CompEcon Week12

Effects of taxes and safety net pensions on life-cycle labor supply, savings and human capital: The case of Australia

Fedor Iskhakov Michael Keane

ANU

October, 2018

Applied structural microeconometrics project

Main research question:

- Labor supply responses to Age Pension and taxes
- Over the whole life cycle (long run effects)
- With anticipated and unanticipated policy changes

Applied structural microeconometrics project

Main research question:

- Labor supply responses to Age Pension and taxes
- Over the whole life cycle (long run effects)
- With anticipated and unanticipated policy changes

Counterfactual scenarios of alternative policy settings:

- Numerical model of individual behavior
  which includes the ..
- .. parameterized retirement and tax policy.
- Study this model under alternative policy parameters.
- Eliminate various elements to assess their induced effects.

Structural approach

Available data:

- Historical data on modeled behavior
- Observations under particular (possibly time variant) policy
- Counterfactual scenarios generally not observed
- HILDA survey panel data: social and economic measures, individual and household level

Structural approach

Available data:

- Historical data on modeled behavior
- Observations under particular (possibly time variant) policy
- Counterfactual scenarios generally not observed
- HILDA survey panel data: social and economic measures, individual and household level

Structural approach:

1. Develop a theoretical model of behavior
   - Structural parameters $\leftrightarrow$ policy invariant (by reasonable assumption)
   - Policy parameters to facilitate counterfactual scenarios
2. Estimate the model using historical data
3. Simulate counterfactual scenarios of alternative policy settings, both anticipated and unanticipated changes

Solution and estimation

Nested structure:

- Estimation — optimization over parameter vector
- Need to solve the model for each value of parameter vector

1. For given values of parameters,
   solve the theoretical model numerically
2. Compute/simulate the entities entering into the estimator
3. Evaluate the criterion function using the data
4. Find admissible direction and step size of the parameter vector for the next iteration
5. Repeat until convergence

Computationally intensive!

Road map

1. Theoretical model
   - Labor supply, consumption and savings
   - Human capital accumulation
   - Parameterized pension and tax policy
2. Data and *first stage* estimation
   - Some relationships can be estimated outside the model
3. Computational method for solving the model
   - Simultaneous continuous-discrete choice..
4. Estimation method
   - Method of Simulated Moments
5. Counterfactual simulations

Why: story about labor supply elasticities

Precise measure of responsiveness of labor to after-tax wages is crucial for effective design of the tax and transfers system.

- High responsiveness $\rightarrow$ taxes are distortionary
- Low responsiveness $\rightarrow$ optimal tax is higher

Reduced form estimation from micro-data $\Rightarrow$ Small elasticities

- MaCurdy (1981, 1983) Altonji (1986) Blundell and Walker (1986) Blomquist (1985) Angrist (1991) Pistaferri (2003) Ziliak and Kniesner (2005) Aaronson and French (2009)

Calibration of macro models $\Rightarrow$ Large elasticities are essential

Reconciliation

Keane, Rogerson 2011 *NBER*
Reconciling Micro and Macro Labor Supply Elasticities:
A Structural Perspective

Sources of the discrepancy:

1. Human capital
2. Borrowing constraints
3. Uninsurable wage risk
4. Optimization frictions
5. Extensive margin (participation)
6. Population heterogeneity and aggregation

Similar issues are important for modeling the effects of pension system

Main model design features

- Labor supply elasticity ⇒ Human capital accumulation
Borrowing constraint
Wage risk and wealth process
- Intensive and extensive margins ⇒ Endogenous retirement without absorption
- Frictions on the labor market ⇒ Discreteness of labor supply
- Heterogeneity in effects ⇒ Observed (education),
Unobserved (types) heterogeneity

Summary of the stochastic life cycle model

1. Discrete time = age from 19 to 100 (stochastic survival)
2. Annual decisions on
   - Consumption (continuous choice)
   - Hours from [0, 24, 40, 45, 50, 60] per week (discrete choice)
3. Stochastic elements in the model
   - Survival (longevity risk)
   - Idiosyncratic wage shock
4. Human capital accumulation
   - Learning-by-doing $\rightarrow$ Accumulating work experience
   - Human capital increases future wage $\rightarrow$ part of compensation
5. Observed and unobserved heterogeneity in the population
   - Education $\rightarrow$ Initial endowment and human capital technology
   - Unobserved types $\rightarrow$ Initial endowment and preference for leisure

Hours, human capital and wage + wealth and consumption

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 $wage_t = K_{t-1} \cdot R_t \cdot \epsilon_t^{wage}$,
$R_t = 1$ is rental rate of human capital,
$\epsilon_t^{wage} \sim lnN(0, \sigma_t^{wage})$

Consumable wealth in the beginning of the period $M_t < M_{max}$

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

Intertemporal budget
$M_{t+1} = (M_t - c_t)(1 + r) + h_t \cdot wage_{t+1} + \text{transfers}_{t+1}$

What should be the state variables?

Definition: Formal account of the information that is necessary and sufficient for making the decisions we model

Art of structural modeling: Choose variables such that interesting and realistic behavior is replicated with tractable set of variables

Tricks:

- Sufficient statistics
- Principle components
- Efficient encoding of information
- Simplifying behavioral assumptions

State variables

1. Consumable wealth M_t
2. Human capital $K_t = f\left(\sum_{\tau=1}^{t-1} h_\tau, \text{age}, \text{education}, \text{type}\right)$
3. Education
4. Unobserved type

$X_t = \left(M_t, K_t, \text{education}, \text{type}\right)$ is not perfect

$0 \leq M_t \leq M_{max}$

$0 \leq K_t \leq \sum_{\tau=1}^{t-1} \max\{h_\tau : h_\tau \in H\} \leq (t-1)h_{max} \leq (T-1)h_{max}$

State variables (more convenient way)

1. Consumable wealth M_t
2. Human capital $K_t = f\left(\sum_{\tau=1}^{t-1} h_\tau, \text{age}, \text{education}, \text{type}\right)$
3. Education
4. Unobserved type

$X_t = \left(M_t, \mathcal{E}_t, \text{education}, \text{type}\right)$, where $\mathcal{E}_t$ is fraction of total working time to total time budget, i.e. work experience

$0 \leq \mathcal{E}_t \leq 1$

$\mathcal{E}_t = \frac{1}{t \cdot h_{max}} \sum_{\tau=1}^{t-1} h_\tau \;\Leftrightarrow\; \mathcal{E}_{t+1} = \frac{1}{t+1}\left(\mathcal{E}_t t + \frac{h_t}{h_{max}}\right), \mathcal{E}_0 = 0$

$K_t = f\left(\mathcal{E}_t \cdot t \cdot h_{max}, \text{age}, \text{education}, \text{type}\right)$

Preferences

Utility of consumption

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

Utility of (accidental) bequests

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

$B_t = M_t - c_t$ bequeathed wealth

$b_{scale} > 0$, $\zeta > 0$, $\xi > 0$ parameters to be estimated

a_0 credit constraint (maximum amount of borrowing)

Disutility of work

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

Type: high ($\kappa_{type} = 1$) and low ($\kappa_{type} = 1 + \kappa_1$)

Simple age effects:

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

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

Correction for the young ages

Labor supply in ages 19 to 25 was persistently overpredicted

1. Scrutinize the definition of schooling and graduation time
   Make college graduates start their career later than everybody else
2. Transfer from parents
3. Additional disutility of work at young age

$$M_{t+1} = (M_t - c_t)(1+r) + h_t \cdot wage_{t+1} + \text{transfers}_{t+1}$$

$$\downarrow$$

$$M_{t+1} = (M_t - c_t)(1+r) + h_t \cdot wage_{t+1} + \text{transfers}_{t+1} + tr_{t+1} \cdot \mathbb{1}\{t+1 \leq 22\}$$

Bellman equation

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

Note: c_t continuous, h_t discrete

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

$H_t(\tau)$ type-specific 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
- Family dynamics, income and labor supply (each year)
- Data on wealth, health and health insurance, retirement, fertility, literacy and numeracy (particular years, reoccurring)
- Approximately 20,000 households in total

Structural estimation sample:

- Single and married men between age 19 and 85
  - 10,133 individuals, unbalanced panel of 81,197 observations
  - Individuals born 1912 - 1997

Australian institutional settings

1. Old Age Pension (safety net)
   - Universal from age 65
   - Not dependent of working history
   - Financed from general revenue
   - Subject to means testing
2. Superannuation (compulsory savings)
   - Defined contribution system, accumulation subject to market risk
   - Individual accounts in private super funds
   - Employers are compelled to contribute a fraction of wage
   - Accessible from age 65

Incorporating institutional settings in the model

Age Pension

$$\text{Pension} = \max\left\{0, \text{full benefit} - \max\{\text{income test}, \text{asset test}\}\right\}$$

$$\begin{aligned} \text{Income test} &= \max\left\{0, \text{income taper rate} \cdot (\text{income} - \text{income threshold})\}\right\} \\ \text{Asset test} &= \max\left\{0, \text{asset taper rate} \cdot (\text{wealth} - \text{asset threshold})\}\right\} \end{aligned}$$

- Need to represent within the state space (with minimal additions)
  - Use simplified institutional rules and formulas?
  - Use approximation obtained from observed data? ◀

Means testing and 2009 reform

- Increase in maximum pension benefit over time
- Big jump in 2009

Means testing and 2009 reform

1. Increase in maximum pension benefit
2. Changes in taper rate (second threshold)

Age pension equation

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

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

Superannuation

- Superannuation is a function of the labor supply throughout career $\rightarrow$ function of accumulated human capital
- We need to simplify the rules:
- Disregard the details of retirement income $\rightarrow$ paid as lump sum at age 65

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

- Not a completely unrealistic assumption: market for annuities is extremely thin

Income tax function

Tax rules

Tax data

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

Intertemporal budget constraint (updated)

$$M_{t+1} = (M_t - c_t)(1+r) + h_t \cdot wage_{t+1} + \text{transfers}_{t+1}$$

$$\downarrow$$

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

$tr_{t+1} \cdot \mathbb{1}\{t+1 \leq 22\}$ transfers from parents to youth

Survival process

$$\delta_t = \begin{cases} 1 & \text{if } age_t < 40 \\ 1 - \underset{(0.0000249)}{0.0006569} \left[\exp\left(\underset{(0.0008145)}{0.1078507} (age_t - 40) \right) - 1 \right] & \text{if } age_t \geq 40 \end{cases}$$

Education levels

Original HILDA classification			Coarsened 3 level classification		
	N obs	%		N obs	%
Postgrad - masters or doctorate	557	5.50	College	2,391	23.60
Grad diploma, grad certificate	503	4.96			
Bachelor or honours	1,331	13.14			
Advanced diploma, diploma	922	9.10	High school	5,254	51.85
Certificate III or IV	3,178	31.36			
Certificate I or II	0	0.0			
Certificate not defined	0	0.0			
Year 12	1,154	11.39			
Year 11 and below	2,488	24.55	Dropouts	2,488	24.55
Undetermined	0	0.0			
Total	10,133	100.00		10,133	100.00

Choice of hours levels

K-medians cluster analysis

h_t	Nobs	med(annual)	med(weekly)	min	max
0	18,168	0	0	0	500.00
1	4,484	1200	24	500.03	1600.00
2	15,930	2000	40	1600.02	2124.90
3	5,466	2250	45	2133.36	2368.03
4	8,735	2500	50	2375.10	2750.00
5	6,259	3000	60	2750.10	4200.00

Choice of discrete levels of hours

K-medians cluster analysis

h_t	Nobs	annual	week
0	26,411	0	0
1	6,711	1200	24
2	23,387	2000	40
3	7,622	2250	45
4	12,115	2500	50
5	8,368	3000	60

Correspondence to HILDA

Empl FT	Empl PT	Unemp	OLF
353	1,877	2,216	21,960
1,303	5,408	0	0
23,212	175	0	0
7,622	0	0	0
12,115	0	0	0
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

- Endogenous gridpoint method for discrete-continuous choice problems (DC-EGM)
- Generalization of EGM algorithm to discrete-continuous choices
- Iterates on the Euler equation $\rightarrow$ avoids root-finding operations $\rightarrow$ fast and accurate solution method
- Discrete choice $\rightarrow$ Non-convex problem $\rightarrow$ First order conditions not sufficient $\rightarrow$ DC-EGM finds all extrema and selects global maxima
- Assume EV taste shocks associated with discrete choices $\rightarrow$ Choice probabilities $\rightarrow$ Smooth the problem to some extent

Endogenous gridpoint method (EGM)

Carroll (2006) *Economics Letters*
The method of endogenous gridpoints for solving dynamic stochastic optimization problems.

Idea

- 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)

1. A1. Strictly concave monotone and differentiable utility function + analytical inverse of the marginal utility
2. One continuous state variable (*wealth*) and one continuous choice (*consumption*)
3. Particular structure of the motion rule (*budget*)

Model of consumption/savings only

$$V_t(M_t) = \max_{0 \leq c \leq M_t} \left[u(c) + \beta E V_{t+1} \left(R(M_t - c) \right) \right]$$

M_t	cash-in-hand, all resources available at period t
$A_t = M_t - c_t$	assets at the end of period t (savings)
R	*deterministic* return on savings
$u(c)$	utility of current consumption

Euler equation (F.O.C.)

$$u'\left(c_t^\star(M_t)\right) = \beta E \left[R \cdot u' \left[c_{t+1}^\star \left(R(M_t - c_t^\star(M_t)) \right) \right] \right]$$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$
2. Intertemporal budget constraint: $A \to M_{t+1}$
   $M_{t+1} = R(M_t - c_t) = R \cdot A$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$
2. Intertemporal budget constraint: $A \to M_{t+1}$
   $M_{t+1} = R(M_t - c_t) = R \cdot A$
3. Policy function at period $t+1$: $M_{t+1} \to c_{t+1}$
   $c_{t+1} = c_{t+1}^\star(M_{t+1})$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$
2. Intertemporal budget constraint: $A \to M_{t+1}$
   $M_{t+1} = R(M_t - c_t) = R \cdot A$
3. Policy function at period $t+1$: $M_{t+1} \to c_{t+1}$
   $c_{t+1} = c_{t+1}^\star (M_{t+1})$
4. Inverted Euler equation: $c_{t+1} \to c_t$
   $c_t = (u')^{-1} \left(\beta E \left[R \cdot u' \left(c_{t+1}^\star (M_{t+1})\right) | A\right]\right)$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$
2. Intertemporal budget constraint: $A \to M_{t+1}$
   $M_{t+1} = R(M_t - c_t) = R \cdot A$
3. Policy function at period $t+1$: $M_{t+1} \to c_{t+1}$
   $c_{t+1} = c_{t+1}^\star (M_{t+1})$
4. Inverted Euler equation: $c_{t+1} \to c_t$
   $c_t = (u')^{-1} \left(\beta E \left[R \cdot u' \left(c_{t+1}^\star (M_{t+1})\right) | A\right]\right)$
5. Intratemporal budget constraint: $c_t + A \to M_t$
   $M_t = c_t + A \to c_t^\star (M_t)$

EGM algorithm

Fix a time invariant grid over current period savings A_t
Start with $c_T^\star = M_T$. In each period $t = T, T-1, .., 1$:

EGM step

1. For each point A = current period savings $(= M_t - c_t)$
2. Intertemporal budget constraint: $A \to M_{t+1}$
   $M_{t+1} = R(M_t - c_t) = R \cdot A$
3. Policy function at period $t+1$: $M_{t+1} \to c_{t+1}$
   $c_{t+1} = c_{t+1}^\star (M_{t+1})$
4. Inverted Euler equation: $c_{t+1} \to c_t$
   $c_t = (u')^{-1} \left(\beta E \left[R \cdot u' \left(c_{t+1}^\star (M_{t+1})\right) | A\right]\right)$
5. Intratemporal budget constraint: $c_t + A \to M_t$
   $M_t = c_t + A \to c_t^\star (M_t)$

Obtain endogenous points $\{M_t\}$ and corresponding policy $\{c_t^\star (M_t)\}$

Accuracy and speed of EGM

	Traditional Euler	EGM
Running time	37 sec.	0.11 sec.
Max abs error, $c_t^\star$	5e-9	4e-14
Mean abs error, $c_t^\star$	1.4e-12	1.5e-14
Max abs error, $V_t(M, \mathbb{R})$	39.466	15.163
Mean abs error, $V_t(M, \mathbb{R})$	2.5e-02	3.2e-02

5000 grid points, 50 time periods
Comparison to the analytic solution

How to approach discrete/continuos choice

The goal: avoid rootfinding

DC-EGM ver. 1.0

1. EGM step for each discrete choice d and every state st
2. Compute d-specific value functions and consumption rules
3. Compare the d-specific value functions to find optimal switching points (compute upper envelope)
4. Reconstruct overall consumption rule and value function from optimal switching points

Clausen & Strub, 2010-2013
A General and Intuitive Envelope Theorem.

Show that Euler equation remains a necessary condition for the optimal continuous consumption.

Period T : choice specific value functions

Period $T-1$: Choice specific VF

Period $T-1$: Optimal consumption

Primary kinks

1. d-specific 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

Period $T-2$: Choice specific VF

Period $T-2$: Choice specific VF

Multiple zeros of Euler residuals

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

Period $T-2$: Secondary upper envelope

DC-EGM updated algorithm

DC-EGM ver. 2.0

1. Start from terminal period, compute optimal consumption rule and value function. Loop backwards over time:
2. EGM step for each discrete choice d and every state st
3. Compute d-specific value functions and consumption rules
4. Compute the "secondary" upper envelope over the non-convex regions of the d-specific value functions and update the corresponding consumption rules
5. Compare the d-specific value functions to find optimal switching points (compute upper envelope)
6. Reconstruct overall consumption rule and value function from optimal switching points

Period $T-2$: VF, primary and secondary kinks

Period $T-2$: Optimal consumption

Multiplication and accumulation of kinks

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

This properties are attributes of the model itself.
Any solution method has to deal with these complexities.

Wage shocks and taste shocks

Random earnings or random returns:

- Smooth out secondary kinks only
- Expectations have to be taken over discontinuous functions

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

Before $T-3$: Optimal consumption with $\sigma_R = .1$

Smoothed Euler equation

Bellman equation with taste shocks

Chocie-specific EV i.i.d. taste shocks ϵ_h

$$V_t(X_t) = \max_{h_t \in H_t(\tau)} \left[v_t(X_t, h_t) + \lambda \epsilon_h \right]$$

$$v_t(X_t, h_t) = \max_{0 \le c_t \le M_t + a_0} \Big\{ u(c_t) - v_t(h_t, \tau_{uh}) + (1 - \delta_t) w(M_t - c_t) + \delta_t \beta(\tau_{edu}) E\Big[\text{LogSum}\big(v_{t+1}(X_{t+1}, h_{t+1}) \big) \Big| X_t, c_t, h_t \Big] \Big\}$$

$$\text{LogSum}\big(v_t(X_t, h_t)\big) = \lambda \log \Big(\sum_{h_t \in H_t(\tau)} \exp \frac{v_t(X_t, h_t)}{\lambda} \Big)$$

$$P(h|X_t) = \exp \frac{v_t(X_t, h)}{\lambda} \Big/ \sum_{h_t \in H_t(\tau)} \exp \frac{v_t(X_t, h_t)}{\lambda}$$

Labor supply becomes probabilistic with standard logit choice probability

DC-EGM with taste shocks

DC-EGM in the space of choice specific value and consumption functions

DC-EGM ver. 3.0

1. Start from terminal period, compute optimal consumption rule and value function. Loop backwards over time:
2. EGM step for each discrete choice d and every state st
3. Compute d-specific value functions and consumption rules
4. Compute the "secondary" upper envelope over the non-convex regions of the d-specific value functions and update the corresponding consumption rules
5. ~~Compare the d-specific value functions to find optimal switching points (compute upper envelope)~~
6. ~~Reconstruct overall consumption rule and value function from optimal switching points~~

Optimal consumption with taste shocks only

Optimal consumption with random returns

DC-EGM with taste shocks

- Problem re-formulated in terms of choice specific value functions (Higher memory requirement for storage)
- Calculation of *primary* upper envelope $\rightarrow$ *logsum*
- Easier computation of expectations (due to less discontinuities)

- Still no root-finding operations (apart from the terminal period)
- Still very accurate solution (zero Euler residuals at grid points)
- With wage and taste shocks the accumulation of kinks is moderated

Iskhakov, Jørgensen, Rust, Schjerning (2017) *QE*
The Endogenous Grid Method For Discrete-Continuous Dynamic Choice Models With (Or Without) Taste Shocks

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
- Non-convex values functions create problems for derivative based methods
- POUNDerS derivative free trust region minimization algorithm

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 2001-2012, Australian national representative sample
- Family dynamics, income and labor supply (each year)
- Data on wealth, health and health insurance, retirement, fertility, literacy and numeracy (particular years, reoccurring)
- Approximately 20,000 people in total

Structural estimation sample:

- Single and married men between age 19 and 85
  - 8,836 individuals, unbalanced panel of 56,090 observations
  - Individuals born 1916 - 1993

Choice of moments to match

	High school		Dropouts		College	
Moments	Ages	N	Ages	N	Ages	N
working	19 - 85	67	19 - 85	67	23 - 85	63
hours|working	19 - 75	57	19 - 78	60	23 - 71	49
wage|working	19 - 73	55	19 - 69	51	23 - 69	47
var of wage	19 - 73	55	19 - 69	51	23 - 69	47
hours20	19 - 85	67	19 - 85	67	23 - 85	63
hours40	19 - 85	67	19 - 85	67	23 - 85	63
hours45	19 - 85	67	19 - 85	67	23 - 85	63
hours50	19 - 85	67	19 - 85	67	23 - 85	63
wealth	19 - 85	55	19 - 85	55	23 - 85	49
work2work	19 - 74	56	19 - 77	59	23 - 70	48
nowork2nowork	19 - 85	67	19 - 85	67	25 - 85	61
super lumpsum	65	1	65	1	65	1
Total		681		679		617

Estimates of the preference parameters

Parameter	Description	Estimate	Std.Err.
ζ	CRRA coefficient in consumption	0.80989	0.06206
γ_1	Disutility of working 1000 hours (20 per week)	0.92654	0.24025
γ_2	Disutility of working 2000 hours (40 per week)	0.82177	0.16702
γ_3	Disutility of working 2250 hours (45 per week)	1.64690	0.39486
γ_4	Disutility of working 2500 hours (50 per week)	1.51608	0.35264
γ_5	Disutility of working 3000 hours (60 per week)	2.16258	0.57946
$\kappa_1(\tau = \text{low})$	Correction coefficient with disutility of work	0.61153	0.58616
κ_2	Quadratic coefficient on age for older workers	0.00142	0.00062
κ_3	Linear coefficient on age for young workers	0.04804	0.03242
ξ	CRRA coefficient in utility of bequest	0.46775	0.48061
b_{scale}	Scale multiplicator of the utility of bequest	0.67227	2.02012
$\beta(\tau = \text{hs})$	Discount factor, highschool	0.96944	0.00297
$\beta(\tau = \text{dr})$	Discount factor, dropouts	0.96970	0.00403
$\beta(\tau = \text{cl})$	Discount factor, college	0.96963	0.00367
λ	Scale of EV taste shocks	0.83949	0.39929

Human capital accumulation process

Parameter	Description	Estimate	Std.Err.
$\eta_0(\tau = \text{cl})$	Constant for college	2.93936	1.37286
$\eta_0(\tau = \text{hs})$	Constant for high school	2.61254	1.56208
$\eta_0(\tau = \text{dr})$	Constant for dropouts	2.38097	1.38154
$\eta_0(\tau = \text{high})$	Constant for high type	0.13360	1.60836
η_1	Age (time index)	0.02753	0.01937
η_2	Age (time index) square	-0.00076	0.00044
$\eta_3(\tau = \text{cl})$	Work experience for college	0.03125	0.02754
$\eta_3(\tau = \text{hs})$	Work experience for high school	0.02200	0.02893
$\eta_3(\tau = \text{dr})$	Work experience for dropout	0.01991	0.03011
$\eta_4(\tau = \text{cl})$	Work experience square for college	-0.00017	0.00130
$\eta_4(\tau = \text{hs})$	Work experience square for high school	-0.00002	0.00120
$\eta_4(\tau = \text{dr})$	Work experience square for dropout	-0.00000	0.00118

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

Validation of structural models

- Gold standard: randomized experiment
  1. Estimate the model using control (treated) group
  2. Match the effect of the treatment by simulation
- Second best: observed policy change
  1. Estimate the model using data from before policy change
  2. Match the effect of the policy by simulation
- May be hard to do due data availability

Validation using 2009 reform

1. Increase in maximum pension benefit
2. Changes in taper rate (second threshold)
3. Gradual increase in the eligibility age

Data before and after 2009 reform: hours

Data before and after 2009 reform: hours | working

Data before and after 2009 reform: working

Validation using 2009 reform

- Data before and after the reform shows very little change!
- Validation exercise not possible due to data availability

Policy simulations

Policies to be simulated:

1. Transitory and permanent wage decrease
2. Maximum old age pension benefit
3. Taper rates in income and asset tests
4. Additional lump sum in superannuation
5. Higher superannuation contributions

Baseline: No policy change

Anticipated: Fully anticipated policy change

Unanticipated: Exogenous shift from regime 1 to regime 2

- 1000 individuals in each education/type
- Identical sequence of (pseudo) random variables in all simulations
- Varying revelation age

Special simulation algorithm

Two policy regimes:

1. Baseline:, $r_t = 0$
   Solve the model under old policy regime
   Simulate data using policy functions with $r_t = 0,\ t = 1, \ldots, T$
2. Under new policy, $r_t = 1$
   Solve the model under new policy regime
   1. Fully anticipated:
      Simulate data using policy functions with $r_t = 1,\ t = 1, \ldots, T$
   2. Unanticipated:
      Simulate data using policy $r_t = 0,\ t = 1, \ldots, t_r - 1$ and
      $r_t = 1,\ t = t_r, \ldots, T$

- Exogenous shift from $r_t = 0$ to $r_t = 1$ at a set of revelation ages t_r

Permanent 10% wage decrease $\rightsquigarrow$ hours

No compensation $\longrightarrow$ Marshall effects

Permanent 10% wage decrease ⇝ hours

Permanent 10% wage decrease $\rightsquigarrow$ hours

No compensation $\longrightarrow$ Marshall effects

Permanent 10% wage decrease ⇝ hours

Permanent 10% wage decrease ⇝ hours

Permanent 10% wage decrease $\rightsquigarrow$ hours (high school)

No compensation $\longrightarrow$ Marshall effects

Permanent 10% wage decrease $\rightsquigarrow$ hours (dropouts)

No compensation $\longrightarrow$ Marshall effects

Permanent 10% wage decrease $\rightsquigarrow$ hours (college)

No compensation $\longrightarrow$ 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
- No increase in hours at any point after the wage decrease if policy is anticipated, sizable increase in hours if wage decrease is unanticipated and happens early in the life cycle

- For dropouts: larger increase before the wage drop if policy is anticipated, no positive effect on hours at any point of life cycle
- For college graduates: much larger positive effect, up to wage change at age 50

Transitory 10% wage decrease $\rightsquigarrow$ hours (high school, %)

Transitory $\longrightarrow$ Frisch effects

Implied labor supply elasticities summary

Marshall (permanent unanticipated, uncompensated)

	< 30	$30 - 50$	$50 - 60$	> 60
College graduates	0	small negative	0 to 0.5	over 1.0
High school	0.05	0.05 to 0.2	0.3 to 0.8	1.2
Dropouts	0.1	0.1 to 0.4	0.6 to 0.9	1.05

Frisch (transitory anticipated, compensated)

	< 30	$30 - 50$	$50 - 60$	> 60
College graduates	<0.2	0.2 to 0.4	0.4 to 0.9	1.2
High school	0.3	0.3 to 0.5	0.5 to 1.0	over 1.2
Dropouts	0.5	0.5 to 0.7	0.8 to 1.0	over 1.0

Intensive vs. extensive margin in labor supply elasticities

- Permanent 10% wage decrease $\rightsquigarrow$ probability of working
- Permanent 10% wage decrease $\rightsquigarrow$ hours conditional on working
- Relative changes (%)
- Unanticipated wage decrease

- Evidence of significantly higher elasticity on the extensive margin

Intensive vs. extensive margin, high school graduates (%)

Intensive vs. extensive margin, dropouts (%)

Intensive vs. extensive margin, college graduates (%)

Effects of changes in age pension rules

Policy parameters in age pension:

- Maximum pension benefit (+25%)
- Taper rate in income test (-10%)
- Taper rate in asset test (-10%)

Effects on:

1. $\rightsquigarrow$ hours of labor supply (hours per annum)
2. $\rightsquigarrow$ wealth ($1000)
3. $\rightsquigarrow$ consumption ($1000)

- Only high school graduates

Maximum age pension +25% ⇝ hours (annual and %)

Maximum age pension +25% ⇝ wealth ($1000 and %)

Maximum age pension +25% ⇝ consumption

Asset taper rate -10% ⇝ hours (annual and %)

Asset taper rate -10% ⇝ wealth ($1000 and %)

Asset taper rate -10% ⇝ consumption ($1000 and %)

Income taper rate -10% ⇝ hours (annual and %)

Income taper rate -10% ⇝ wealth ($1000 and %)

Income taper rate -10% ⇝ consumption ($1000 and %)

Cumulative effects of the age pension

Simulate the world without Age Pension

Compute effects on:

- $\rightsquigarrow$ hours of labor supply (hours per annum)
- $\rightsquigarrow$ participation (extensive margin)
- $\rightsquigarrow$ consumption ($1000)
- $\rightsquigarrow$ wealth ($1000)
- $\rightsquigarrow$ welfare effects (changes in value function) at different ages

Unchanged taxes vs. Revenue neutral

Elimination of Age Pension ⇝ participation

Elimination of Age Pension ⇝ total hours

Elimination of Age Pension ⇝ consumption ($1000)

Elimination of Age Pension ⇝ wealth ($1000)

Revenue neutral ⇝ expected discounted utility at 19

Revenue neutral ⇝ expected discounted utility at 65

Effects of the age pension

The world without Age Pension

Main effects (revenue neutral):

- Labor supply increases for dropouts and high school graduates
- Labor supply decreases for college graduates
- Tax rates fall by 42%
- Welfare improves for all education groups on average
- Small fraction of low skill type individuals experience decrease in welfare

Preliminary results

Labor supply

- Estimates of labor supply elasticity correspond well to the existing literature
- Large variation of labor supply elasticities over the lifecycle and among education levels:
  - Increases over the lifecycle
  - Smaller for the higher education groups
- Stronger effects on the extensive margin for all education levels

Age Pension

- The program has large negative labor supply effects
- The program may be poorly targeted $\rightarrow$ Very low taper rates
- Elimination would allow 42% reduction in income tax rates and improve welfare on average