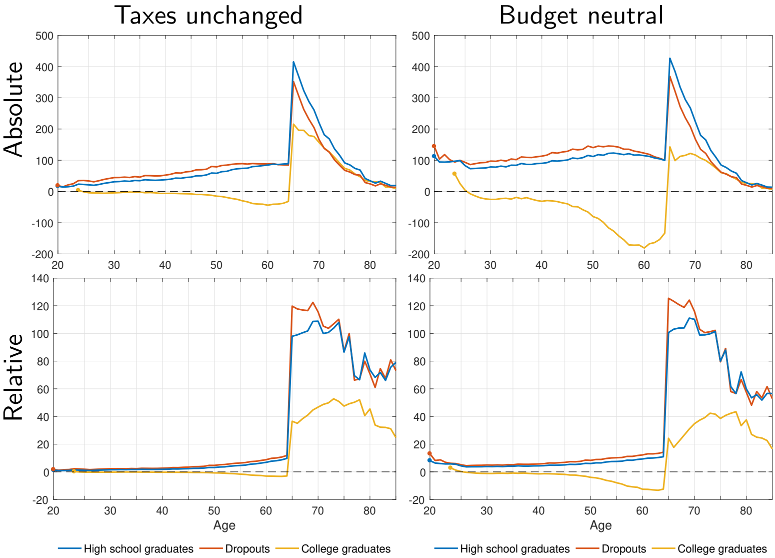
Simulate the world without Age Pension

1. Cost of program is 1/3 of income tax revenue
2. Elimination allows 33% tax cut (if no behavioral response)

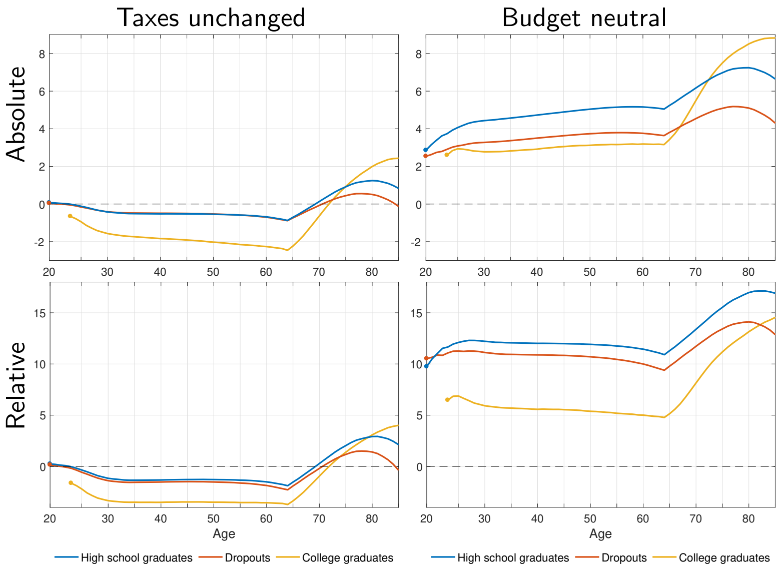
Unchanged taxes vs. Revenue neutral

- Elimination of Age Pension generates 5.8% increase in labor supply
- This allows a 37% cut in income tax rates in budget neutral simulation

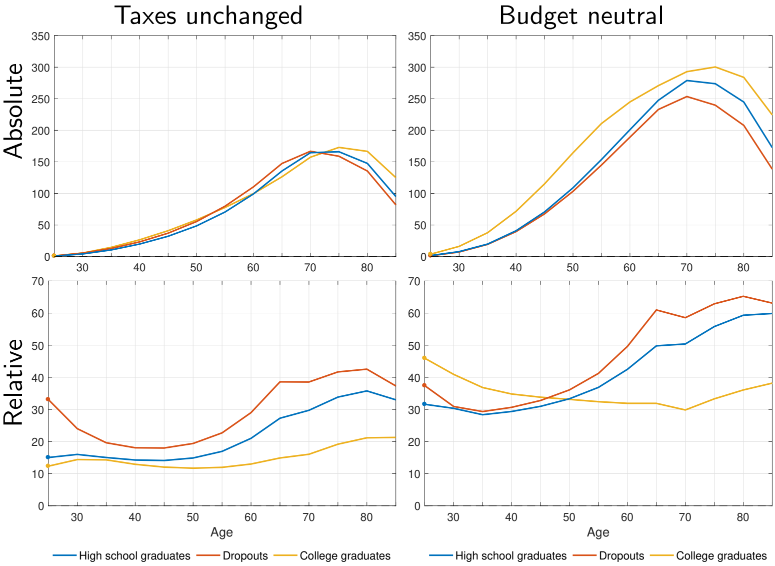
Elimination of Age Pension ⇝ hours



Elimination of Age Pension ⇝ consumption, $1000 AUD



Elimination of Age Pension $\rightsquigarrow$ wealth, \$1000 AUD



Effects of the age pension

The world without the age pension (revenue neutral):

- Labor supply increases for dropouts and HS graduates at all ages
- Labor supply decreases for college graduates pre 65 (income effects), but increases greatly at 65+
- Tax rates fall by 37% in budget neutral simulation
- About 90% of workers prefer to live in a world with no age pension and lower taxes
- Only 10% of low skill type individuals experience decrease in welfare
- This result reflects the poor targeting of the Age Pension program and large labor supply distortion it creates

Better Targeting the Age Pension

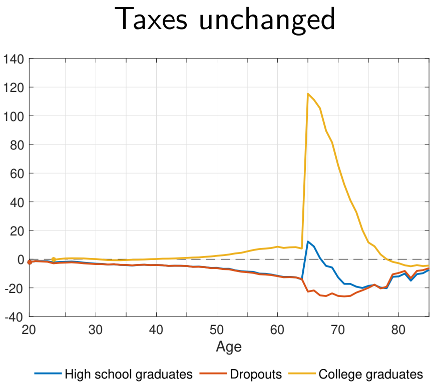
The Change in Program Rules that we simulate is:

- Double the Income and Asset taper rates.
- Specifically:
  - Double effective income taper rate from 27.7% to 55.5%
  - Double effective asset taper rate from 1/2 cent on the dollar to one cent on the dollar

- In budget neutral simulation we can cut income tax rates by 5.9% - i.e., top rate reduced from 37.9% to 35.7%

Better Targeting the Age Pension

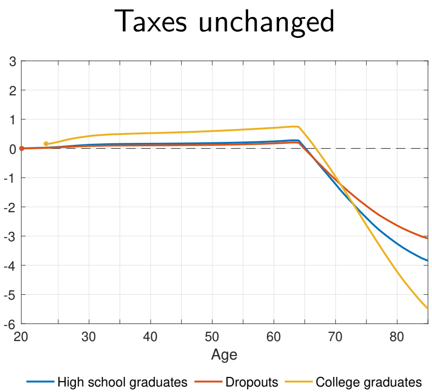
Doubling of income/asset tapers ⇝ Effects on Hours of Work

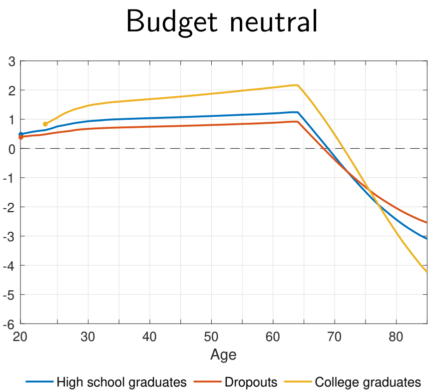


(Note: Change in annual hours)

Better Targeting the Age Pension

Doubling of income/asset tapers ⇝ Effects on Consumption

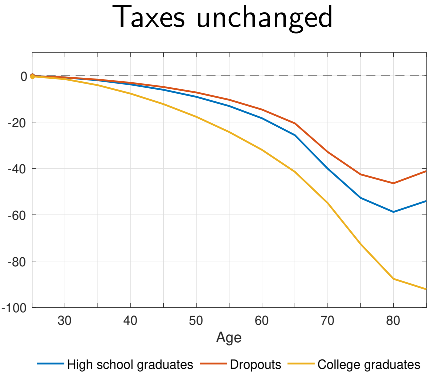


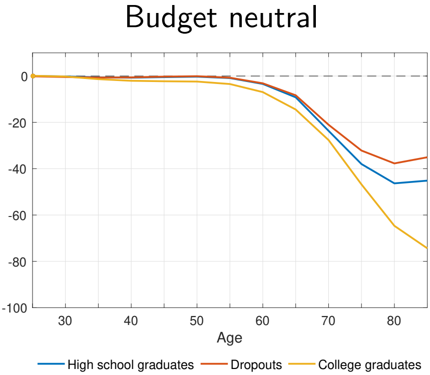


(Note: Change in $1000 AUD)

Better Targeting the Age Pension

Doubling of income/asset tapers ⇝ Effects on Assets

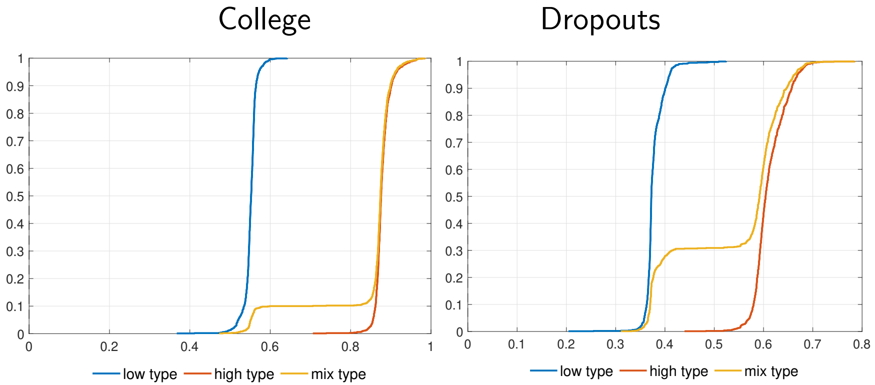




(Note: Change in $1000 AUD)

Better Targeting the Age Pension

Doubling Tapers + Tax Cut $\rightsquigarrow$ Effects on Utility



(Note: Change in ex-ante utility)

Better Targeting the Age Pension

Double Taper Rates + Tax cut $\rightsquigarrow$ Results:

- At age 65+ labor supply of college grads increases by 20% while that of dropouts falls by 8%
- College grads rely on age pension less while dropouts rely on it more - better targeting
- In budget neutral simulation we cut income tax rates by 5.9% - i.e., top rate reduced from 37.9% to 35.7%
  - This causes small increase in labor supply prior to age 65
- All types better off - CEVs are \$1.4k, \$1.5k, \$1.7k for dropouts, HS, college types, respectively

Results and conclusions

Labor supply

- Large variation of labor supply elasticities by age and education:
  - Labor supply elasticities increase with Age
  - Elasticities are smaller for higher education groups

Age Pension

- The program has large negative labor supply effects
- The program is expensive (Largest welfare item in budget)
- It is poorly targeted $\Rightarrow$ Very low effective taper rates
- Doubling of Taper Rates combined with 5.9% tax cut would be Pareto improvement