Economic inequality over the life cycle in Australia

by Owen Freestone

Submitted September 2019

Re-submitted with corrections February 2020

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

This thesis is entirely my own work except for:

- Chapter 2 which is based on joint research with Robert Breunig, the chair of my PhD panel; and
- material that is explicitly cited throughout the body of each chapter and referenced in the bibliography.

The total world count is 36 000 excluding tables, figures, footnotes, appendices and bibliography.

Thank you to my supervisor Bob Breunig for excellent advice and input throughout my PhD, for his willingness and ability to work flexible, and being a leading example of successful academic-government collaboration on economic policy development in Australia. I would also like to thank the other members of my advisory panel: Chung Tran, for his invaluable advice on Chapter 5; and Garry Barrett, for his willingness to engage and provide feedback. Thank you also to my fellow students, Joseph Chien, Nathan Deutscher and Nabeeh Zakariyya, for providing valuable input as discussants for my seminars, and to Paul Burke for engendering a constructive and supportive seminar series within the Crawford School of Public Policy. I am grateful to Michael Kouparitsas for guidance and inspiration in selecting my PhD topic, and to my other Australian Public Service mentors notably Angelia Grant. I would also like to acknowledge the generous financial support of the Australian Treasury and the Sir Roland Wilson Foundation via the Foundation's scholarship program. Last but not least, thank you to my wife and son for their understanding during my PhD.

This thesis was proofread by Kathie Brown, A+ Academic Editing, Canberra.

This thesis investigates the key factors that govern the economic behaviour of Australian households over their life cycle and contributes to the study of optimal income taxation within an overlapping generations (OLG) framework. Of the thesis's five chapters, three are empirical in nature making use of Australian household panel data[1] to estimate the underlying preference parameters of Australian households and key aspects of the economic environment in which they make their decisions. These chapters go some way to addressing the lack of available Australian evidence on these crucially important parameters. The other two chapters comprise: a short but novel Monte Carlo study that compares the statistical performance of dynamic wage models estimated in levels versus growth rates; and a study of optimal income taxation in the presence of dual-earner households and marital risk which extends the existing OLG literature in this area.

A brief summary of each chapter follows.

Chapter 1 explores the life cycle pattern of wage inequality in Australia, specifically, the rise in wage inequality with age. Using panel data on male wages, it explores the relative importance of unobserved worker heterogeneity versus random wage shocks in explaining this empirical regularity for Australia. It finds significant heterogeneity in wage *levels* (via differences in starting wages) but no evidence of systematic heterogeneity in wages *growth*. Instead, highly persistent wage shocks are found to account entirely for the rise in wage inequality with age. The chapter also considers the factors behind the upward trend in wage inequality since the early 2000s, providing evidence that this trend reflects permanent rather than transitory factors.

Chapter 2 explores the statistical properties of the generalised method of moments (GMM) estimator which underpins the empirical results of Chapter 1. In particular, this Monte Carlo study explores whether GMM estimation of dynamic wage models is more accurate at recovering key parameters when specified in levels or growth rates in the context of relatively short panel data sets.

Chapter 3 uses panel data to explore risk aversion among Australian households. Using households' share of risky assets, it tests whether relative risk aversion is constant in wealth. After accounting for measurement error, the constant relative risk aversion (CRRA) assumption cannot be rejected. Using a consumption Euler equation that adjusts for measurement error in consumption data, I estimate the coefficient of relative risk aversion in the CRRA utility function. Point estimates from the preferred non-linear models suggest a moderate degree of risk aversion for the typical Australian household, with values ranging from 1.2 to 1.4. These findings can provide guidance for calibrating household preferences in macroeconomic models of the Australian economy.

Chapter 4 provides estimates of the Frisch (or wealth-compensated) elasticity of labour supply for Australian workers using 17 waves of panel data. This wage elasticity is central to many economic and policy questions, yet very little empirical evidence exists on its value in Australia. Critically, the chapter

[1]In particular, I make extensive use of the Household, Income and Labour Dynamics in Australia (HILDA) survey. As elaborated upon elsewhere in this thesis, the HILDA survey collects data annually on a broad range of economic and social topics. It began in 2001 and the latest year of data is 2017.

uses an estimation approach that incorporates the extensive (employment) as well as the intensive (hours worked) labour supply response, and finds that the resulting aggregate elasticity is large, ranging from 1.5 to 2.4 using a variety of techniques. This is well above the 0 to 0.5 range typically found in international labour supply studies of prime-age men that ignore the extensive margin. The chapter's sub-aggregate results confirm a relatively large wage elasticity for women and older age groups, especially along the extensive margin of labour supply.

Chapter 5 re-examines optimal taxation of capital in the long run. Recent studies using OLG models cast doubt on the classic public finance result of Atkinson and Stiglitz (1976) that the optimal tax is zero. The chapter explores whether these findings carry over to a more realistic environment that includes dual-earner households and life cycle marital risk. Analytical results are first derived from a simple two-period model. For plausible parameter values, the saving of married households is *higher* and, conversely, the saving of single households is *lower* than in a world without divorce and marriage risk. The net impact of these competing effects is to reduce saving in the most empirically relevant cases. The chapter then examines optimal tax in a fully-fledged life cycle model with dual earners and marital risk (in addition to other standard features) calibrated to the United States' economy. The optimal steady state tax on capital is found to be zero. The main reasons are twofold: first, a relatively low tax on capital mitigates the tendency for lower saving induced by marital risk demonstrated in the two period model. Second, the dual labour supply margin available to married households acts as a substitute for precautionary saving, making aggregate saving more elastic to interest rate changes. This increases the welfare costs of capital taxation. This chapter also highlights the role that dual labour supply plays as an insurance mechanism (in addition to savings) against idiosyncratic wage uncertainty. In doing so, it establishes an important link between optimal taxation and the study of partial insurance in incomplete market environments (e.g. Blundell et al. (2016)).

Contents

1 The drivers of life-cycle wage inequality in Australia **1**

1.1 Introduction and related literature 1

1.2 Data and empirical context 3

1.2.1 Data 3

1.2.2 Empirical context 4

1.3 The empirical model 4

1.3.1 Modelling unobserved worker heterogeneity 5

1.3.2 Modelling the stochastic component of wages 6

1.3.3 The overall model 6

1.3.4 Allowing for time and cohort effects 8

1.4 Identification strategy 9

1.4.1 Problems with estimation based wage growth rates 9

1.4.2 Estimation using wage levels and cohort groupings 10

1.4.3 Visual evidence from sample covariances 11

1.5 Results 12

1.5.1 Main results 12

1.5.2 Discrepancy between wage level and growth estimates 13

1.5.3 Components of life-cycle wage inequality 14

1.5.4 Trends in wage inequality and mobility over time 15

1.5.5 Comparison with US and other overseas studies 16

Appendices **19**

Appendix 1.A Data and sample selection 19

Appendix 1.B Estimation details 20

Appendix 1.C Sensitivity analysis 20

2 GMM estimation of dynamic wage models: levels or growth rates? **23**

2.1 Introduction 23

2.2 Methodology 24

2.3 Results 25

3 Risk aversion among Australian households **29**

3.1 Introduction 29

3.2 Background 30

3.3 Data 32

3.3.1 Risky asset share regressions 33

3.3.2 Consumption Euler equations 33

3.4 Testing the CRRA assumption using share of risky assets 34

3.4.1 Previous estimates 34

3.4.2 Our approach 35

3.4.3 Results 36

3.5 Estimates of risk aversion using a consumption Euler equation 38

3.5.1 Exact non-linear equation versus the linearised version 39

3.5.2 Our approach . . . 40

3.5.3 Results . . . 42

3.6 Conclusion . . . 44

Appendices **45**

Appendix 3.A Sample selection . . . 45

Appendix 3.B Relative risk aversion estimates by net worth . . . 45

Appendix 3.C Impact of measurement error in wealth on risky asset regressions . . . 46

Appendix 3.D Estimates of measurement error in grocery data . . . 48

4 Empirical evidence on the Frisch elasticity of labour supply for Australia **51**

4.1 Introduction and related literature . . . 51

4.2 Standard life cycle model of labour supply . . . 53

4.3 Estimation issues . . . 55

4.3.1 Estimation challenges and instruments . . . 55

4.3.2 Accounting for the extensive margin . . . 55

4.4 Data . . . 58

4.5 Results . . . 58

4.5.1 Further exploration . . . 62

4.5.2 Omissions and caveats . . . 63

4.5.3 Implication for calibrating preference parameters and policy experiments . . . 64

4.6 Conclusion . . . 64

Appendices **67**

Appendix 4.A Sample selection . . . 67

Appendix 4.B Sensitivity checks . . . 67

5 Optimal income taxation, marital risk and family labour supply **69**

5.1 Introduction 69

5.2 Impact of adding marital risk and dual earner households 72

5.2.1 A two-period model with marital risk 72

5.2.2 Dual labour supply and partial insurance 76

5.2.3 Empirical evidence on saving by marital status 77

5.3 The model 78

5.4 Calibration 82

5.5 The policy experiment 87

5.6 Results and discussion 87

5.6.1 The optimal tax system in the benchmark economy 87

5.6.2 Sensitivity 88

5.6.3 Dual labour supply and insurance 89

5.6.4 Relationship to Erosa and Gervais (2002) and Gervais (2012) 92

5.7 Further work 95

Appendices **97**

Appendix 5.A Computational appendix 97

5.A.1 Household decision rules and marginal value functions 97

5.A.2 Calculation of stationary asset distribution and optimal tax mix 98

Disclaimer **109**

Chapter 1

The drivers of life-cycle wage inequality in Australia

1.1 Introduction and related literature

It is a well established empirical fact that wage inequality rises with age. But what are the factors that lie behind this widening dispersion over the life cycle? What is the relative importance of worker heterogeneity (such as skills and talents) versus random wage shocks? What portion of life-cycle wage inequality can be attributed to initial differences in wages when a cohort of workers first enters the labour market? We address these questions in an Australian context, and in doing so become the first Australian study to directly confront two major controversies raised in the empirical literature on life-cycle wage dynamics.

The first of these controversies involves two fundamentally different views of individual wage dynamics. In the first view, known as the heterogeneous income profiles (HIP) model, fixed differences between workers (in talents, skills and other characteristics) result in individual-specific wage trajectories over the life cycle. While not denying a significant role for random wage shocks, the HIP model interprets these fixed differences in wage growth rates between workers as a large contributor to the observed increase in wage inequality with age.[1] In the rival model, known as the restricted income profiles (RIP) model, workers are viewed as having a common rate of expected wage growth over their life cycle, with any *ex post* deviation from this expected growth rate entirely driven by idiosyncratic wage shocks. The RIP model thus attributes rising wage inequality purely to the cumulative effect of random wage shocks.[2]

A second related debate concerns the degree of persistence of wage shocks faced by individuals over their

[1]The HIP model is usually defined in terms of years of labour market experience rather than age. In this paper, we use age as a shorthand for experience. We show in Appendix 1.C that this is an innocuous assumption.

[2]The RIP model, like the HIP model, does not rule out an important role for worker heterogeneity in explaining differences in starting wages, and therefore differences in wage *levels*. Indeed, this is exactly what we find in our results below.

working lives. At one extreme, the stochastic component of wages is modelled as a unit root process, so that wage shocks have a permanent effect on a worker's subsequent earnings level (as in the classical permanent income hypothesis). In the alternative model, the wage process is a persistent, but ultimately mean-reverting process.

Using 14 years of Australian panel data on male wages, we estimate a model that encompasses these competing theories of individual wage dynamics. Our results reject the HIP model. This finding is robust if we replace hourly wage rates with labour income (or 'earnings'), and holds across sub-samples broken down by educational attainment. It is also robust to whether the model is estimated using moments constructed from wage levels or wage growth rates. Second, we reject a unit root in the wage process, with an estimated autocorrelation coefficient of around 0.95, implying highly persistent but not permanent wage shocks. Given the well-established difficulty of detecting heterogeneous wage growth rates and a unit root in short, unbalanced panels (Gustavsson and Österholm, 2014), we address this concern using evidence from Monte Carlo simulations. We show that our particular estimation strategy is relatively effective at delineating between the RIP and HIP models, and generally provides more accurate estimates of the autocorrelation coefficient than the alternative strategies.[3]

Our study goes beyond the few previous Australian studies on life-cycle income dynamics. Chatterjee, Singh and Stone (2016), the Australian study most comparable to ours, impose rather than test both the RIP and unit root assumptions in their model of individual wages.[4] Our econometric approach also differs to theirs, as they use a small subset of information provided by wage autocovariances, whereas we fit our model to the entire autocovariance matrix using generalised method of moments (GMM). Guvenen (2009) shows that these higher-order autocovariances can shed light on the true model. We also divide the sample into four cohort groupings, which enables us to exploit the resulting age variation across the covariance matrix.[5] Our estimation approach also allows us to undertake a full decomposition of wage inequality, not just changes in wage inequality as in Chatterjee et al.[6] More generally, our results confirm Chatterjee et al. insofar as we find a roughly similar relativity in the variances of transitory and persistent wage shocks. We are also able to reconcile much of the discrepancy noted in Chatterjee et al. (2016) between estimates based on wage levels and wage growth rates by making the sample adjustments suggested in Manovskii et al. (2015).

Our study also adds to a more extensive international literature on individual wage dynamics.[7] Despite

[3]The relative performance of levels versus growth-based GMM estimates has not previously been examined to our knowledge. Doris et al. (2010) look at the performance only of levels-based estimates, and their wage model does not allow for either heterogeneous growth or purely transitory shocks. Likewise, Dmytro (2012) focuses only on the performance of growth-based estimation, whereas Manovskii et al. (2015) look at levels versus growth-based estimation but only in relation to reconciling differences in estimates of persistent and transitory shocks.

[4]Dollman et al. (2015) allow for a non-unit root but do not test for heterogeneous growth. However, the HIP model seems less relevant in their application, as they were attempting to explain the dynamics of disposable household income rather than individual wages or earnings.

[5]This point is also demonstrated in our Monte Carlo analysis in Chapter 2.

[6]Chatterjee et al. also use wage levels, but proceed to first-difference the moments in adjacent periods, which removes the accumulated impact of the persistent wage shock and any fixed effect. This means they can only comment on *changes* in inequality, not the *level* of inequality.

[7]Early examples include Lillard and Weiss (1979) and MaCurdy (1982). Meghir and Pistaferri (2011) provide a thorough survey of the key analytical techniques, main findings and ongoing controversies with a focus on the US literature. Recent examples outside the US include the Canadian study by Baker and Solon (2003), the UK study by Blundell and Etheridge

the large number of studies, particularly in the United States (US), no consensus has been reached on either the heterogeneous growth or unit root questions.[8] Guvenen (2009), whose methodology we follow closely, finds strong support for the HIP model in the US, but this finding has been recently challenged by Dmytro (2012). Nonetheless, in Section 5.6 we discuss why our finding with regard to the HIP model may differ to Guvenen's US finding, including the influence that differing labour market policies may have on each country's life-cycle wage dynamics.

We also see our attempt to gain a deeper understanding of life-cycle wage dynamics as a useful input into broader economic and policy analysis in Australia. This recognises that life-cycle wage dynamics have important implications for many other economic questions, and are a primary determinant of lifetime welfare (Storesletten et al., 2004; Huggett et al., 2011). A realistic depiction of individual wage dynamics is also a key input for dynamic stochastic models with incomplete markets, which are becoming a standard tool for quantitative tax and broader policy analysis where distributional effects are important.[9]

In Australia, our study of life-cycle wage dynamics complements wider research into wage and income inequality.[10] Indeed, it is worth noting that our paper only investigates the *proximate* causes of wage inequality in a reduced form, and no attempt is made to add to the distinct literature that explores the underlying causes of wage inequality, be it socio-economic (for example, Whiteford (2015)) or innate worker heterogeneity (for example, Barrett (2012)).

The rest of the paper is organised as follows. Section 4.4 presents the key life-cycle features of the wages data, including the need to control for possible time and cohort effects. Section 4.2 presents the model for individual wages and discusses ways of allowing for non-stationary features of the wages data, particularly time effects. Section 1.4 emphasises the difficulty in delineating between the HIP and RIP models empirically, and discusses alternative identification strategies. Section 5.6 presents the main results, compares them with other studies and discusses the implications for trends in wage inequality and mobility. Section IV provides some concluding remarks.

1.2 Data and empirical context

1.2.1 Data

This paper uses individual-level wage data from HILDA. The main results are based on the real hourly wage rate of working-age men.[11] This is constructed by dividing weekly gross wages and salaries by

(2010) and Manovskii et al. (2015) who investigate both Danish and German administrative data. The 2010 Special Issue of the *Review of Economic Dynamics* by Krueger et al. (2010) summaries inequality trends across a range of other countries.

[8]See Table 2 in Meghir and Pistaferri (2011), which summarises the findings of studies based on the US Panel Survey of Income Dynamics (PSID).

[9]Guvenen (2011) and Heathcote et al. (2009) provide recent surveys of this research area.

[10]Borland and Coelli (2016) provide a recent survey of Australian empirical studies in this area. Some earlier Australian studies, such as Barrett et al. (2000) have also looked at wage dynamics, but were limited by the lack of panel data at the time.

[11]Appendix 1.C demonstrates that the paper's main results regarding the HIP model are unaffected if earnings are used in place of hourly wages.

hours per week usually worked in all jobs then deflating by the Australian Consumer Price Index. We aid comparison with previous studies by selecting a similar sample of men aged between 23 and 60 inclusive whose, usual hours are between 10 and 98 hours per week.[12] The estimation period spans 2001 to 2014 inclusive. To maximise the number of observations, we opt for an unbalanced panel and allow anyone who appears in at least two consecutive waves to feature in the sample. This leaves a total of around 40,000 observations across 14 years to be used in the main results.[13]

1.2.2 Empirical context

Focusing on a few key dimensions of the wages data will guide our choice of model specification below. First, *average* wages display an upward trend over the life cycle, peaking when workers reach their mid-50s. This is a familiar empirical fact, consistent with workers' rising productivity as they gain experience and on-the-job skills with more time in the labour market. This common age profile (along with other observable factors) explains only a small portion of wage variability and is removed in a first stage regression as explained further below.

Second, the dispersion in both log hourly wages and earnings increases almost monotonically with age: the key motivation for our study (Figure 1.1).[14] This aggregate picture is confirmed by Figure 1.2, which plots the variance of log wages for four equally-populated cohort groupings born in the years indicated. Although each cohort appears to experience a widening wage dispersion over time, the increase seems to be more pronounced at younger ages. Notwithstanding some variation between cohorts, the age profile appears to be broadly concave, although not far from being linear.[15] Also, the rise in the variance of log wages is concave in age, regardless of whether one controls for time or cohort effects, although the profile is somewhat steeper in the latter case (Chatterjee et al., 2016).[16]

Lastly, Figure 1.3 shows a moderate but clear upward trend in wage inequality over the sample period. This highlights the possible need to allow for non-stationarity in our model of wages below. To summarise, the wages data suggest that any model of wages will need to allow for: a) life-cycle trends in both the mean and variance of wages; b) an age profile in the variance of wages that is non-linear (that is, concave); and c) non-stationary variances over the sample period.

1.3 The empirical model

Our model specification nests the alternative models of individual wages discussed in the introduction, and is similar to that used in Guvenen (2009).

[12] In particular, our sample selection follows Guvenen (2009) and Chatterjee et al. (2016) where practical.
[13] Further details of the sample selection process are contained in Appendix 4.A.
[14] The similar age-profile for hourly wages and earnings (or labour income) is consistent with there being no obvious life-cycle trend in average hours, at least among prime age males. The hours data show a similar pattern in the US (Heathcote et al., 2010).
[15] Hahn, Miller and Yang (2016) also find a concave life-cycle profile when examining earnings data from HILDA.
[16] In the US, it is well known that controlling for cohort effects yields a much steeper age profile than controlling for time effects (Guvenen, 2009; Heathcote et al., 2010, 2005).

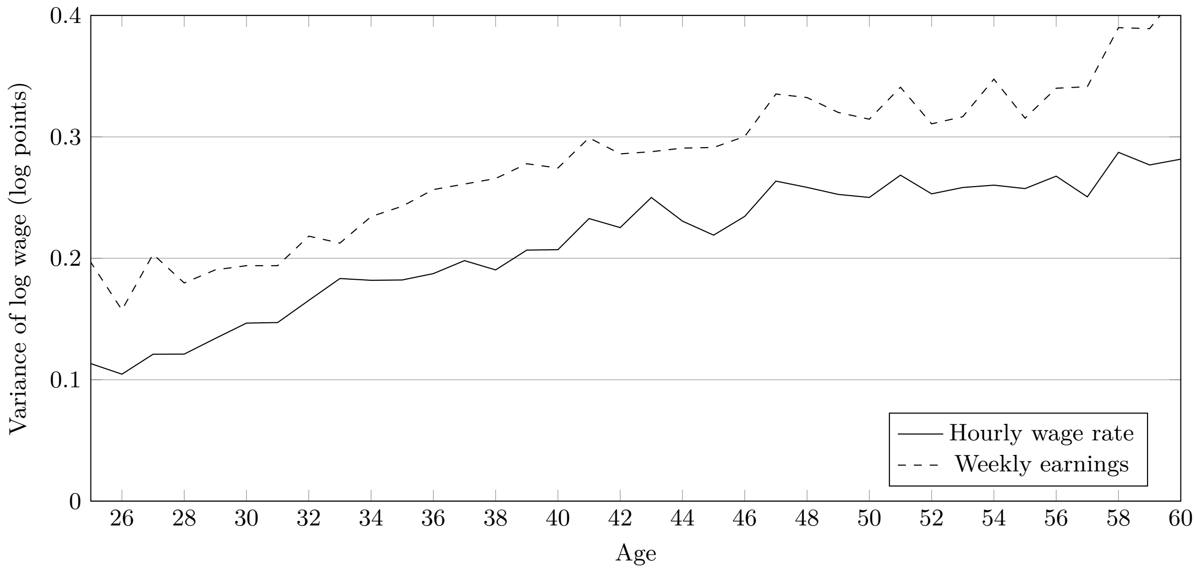

Figure 1.1: Dispersion in male wages

Source: HILDA Release 14; ABS Consumer Price Index

We begin by making the usual distinction between the part of wages that can be attributed to common observable factors and the part that is unobservable. The observable part of wages is removed through a first stage regression of log wages on a second degree polynomial in age (interacted with year dummies) and dummies for year, cohort, marital status, region and non-English speaking background (NESB).[17] The remaining, unobservable component of wages forms the basis of our analysis. It can be broken down further into a part that reflects individual heterogeneity and the portion that reflects idiosyncratic 'wage uncertainty' or 'risk', which we outline in turn.

1.3.1 Modelling unobserved worker heterogeneity

We allow the life-cycle wage profile for each individual i to have its own intercept (α_i) and slope (β_i):

$$y_{i,h}^{Hetero.} = \alpha_i + \beta_i h \tag{1.1}$$

where h is actual age less 22.[18] The idea behind the HIP model is demonstrated in Figure 1.4, which presents stylised age profiles of log wages for two workers, A and B, as well as the mean profile.[19] In essence, testing the HIP model amounts to testing whether there is systematic differences between workers' slope term β_i, that is, whether the variance, σ^2_β, differs from zero. As is clear in Figure 1.4, any systematic difference in the slope of workers' wage profiles will tend to result in rising wage dispersion as

[17]The first stage regression only explains a small portion of wage inequality with an R^2 of 0.18. Adding extra controls for educational attainment, family type, occupational status and union membership increases the R-squared to 0.33. This results in a lower estimated transitory variance (around 0.035) in GMM estimation, but does not materially affect the estimates of other parameters.

[18]h is therefore a simple measure of (potential) labour market experience. In Appendix 1.C we show that our results are largely unchanged when we substitute a measure of actual labour market experience.

[19]Of course, the actual mean life-cycle wage profile is closer to being 'hump shaped' than log-linear.

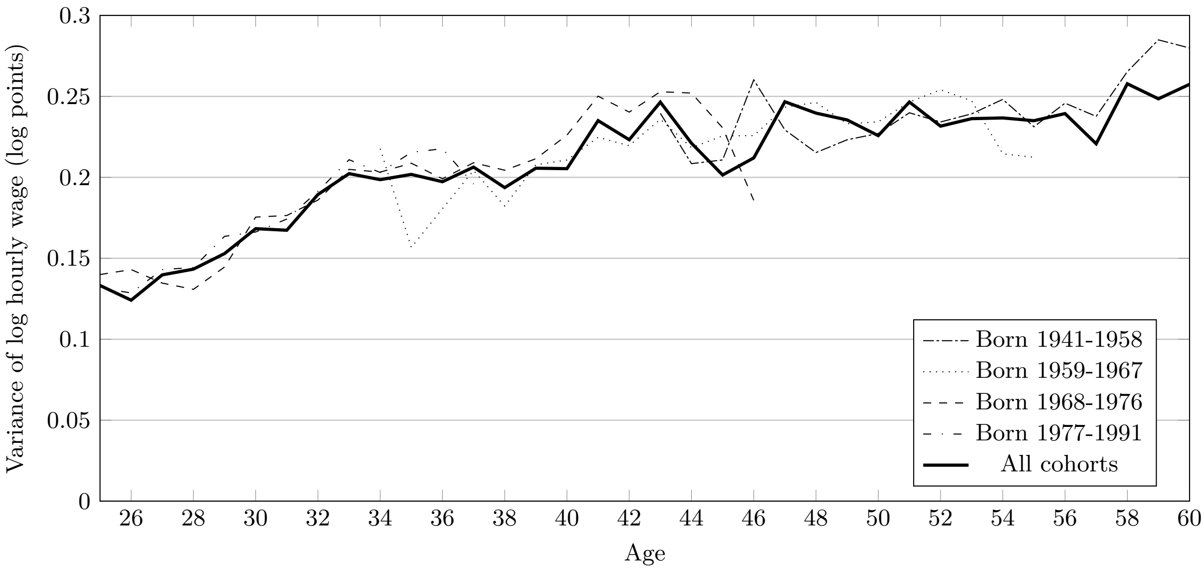

Figure 1.2: Wage dispersion by cohort

the workers age.

The intercept term α_i allows for heterogeneity in wage *levels* even when there is no heterogeneity in wage growth. Whether the variance σ_α^2 is different from zero is therefore a conceptually distinct question and is neither necessary nor sufficient for the HIP model to hold. The intercept (α_i) and slope (β_i) are permitted to be correlated in either direction.[20]

1.3.2 Modelling the stochastic component of wages

The stochastic part of wages is modelled as a first-order autoregressive process ($z_{i,h} = \rho z_{i,h-1} + \eta_{i,h}$) plus a purely transitory shock ($\nu_{i,h}$). We assume each shock, $\eta_{i,h}$ and $\nu_{i,h}$, is independently and identically distributed with mean zero and variance σ_η^2 and σ_ν^2. Allowing for purely transitory wage shocks $\nu_{i,h}$ recognises that most previous empirical studies have found an important role for such shocks. As well as genuine year-to-year wage instability, this component captures measurement error in the income data that inevitably arises in household surveys.[21] The autoregressive coefficient ρ is freely estimated rather than imposing $\rho = 1$ (a unit root) as in many other studies.

1.3.3 The overall model

Putting this altogether, the unobserved wage for individual i of age h of a particular cohort is given by:

[20] This parameter could therefore be used to test competing labour economic theories. For example, a negative sign would be consistent with the theory that individuals who undertake larger investments in human capital might experience lower wages initially but higher wage growth rates as their careers progress.

[21] As noted above, it is well recognised that (classical) measurement error cannot be separately identified without additional assumptions. See Meghir and Pistaferri (2004) for a discussion of this point.

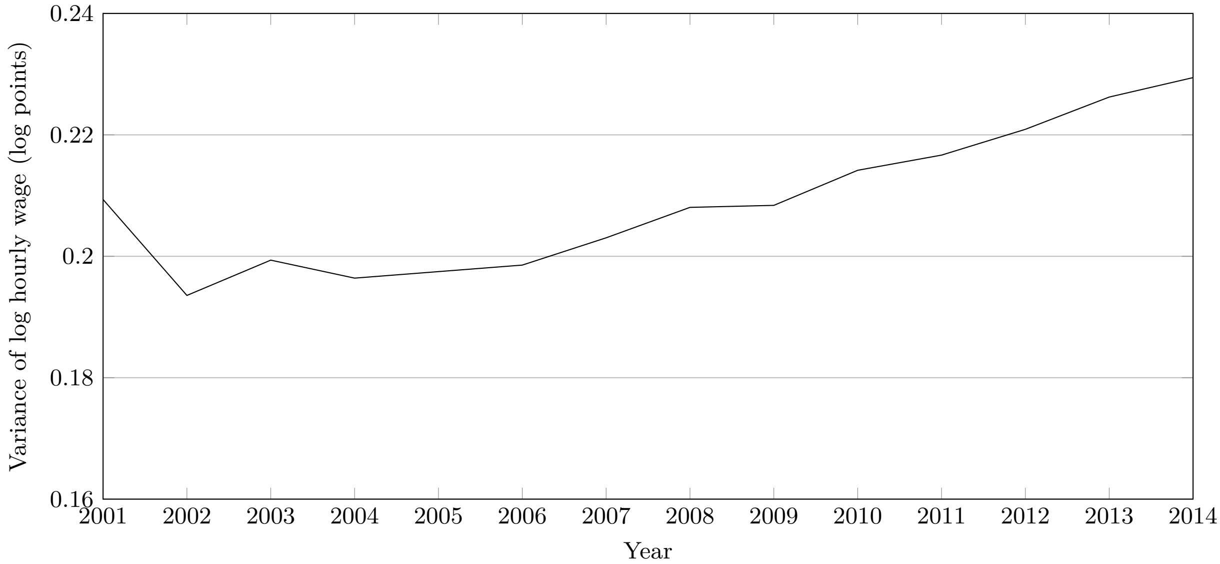

Figure 1.3: Wage inequality over time

$$\hat{y}_{i,h} = \overbrace{\underbrace{\overbrace{\underbrace{\alpha_i}_{\textit{Ind. fixed effect}} + \underbrace{\beta_i h}_{\textit{Ind. growth}}}^{\text{Ind. heterogeneity}}}_{} + \overbrace{\underbrace{z_{i,h}}_{Persistent(AR(1))} + \underbrace{\nu_{i,h}}_{Transitory}}^{\text{Wage uncertainty}}}^{\text{Unobserved part of wages}} \tag{1.2}$$

where $z_{i,h} = \rho z_{i,h-1} + \eta_{i,h}$. The variances of the individual-specific intercept and slopes are σ^2_α and σ^2_β.

This produces the following variance-covariance matrix:

$$\begin{aligned} \mathrm{var}(\hat{y}_{i,h}) &= \sigma^2_\alpha + h^2\sigma^2_\beta + 2\sigma_{\alpha\beta}h + \rho^2\mathrm{var}(z_{i,h-1}) + \sigma^2_\eta + \sigma^2_\nu \\ \mathrm{cov}(\hat{y}_{i,h}, \hat{y}_{i,h+n}) &= \sigma^2_\alpha + h(h+n)\sigma^2_\beta + (2h+n)\sigma_{\alpha\beta} + \rho^n\mathrm{var}(z_{i,h}) \end{aligned} \tag{1.3}$$

for n=1,...,13.

Given the recursive structure of the AR(1) process, $\mathrm{var}(z_{i,h})$ incorporates the accumulation of past shocks up to h. In practice, we do not have complete work histories for each cohort and must make an assumption about the size of shocks prior to the beginning of the sample. For this purpose, we assume that:

$$\begin{aligned} var(z_{i,1}) &= \sigma^2_\eta \\ var(z_{i,h_1}) &= \sigma^2_\eta \sum_{j=0}^{h_1-1} \rho^{2j} \end{aligned} \tag{1.4}$$

In other words, we assume that the variance of $z_{i,h}$ is zero at age 22 and that the variance of $\eta_{i,h}$ in years prior to the person being first observed is equal to the variance of $\eta_{i,h}$ over the estimation period. The assumption that the variance of the persistent process $z_{i,h}$ is zero at age 22 enables σ^2_α to be identified.[22]

[22]Choosing a slightly different age (for example age 24) to initialise $\mathrm{var}(z_{i,h-1})$ has a negligible impact on the paper's main results.

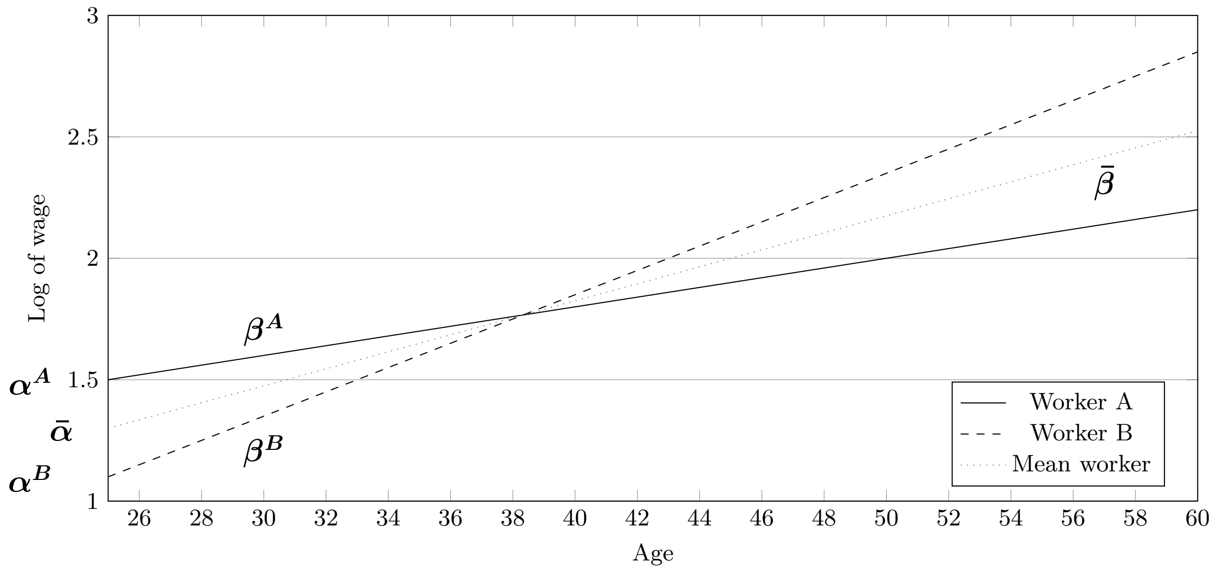

Figure 1.4: Stylised heterogeneous wage profiles

1.3.4 Allowing for time and cohort effects

We also consider numerous extensions of this basic model. In particular, we attempt to capture the observed non-stationarity in the wages data (as evidenced by the upward trend in wage inequality shown in Figure 1.3) through the inclusion of calendar time factor loadings π_t and λ_t on the permanent (α_i) and transitory ($\nu_{i,h}$) components of wages.[23] In this context, higher values of π_t can be interpreted as increased returns to unobserved skills, and associated with greater inequality in long-run earnings. In contrast, higher values of λ_t imply greater year-to-year 'scrambling' in the income distribution which is then unwound in the following year (Baker and Solon, 2003). We also include cohort factor loadings on the permanent component of wages, but omit these for brevity.[24] This results in the following modified model variances and covariances:

$$
\begin{aligned}
\text{var}(\hat{y}_{i,h,t}) &= \pi_t^2\sigma_\alpha^2 + h^2\sigma_\beta^2 + 2\sigma_{\alpha\beta}h + \text{var}(z_{i,h}) + \lambda_t\sigma_\nu^2 \\
\text{cov}(\hat{y}_{i,h,t}, \hat{y}_{i,h+n,t+n}) &= \pi_t\pi_{t+n}\sigma_\alpha^2 + h(h+n)\sigma_\beta^2 + (2h+n)\sigma_{\alpha\beta} + \rho^n\text{var}(z_{i,h,t})
\end{aligned}
\tag{1.5}
$$

for n=1,...,13[25]

[23]We also experimented with time factor loadings on $\eta_{i,h}$ instead of α_i as an alternative way of capturing the non-stationary in the wages data. The results were very similar and did not affect our conclusion regarding the HIP model.

[24]It is well-known that one cannot simultaneously control for age, time and cohort effects in a linear regression. However, the age terms in our model are non-linear, making it possible to include all three in this case.

[25]We normalise each set of factor loadings to be equal to 1 in first year/cohort, which permits identification.

1.4 Identification strategy

1.4.1 Problems with estimation based wage growth rates

Before proceeding further, it is worth demonstrating why it is inherently difficult to distinguish empirically between the RIP and HIP models, particularly using moments based on wage growth rates. The problem arises because hypothesised values of σ_β^2 under the HIP model, while economically large, are numerically small (Baker, 1997).[26] This means that empirical tests designed to detect heterogeneous growth are inherently low-powered (Guvenen, 2009), especially early versions that were based on wage growth moments, such as the test proposed by MaCurdy (1982). To see this, observe that the wage *growth* autocovariances for our basic model in Eq. (1.2) are given by:

$$\mathrm{cov}(\Delta\hat{y}_{i,h}, \Delta\hat{y}_{i,h+n}) = \sigma_\beta^2 - \rho^{n-1}\left(\frac{1-\rho}{1+\rho}\sigma_\eta^2\right) \tag{1.6}$$

for $n \geq 2$.

In the case of a unit root, the second term is zero and the HIP model makes a very clear prediction: all autocovariances beyond order 1 should be positive. Although potentially lacking in power due to the small size of σ_β^2, a simple test against the null of these terms being equal to zero is likely to detect the existence of heterogeneous growth in a moderately-long panel (Baker, 1997; Hryshko, 2012).

However, for $\rho < 1$, a test based on Eq. (1.6) becomes less clear-cut. This is because the second term is now negative, which makes the sign of these autocovariance terms parameter- and lag-dependent. In fact, the sign of Eq. (1.6) is now likely to switch from negative to positive at some critical lag order, say $n^\star$, with a higher value of ρ (for $\rho < 1$) implying a higher $n^\star$ all else equal. In such cases, Guvenen (2009) shows that a test based on these autocovariance terms being positive will often fail to detect heterogeneous growth even when it is present.

A more promising approach explored by Hryshko (2012) is to fit the wages model to the entire covariance matrix of wage growth rates using GMM. His Monte Carlo study shows that this approach can accurately recover the true value of σ_β^2 (and the model's other parameters) across a wide range of specifications and parameter values. However, he also finds that this approach is less able to detect the heterogeneous growth term when the true model contains a highly persistent autoregressive process (ie, $\rho \approx 0.95$).[27] Unfortunately, the international evidence – as well as the Australian evidence presented below – suggests that the true value of ρ may actually lie in this corridor of uncertainty.[28]

[26] For example, the estimate for σ_β^2 of just 0.0008 found in Baker is considered to be economically large. It implies that a worker with wage growth rate one standard deviation above the mean would open up a 30 per cent wage advantage after 10 years compared with a worker experiencing average wage growth rate over the same time.

[27] In particular, Hryshko shows that GMM estimation based on wage growth rates from a simulated small, unbalanced panel produces insignificant estimates of σ_β^2 when the true model contained heterogeneous growth (σ_β^2=0.0004).

[28] Two recent panel studies (Gustavsson and Österholm, 2014; Browning et al., 2010), which employ more sophisticated methodologies that allow for heterogeneity in ρ, both reject a unit root in wages for most workers. They instead find support for a moderately to highly persistent wage process. While a Dickey-Fuller style test might seem like another good way of proceeding, these tests offer little insight as they overwhelmingly favour the null of a unit root in short panels (Andrews,

Manovskii et al. (2015) raise another concern in using wage growth rates to estimate the model's parameters. They find that the well-documented discrepancy between growth- and levels-based estimates of the persistent wage shock variance in unbalanced panels is caused by greater wage volatility among individuals in their first (last) year after (before) entering (leaving) the panel.[29] They show that this results in a substantial upward bias in estimates of σ_η^2 when estimation is based on wage growth rates. We investigate this issue below for Australia.

1.4.2 Estimation using wage levels and cohort groupings

We follow Baker and Solon (2003), Heathcote et al. (2005) and Guvenen (2009) among others in using covariance information from wage *levels* to identify the model's parameters, although we also estimate our model using wage growth rates for comparison.[30] Besides being able to recover additional parameters (the accumulated value of z_h, σ_α^2 and $\sigma_{\alpha\beta}$), using wages levels exploits covariance information that is lost when the data is differenced. In particular, the HIP model should generate a convex increase in the variance of log wages by age, which follows from the quadratic term h^2 contained in Eq. (1.3) (Guvenen, 2009). Intuitively, even a small difference in wage growth rates between workers will contribute to a larger and larger gap *in levels* over time due to the effects of compounding. By contrast, a purely autoregressive process can generate at most a linear increase with age (in the case of a unit root with $\rho = 1$), and will be strictly concave if wage shocks are anything less than permanent ($\rho < 1$).

The autocovariance function can similarly provide evidence of the underlying wage process. The HIP model predicts that the autocovariance between wage levels multiple periods apart should eventually rise, especially at higher ages (again because of a h^2 term in Eq. (1.3)). On the other hand, the RIP model suggests that the covariance function should be strictly declining at a rate that depends directly on ρ.

In order to exploit the information in wage levels, we need sufficient variation in average age across the empirical autocovariance matrix.[31] To achieve this, we divide the sample into four cohorts, as defined in Figure 1.2, with each sub-sample containing a roughly equal number of observations. We retain enough observations to reliably calculate sample moments for each year-year-cohort cell.[32] Our approach departs from studies that pool all cohorts into a single estimation sample, which generates very little variation in average age.[33]

1993).

[29]This discrepancy is found in Heathcote et al. (2010) for the US. Table 2 in Chatterjee et al. (2016) documents the discrepancy for Australia and a number of other countries.

[30]In addition to Eq. 1.6, the model moments in growth rates are $\mathrm{var}(\Delta\hat{y}_{i,h}) = \sigma_\beta^2 + \frac{2\sigma_\eta^2}{1+\rho} + 2\sigma_\nu^2$ and $\mathrm{cov}(\Delta\hat{y}_{i,h}, \Delta\hat{y}_{i,h+1}) = \sigma_\beta^2 - \frac{(1-\rho)\sigma_\eta^2}{1+\rho} - \sigma_\nu^2$.

[31]As noted by Baker and Solon (2003) "this is just an application of the point, familiar from introductory econometrics textbooks,...that aggregation causes inefficiency in parameter estimation when the aggregation discards within-aggregate variation in the relevant variables."

[32]Ideally, we would estimate a TxT covariance matrix for each birth-year cohort, but the small sample and impact of attrition make this approach impractical with survey data. Using a large administrative dataset drawn from Canadian income tax records, Baker and Solon (2003) are able to estimate a much richer model than would be possible here.

[33]Specifically, an unbalanced panel generates only minor variation in age between years, while a balanced panel implies that (average) age rises by exactly 1 each year making it difficult to disentangle age from time effects.

How well does our estimation strategy work? In Chapter 2, we present evidence from Monte Carlo simulations that our strategy produces more accurate GMM estimates of ρ than GMM estimation based on wage growth rates, regardless of whether or not the true model contains heterogeneous growth.[34] We also find that our estimation strategy is more likely to correctly detect the HIP model when the true wages model contains heterogeneous growth.

1.4.3 Visual evidence from sample covariances

Before turning to GMM estimation, we briefly examine the wage data visually for any clues about the underlying wage process. Firstly, Figure 1.1 shows that the age profile of the variance of log wages appears to be strictly concave. This provides preliminary support for an RIP model and a model in which ρ is somewhat less than 1. Similarly, Figure 1.5 plots the covariance function (Cov(y_h,y_{h+n})) for each age beginning with those aged 25. The plotted covariances have been smoothed by age to reduce the noise in the raw covariances and are adjusted for cohort effects.[35] Again, the graph looks generally inconsistent with the HIP model. While there is an occasional up-tick in higher-order autocovariances, this is likely to be the result of fewer observations at higher orders leading to greater sampling error. The other notable feature of Figure 1.5 is that the covariance functions seem to decline almost linearly. Again, this points to a value of ρ close to one.

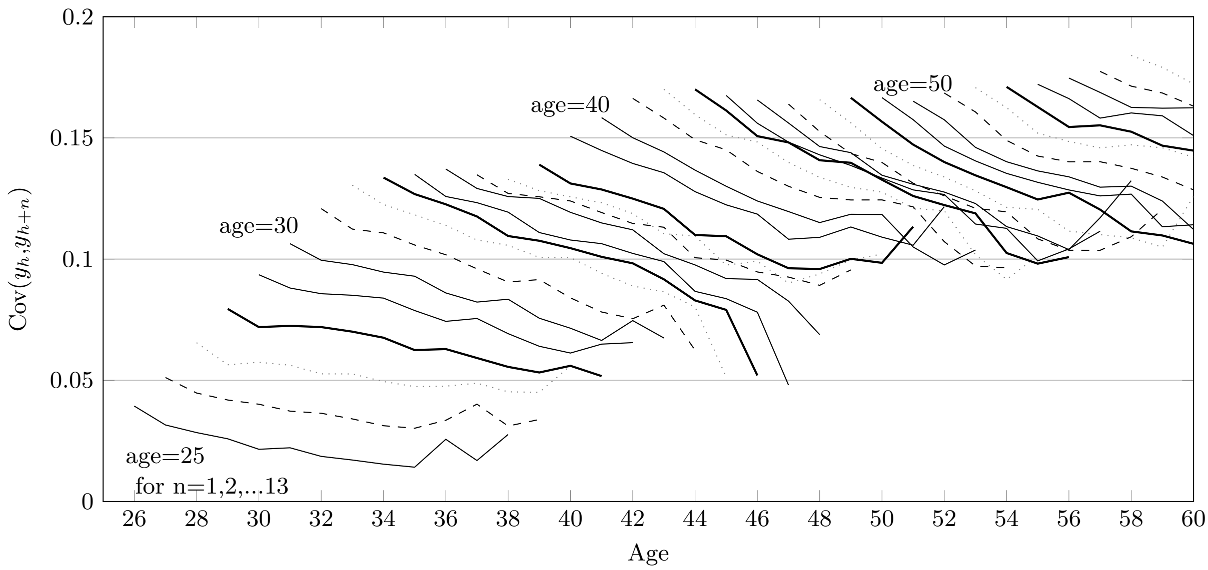

Figure 1.5: Sample autocovariance function by age

[34] This finding extends Hryshko (2012), which only considers the performance of estimation based on growth rates.

[35] Specifically, we follow the approach in Guvenen (2009) by using a rolling 5-year window for age to increase the number of observations contributing to each covariance calculation. The covariances for each age-cohort cell are then regressed on a full set of year-of-birth dummies. The resulting residuals are averaged across cohorts, and it is these covariances that are plotted in Figure 1.5.

1.5 Results

1.5.1 Main results

We use minimum distance methods (or GMM) to estimate the models parameters, details of which are provided in Appendix 1.B. Columns (1) to (3) in Table 1.1 show the results when we allow for heterogeneous growth, while column (4) shows estimates for the restricted (RIP) model. Columns (1), (2) and (4) use moments based on wage levels while the third column estimates the model based on first-differenced wage data. Column (2) uses a diagonal weight matrix for comparative purposes, while all other GMM estimates are based on the identity weight matrix consistent with most other 'covariance' studies.[36]

Beginning with the HIP v RIP question, all three specifications that allow for heterogeneous growth produce estimates of σ^2_β that are small and insignificantly different from zero, *providing no support for the HIP model*.[37] In contrast, estimates of σ^2_α – the variance of the intercept term – are significant and economically large in all levels-based estimations. Overall, this suggests that unobserved heterogeneity is important in explaining the level, but not the profile, of inequality over the life cycle. Unsurprisingly, the estimated covariance between starting wages and wage growth rates ($\sigma_{\alpha\beta}$) is also insignificantly different from zero (see columns (1) and (2)).

For the estimates based on wage levels, the variance of the transitory shock is four to five times that of the persistent shock. In contrast, estimation based on wage growth rates data (column (3)) yields a smaller transitory variance and a much larger persistent variance. These findings, including the discrepancy between levels- and growth-based estimates, are consistent with most past studies including the Australian study by Chatterjee et al. (2016).[38] We attempt to reconcile this discrepancy in the next subsection.[39] In relation to the estimated transitory variance, we again note that a sizeable (but unidentifiable) proportion is likely to be measurement error.

The point estimates for the autoregressive coefficient ρ range from of 0.92 to 0.95, and these are significantly different from one in the levels-based estimations, but insignificant in the growth rate-based estimation (column (3)). Using differenced wage data seems to reduce the precision of the ρ estimates, a proposition that we confirm in our Monte Carlo simulations in Chapter 2. On balance, we interpret these estimates of ρ as moderate evidence against a unit root in wages.[40]

[36]The time effects on the transitory shock and cohort effects on the permanent shock are generally insignificant (against a null hypothesis of 1), so are omitted for brevity.

[37]We experimented with many other levels-based specifications that allowed for non-stationarity through various combinations of factor loadings on the variance terms. All specifications failed to reject $\sigma^2_\beta = 0$ at standard significance levels. These results are available from the authors on request.

[38]See their Table 2 for summary of international findings.

[39]For example, our estimates of the persistent and transitory variances are similar to those in Heathcote et al. (2010) (when they use wage levels) and in Guvenen (2009) when the RIP model is imposed.

[40]The availability of large administrative datasets in a number of countries has allowed more sophisticated ways of testing for a unit root in wages. See for example Gustavsson and Österholm (2014) who examine this question using Swedish administrative data on male wages. They reject the unit root model.

Table 1.1: Minimum distance estimates - test for HIP

	(1) HIP	(2) HIP	(3) HIP	(4) RIP
Fixed effect (σ^2_α)	0.0305*	0.0298*		0.0335**
	(0.0183)	(0.0178)		(0.0042)
Heterogenous growth (σ^2_β)	0.0000	0.0000	0.0002	
	(0.0001)	(0.0000)	(0.0003)	
Correlation ($\sigma_{\alpha\beta}$)	0.0007	0.0005		
	(0.0014)	(0.0011)		
Var. transitory shock (σ^2_ν)	0.0470**	0.0472**	0.0382**	0.0509**
	(0.0024)	(0.0026)	(0.0064)	(0.0012)
Var. persistent shock (σ^2_η)	0.0137**	0.0147**	0.0395**	0.0121**
	(0.0026)	(0.0023)	(0.0100)	(0.0018)
Persistence of AR(1) (ρ)	0.9264*	0.9332**	0.9063**	0.9522**
	(0.0324)	(0.0202)	(0.0380)	(0.0054)
Year effects on σ^2_α - π_{2002}	0.9078	0.9121		0.9204*
	(0.0620)	(0.0578)		(0.0323)
π_{2003}	0.9100	0.8375*		0.9342
	(0.0565)	(0.0714)		(0.0387)
π_{2004}	0.8534*	0.7672*		0.8987*
	(0.0697)	(0.0995)		(0.0471)
π_{2005}	1.0426	0.9748		1.0745
	(0.0853)	(0.0578)		(0.0429)
π_{2006}	0.9789	0.7013*		1.0323
	(0.0728)	(0.1284)		(0.0489)
π_{2007}	1.0172	0.8389		1.0749
	(0.0972)	(0.0927)		(0.0559)
π_{2008}	1.0059	0.8057		1.0880
	(0.1074)	(0.1108)		(0.0599)
π_{2009}	1.0399	0.7847		1.1027
	(0.1140)	(0.1150)		(0.0614)
π_{2010}	1.1428	0.9602		1.1904**
	(0.1614)	(0.0949)		(0.0604)
π_{2011}	1.0969	0.8806		1.1881**
	(0.1432)	(0.0977)		(0.0644)
π_{2012}	1.1503	1.1384		1.2469**
	(0.1622)	(0.1560)		(0.0659)
π_{2013}	1.1044	1.0286		1.2050**
	(0.1523)	(0.1136)		(0.0670)
Data	Levels	Levels	First Diff.	Levels
GMM weight matrix	Identity	Diagonal	Identity	Identity

*significant at 5 per cent; **significant at 1 per cent; one-sided test against null hypothesis of zero for variance parameters and one for ρ and factor loadings.
Standard errors in parentheses.

While insignificant (compared to a null of 1) in some years and models, the estimates of the time effect loadings on σ^2_α in columns (1), (2) and (4) point to a general upward trend in wage inequality, which we discuss below.[41] Finally, we find that the key estimates are very similar when we use the identity and diagnoal weight matrix (column (1) versus column (2)).

1.5.2 Discrepancy between wage level and growth estimates

The large discrepancy in our estimates of the persistent variance when using wage levels versus growth rates is similar to the finding in Chatterjee et al., and is a well-known puzzle in the international literature (see Table 2 in Chatterjee et al. for cross-country results). A recent study by Manovskii et al. (2015) exploring the reasons for the discrepancy finds that a large part is caused by ‘rare’ transitory shocks to wages at the beginning and end of spells within the panel.[42] They find that removing these

[41] The absolute value of the time factor loadings, and therefore the significance in particular years, is sensitive to the choice of numeraire. In any case, we strongly reject the hypothesis that $\pi_t = 1$ over the whole period, a test that is invariant to the choice of numeraire.

[42] In effect, these transitory shocks are misclassified as persistent when estimation is based on wage growth rates.

noisy observations and relying only on 'interior' observations brings the two sets of estimates into closer alignment.

We investigate whether this factor can account for the discrepancy in our estimates. To do so, we re-estimate the RIP model after dropping the first and last observations in an individual's earnings spell as well as improbably large annual increases (> 400 per cent) and decreases (< 90 per cent) in hourly wages.[43] Table 1.2, which compares estimates using the 'full' and 'interior' samples, shows that the interior sample leads to much closer estimates in levels and growth. In particular, the discrepancy in the estimates of the persistent variance σ_η^2 is largely eliminated. The estimates of ρ are also brought into closer alignment.

Table 1.2: Minimum distance estimates - interior spells only

	(1) Full	(2) Full	(3) Interior	(4) Interior
Var. transitory shock (σ_ν^2)	0.0504**	0.0384**	0.0341**	0.0212**
	(0.0012)	(0.0063)	(0.0009)	(0.0023)
Var. persistent shock (σ_η^2)	0.0124**	0.0398**	0.0131**	0.0157**
	(0.0021)	(0.0100)	(0.0018)	(0.0039)
Persistence of AR(1) (ρ)	0.9483**	0.9242*	0.9528**	0.9422
	(0.0060)	(0.0416)	(0.0061)	(0.0410)
Data	Levels	First Diff.	Levels	First Diff.
GMM weight matrix	Identity	Identity	Identity	Identity

*significant at 5 per cent; **significant at 1 per cent; one-sided test against null hypotheses of zero for variance parameters and one for ρ.
Standard errors in parentheses.

1.5.3 Components of life-cycle wage inequality

Having rejected the HIP model, we now use our estimates of the RIP model to decompose life-cycle wage inequality into its component parts. In doing so, we evaluate σ_α^2 at the average value of the time factor loadings. This gives the following empirical equation for wage inequality at age h:

$$\mathrm{var}(\hat{y}_{i,h}) = \underbrace{\sigma_\alpha^2}_{0.05} + \underbrace{\rho^2}_{0.95} \mathrm{var}(z_{i,h-1}) + \underbrace{\sigma_\eta^2}_{0.012} + \underbrace{\sigma_\nu^2}_{0.05} \tag{1.7}$$

where $\mathrm{var}(z_{i,h-1})$ captures the accumulation of shocks since the individual entered the labour market at age 23. This decomposition is shown in Figure 1.6, and suggests that, from about age 45, around one quarter of wage inequality can be attributed to variability in workers' starting wages. A further half can be attributed to the accumulation of persistent wage shocks, while the remaining quarter is attributable to purely transitory fluctuations. In other words, the majority (about three quarters) of wage inequality from the mid-20s onwards is the result of random factors as opposed to unobserved heterogeneity. While there are no directly comparable Australian studies, the US study by Guvenen (2009), which finds support for the HIP model, implies a very different life-cycle decomposition: by age 55, worker heterogeneity accounts for almost 75 per cent of wages inequality in the US (cf. Huggett et

[43] This has the effect of truncating the sample to 2002 to 2013, and reduces the sample to 28,584 observations. Given the smaller sample, we choose to omit any time or cohort factor loadings in the interior estimations.

al. (2011); Storesletten et al. (2004))

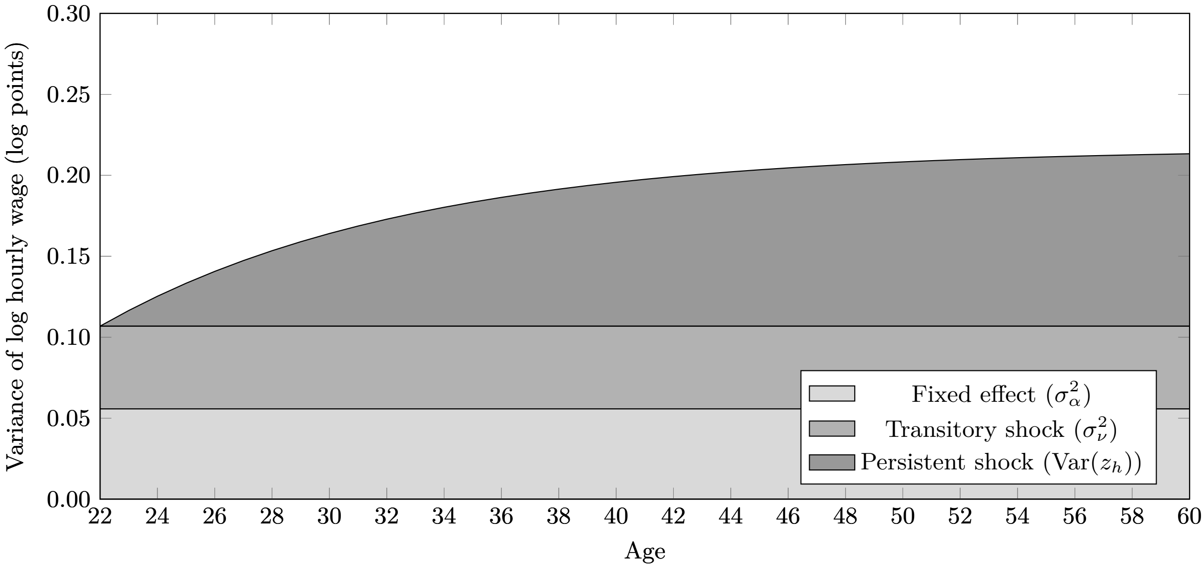

Figure 1.6: Decomposition of life-cycle wage inequality

1.5.4 Trends in wage inequality and mobility over time

In our model, the upward trend in wage inequality seems to be well captured by the rising estimates of π_t over the sample period, as shown in Figure 1.3.[44] As mentioned earlier, a rise in π_t can be interpreted as an increased return to unobserved skills and talents. While this finding may seem at odds with the numerous Australian studies that have found no increase in the return to education over this period (for example Coelli and Wilkins (2009) and Chatterjee et al. (2016)), these studies have focused on returns to *observed* differences in education. On the other hand, movements in π_t capture changes in the return to *unobserved* differences in human capital and other worker attributes, which may well have risen over this period.[45] The increases in π_t also imply that the upward trend in observed wage inequality has been driven mainly by greater inequality in long-run wages rather than increases in short-term (transitory) wage inequality.[46] This is consistent with a reduction in earnings mobility over this period (Baker and Solon, 2003).[47] While identifying the underlying causes of this trend is beyond the scope of this paper, it is worth noting that increased wages inequality is a common feature in advanced economies over the past 20 years.[48] It is also worth emphasising that this up-trend in wage inequality among working-age men does not seem to extend to measures of inequality at the household-level in Australia (Wilkins, 2015, 2016).

[44] A similar fit is obtained if time factor loadings are attached to $\eta_{i,h}$ instead of α_i.

[45] In fact, Baker and Solon (2003) make a similar finding in relation to Canadian male workers.

[46] When factor loadings on the transitory shock were included in the model, they showed a small downward trend over the sample period.

[47] As Baker and Solon note, "an increase in π_t preserves the order of individuals in the earnings distribution, but spreads them out further, and this greater spread persists from year to year."

[48] Empirical studies have linked this upward trend to numerous factors including the globalisation of e-commerce and broader forms of skills-biased technical change (see for example Dabla-Norris et al. (2015)).

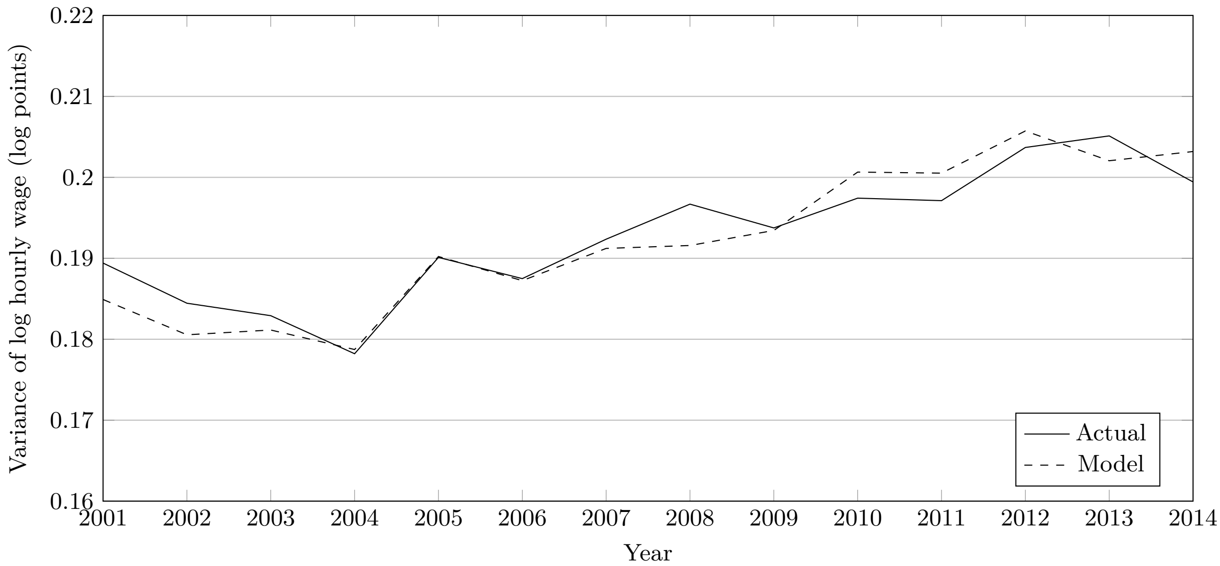

Figure 1.7: Model fit

1.5.5 Comparison with US and other overseas studies

As noted at the outset, the HIP v RIP question remains unsettled in the US where the bulk of studies have taken place. However, given we have closely followed the analytical approach in Guvenen (2009) – which found strong support for the HIP model – it would be interesting to investigate some possible reasons for our contrasting findings.

To begin with, the life-cycle features of the US wages data appear very different to those in Australia. Figure 1.8 plots the age-variance profile for the US and Australian working age men, where the US data is taken directly from Guvenen.[49] Besides exhibiting a much larger rise in wage inequality over the life cycle, the US age profile is broadly convex particularly among college-educated workers.[50] This is only consistent with the HIP model, which predicts that inequality should rise at an increasing rate as sustained differences in growth rates between workers accumulate. This picture is confirmed when we re-apply our estimation methodology to the US dataset. Specifically, we find significant estimates of σ_β^2 for both the entire and college-only samples, with our numerical results broadly consistent with those in Guvenen.[51]

In short, our contrasting finding with Guvenen appears to be the result of differences in the two countries' underlying wage process rather than merely differences in methodology. So why might Australia and the US have different processes for individual wages? A standard theoretical justification for the HIP model is that workers who are more talented or make larger investments in human capital when young should enjoy a premium in wage growth over their careers (Baker, 1997). This of course relies on the ability of

[49] To aid comparability, both lines in Figure 1.8 are based on earnings (ie, wages and salaries) after controlling for cohort effects through dummy variable regressions. Controlling for time effects paints a similar picture. The US data cover the period from 1968 to 1993, while the Australian data cover 2001 to 2014.

[50] Guvenen defines college educated' as at least 16 years of formal education.

[51] Our estimate of σ_β^2 of 0.0002 for the whole sample was somewhat smaller than Guvenen's estimate of 0.00039, but was very similar to Guvenen's estimate of 0.00049 for the college-only sample.

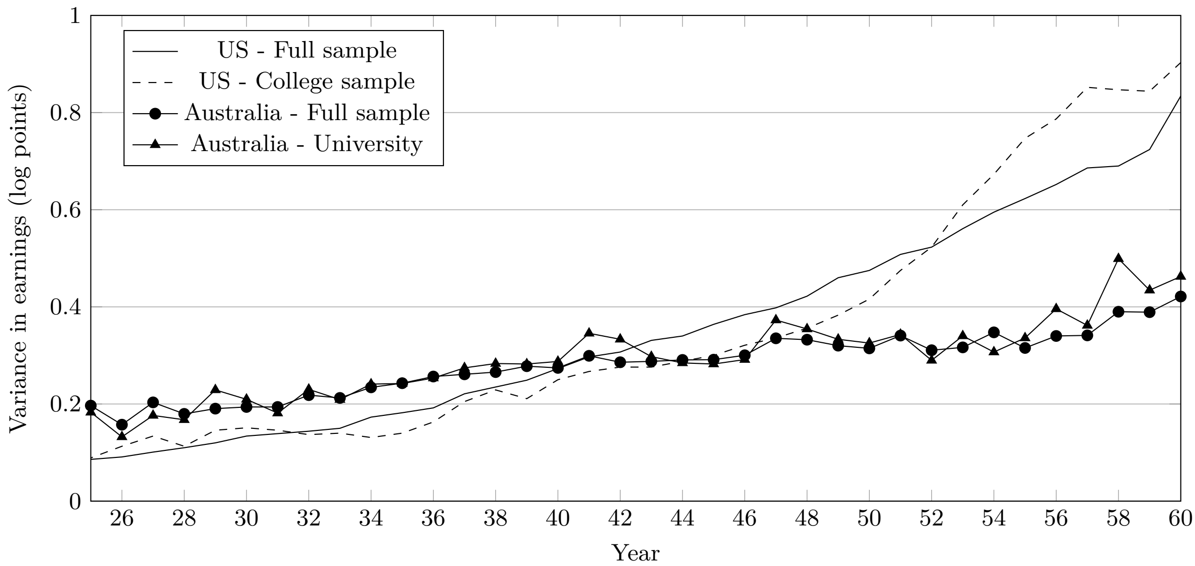

Figure 1.8: Model fit
Source: HILDA Waves 1 to 14, ABS and Guvenen (2009)

firms to detect and reward those workers with higher productivity. There is evidence that this may occur more readily in the US where the labour market is considered to be more competitive and there are less onerous hiring and firing rules. For example, Australia (28th) ranks well below the US (3rd) on the labour market efficiency component of the Global Competitiveness Index (Schwab and Sala-i Martin, 2017), and has had consistently more rigid employment protection legislation than the US, based on OECD analysis (OECD, 2013).

To explore this question in a more rigorous manner would require detailed cross-country evidence. Unfortunately, there are relatively few studies outside the US that have specifically tested the HIP model, and there is no obvious pattern among the few studies that have. For example, Baker and Solon (2003) in Canada and Hoffmann (2013) in Germany both examine male earnings using large administrative datasets, with the former study finding strong support for heterogeneous growth but the latter finding little evidence. In the United Kingdom, Hoffmann (2013) rejects the HIP model using data from the New Earnings Survey. Clearly, further cross-country research will be needed to understand whether labour market features have an influence on individual wage dynamics.

Conclusion

This paper has sought to uncover the key parameters underpinning life-cycle wage dynamics in Australia. Using 14 years of panel data, our analysis suggests that individual wages can be well summarised as the sum of three components: 1) a fixed effect in levels driven by unobserved heterogeneity; 2) a process consisting of highly persistent wage shocks; and 3) purely transitory wage shocks that include any measurement error in the data.

The lack of evidence in HILDA for heterogeneous wage growth rates among male workers is a provocative finding, and it would be worth confronting our results with alternative sources of data such as individual tax records. Also, while we have found no evidence of growth heterogeneity in either the overall labour market or within broad education groupings (see Appendix 1.C), this does not preclude growth heterogeneity being an important factor within particular occupations or industries.[52]

Besides shedding light on the observed trends in wage inequality, our results can better inform the way that heterogeneity and income uncertainty are incorporated into distributional and related policy analysis in Australia. This is particularly the case for heterogeneous agent models, which depend on a realistic representation of wage uncertainty at the individual and household level. Such models are becoming essential tools in the quantitative analysis of government policies and their distributional impact.[53]

[52]Indeed the seminal study by Lillard and Weiss (1979) examined heterogeneous wage growth rates among US scientists.
[53]See Tran and Woodland (2011) for a recent Australian example.

Appendix

Appendix 1.A Data and sample selection

Table 1.A.1: Key variables

Variable	Description	HILDA code	Mean
Earnings ($ per week)	Weekly gross wages + salary all jobs	nwscei	1334
Hours	Hours per week usually worked all jobs	njbhruc	45
Real hourly wage ($ per week)	Wages divided by hours deflated by CPI	derived	33
Age	Age last birthday at June 30	nhgage	40
Weight	Responding person sample weight	hhwtrps	
Region dummy	ASGC 2001 Section of State	hhsos	0.13
NESB dummy	Derived from Country of birth	ancob	0.11
Married dummy	Marital status from person questionnaire	mrcurr	0.59

Table 1.A.2: Sample selection

Removal criteria	Observations dropped	Observations remaining
Initial unbalanced panel		270,942
Missing wages and hours data	156,652	114,290
Implausible wage and hours	10,702	103,588
Not female	49,115	54,473
Missing demographic data	868	53,605
Not in at least two consecutive waves	1,617	46,152
Not working age	6,576	39,576

Appendix 1.B Estimation details

We follow many other studies in this area by using minimum distance methods (or GMM) to estimate the model's parameters. We compute a T x T covariance matrix weighted by HILDA's cross-sectional person weights for each of the four cohorts, and stack the resulting 420 unique moments into a column vector $\mathbf{m}$.[54] Likewise, the 420 model covariances, which are a function of the model's parameter vector θ are stacked into a column vector $\Omega(\theta)$.

The problem is to choose a parameter vector θ that minimises the distance between the moments implied by the model, $\Omega(\theta)$, and their empirical counterparts $\mathbf{m}$:

$$\min_{\theta}(\mathbf{m} - \Omega(\theta))' A(\mathbf{m} - \Omega(\theta))$$

where A is a positive semi-definite weight matrix, and the sample variance of moment conditions is given by V. We follow most other studies by using the identity weight matrix, which Altonji and Segal (1996) show is superior to the optimal weight matrix V^{-1} when estimating models of covariance structures in short panels.[55] We also show results in the main text when the model is re-estimated using the inverse of the main diagonal of V as the weight matrix as in Blundell et al. (2008). Subject to standard regularity conditions (see Chamberlain (1984)), the estimator has an asymptotically normal distribution given by:

$$\sqrt{N}(\hat{\theta} - \theta) \sim \mathcal{N}(0, W)$$

where the asymptotic variance W is:

$$W = (D'AD)^{-1}D'AVAD(D'AD)^{-1}$$

and D is the Jacobian matrix evaluated at $c = \theta$:

$$D = E\left(\frac{\partial\Omega(\theta)}{\partial c'}\right)|_{c=\theta}$$

Appendix 1.C Sensitivity analysis

This Appendix shows that the inability to reject the RIP model is robust to: using work experience instead of age; earnings instead of hourly wages; restricting the sample to tertiary educated workers; using household level earnings instead of individual wages; and imposing a unit root (columns (1) to (5) in Table 5.6.2 below). We now discuss each of these in turn.[56]

[54] There are fewer moments when the model is estimated using wage growth rates, because one year is lost in the differencing process.

[55] In particular, Altonji and Segal show that using (V^{-1}) in short panels biases parameter estimates towards zero and that this problem is most severe when the V matrix is large, as is the case here. There is also no guarantee that the variance of the sample moment conditions will be positive semi-definite matrix when working with unbalanced panel data.

[56] We undertook a range of other sensitivity checks, including testing whether the main results apply to women and whether our use of sampling weights in forming empirical moments affected the results. In all cases, we could not reject the

Experience instead of age

Our main results use age (less 22) as a proxy for labour market experience, rather than a more direct measure.[57] Conveniently, the HILDA survey includes a more direct measure of labour market experience (coded as *ehtjb*) that is constructed based on each respondent's work history. After re-estimating the model using this variable in place of age, the estimated variance of the fixed effect σ_α^2 is somewhat lower and the estimated autocorrelation coefficient is lower at about 0.87. However, other estimates are largely unchanged with the heterogeneous growth term remaining insignificant.

Earnings instead of hourly wage rate

In testing for heterogeneous wage growth rates, we have focused on hourly wages. However, as some other studies have chosen to use earnings (or equivalently 'labour income') in testing the HIP model, we re-estimate our model using HILDA's current weekly gross wages and salaries (coded *wscei*) instead. Following other studies (for example, Chatterjee et al. (2016)), we confine the sample to full time workers only (as defined by the variable *esdtl*) to avoid the volatility in hours created by older workers transitioning into retirement. Again, the results show no evidence of heterogeneous growth. The estimated variance of the persistent wage shock is slightly higher. Other estimates are largely unchanged, including the values of the time factor loadings π_t (omitted for brevity in Table 5.6.2).

Dividing sample by educational attainment

Guvenen (2009) finds a greater degree of wage growth rate heterogeneity in his college educated sub-sample. We therefore explore whether there is support for the HIP model among tertiary educated workers in Australia.[58] When Model 3 is re-estimated on tertiary-educated male workers only, the main results are largely unchanged. Again, the heterogeneous growth term remains insignificant.

Household instead of individual-level earnings

Whether wage growth rate heterogeneity at the individual level – if it exists – should carry over to household-level income is unclear. Nonetheless, we re-estimate the model using household labour income (or earnings).[59] To avoid double-counting, the sample now consists only of household heads, defined as the oldest male of each household, or oldest female where no male is present. The estimated variance of the transitory wage shock is now higher, as is that of the persistence shock. This is not surprising, as the overall variance in household earnings is on average about 80 per cent higher than that of individual earnings in the sample period. Again, the heterogeneous growth term remains insignificant.

RIP model.

[57] Most other studies also use a person's age or construct an indirect measure of experience by subtracting total years of schooling from age, as in Guvenen (2009).

[58] We categorise individuals based on their highest qualification over the whole panel.

[59] This uses the HILDA variable *hiwsfei*.

Imposing unit root assumption

Finally, we re-estimate the model imposing $\rho = 1$, noting the difficulties associated with identifying a unit root in short panel contexts. This results in a significantly smaller estimated variance for the persistent shock, bringing it closer to the estimates of Chatterjee et al. (2016) who also impose a unit root.[60] However, the heterogeneous growth term remains insignificant.

Table 1.C.1: Minimum distance estimates - sensitivity analysis

	Exper.	Earnings	Tertiary	HH level	$\rho = 1$
Fixed effect (σ^2_α)	0.0868**	0.0515**	0.0271	0.0781	0.0762**
	(0.0055)	(0.0208)	(0.0200)	(0.0487)	(0.0145)
Heterogenous growth (σ^2_β)	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0000)
Correlation ($\sigma_{\alpha\beta}$)	0.0009	-0.0007	0.0004	-0.0021	-0.0003
	(0.0007)	(0.0014)	(0.0015)	(0.0035)	(0.0008)
Var. transitory shock (σ^2_ν)	0.0361**	0.0296**	0.0333**	0.1507**	0.0636**
	(0.0015)	(0.0018)	(0.0026)	(0.0093)	(0.0021)
Var. persistent shock (σ^2_η)	0.0141**	0.0151**	0.0153**	0.0278**	0.0033**
	(0.0024)	(0.0021)	(0.0028)	(0.0051)	(0.0003)
Persistence of AR(1) (ρ)	0.8687**	0.9604*	0.9347**	0.9528*	
	(0.0207)	(0.0180)	(0.0254)	(0.0246)	
Data	Levels	Levels	Levels	Levels	Levels
GMM weight matrix	Identity	Identity	Identity	Identity	Identity

*significant at 5 per cent; **significant at 1 per cent; one-sided test against null hypotheses of zero for variance parameters and one for ρ.
Standard errors in parentheses.

[60] Imposing a unit root also results in a substantially poorer fit, with a general Wald test rejecting the restriction.

Chapter 2

GMM estimation of dynamic wage models: levels or growth rates?

2.1 Introduction

There is a long history of using panel data to estimate dynamic wage models (e.g. Lillard and Weiss (1979), MaCurdy (1982) and Abowd and Card (1989)). Out of this literature have emerged two main empirical controversies: firstly, does the stochastic component of individual wages contain a unit root? And second, does the deterministic trend in individual wages vary between workers? The latter question is sometimes known as the heterogeneous income profiles (HIP) versus restricted income profiles (RIP) debate (Guvenen, 2009). Guvenen (2009) demonstrates that distinguishing between these various models using generalised method of moments (GMM) estimation is very difficult, especially if the true model contains a highly persistent stochastic process. This is because the pattern of wage covariances generated by a model with a highly persistent shock that does *not* contain heterogeneous growth looks very similar to the pattern generated by a model with a less persistent shock that *does* contain heterogeneous growth. This is exacerbated by the inherent noisiness of the data in short unbalanced panels drawn from household surveys.

In confronting this difficulty, the researcher faces an important choice in estimation: should the model be estimated in wage levels or wage growth? While a roughly equal number of empirical studies have opted for each of these alternatives,[1] there is very little econometric evidence on their relative performance. In a recent Monte Carlo study, Hryshko (2012) argues that GMM estimation using wages growth can accurately recover the true value of all model parameters across a wide range of specifications and parameter values. However, his results also show that this estimation approach struggles to correctly detect heterogeneous growth (i.e. the HIP model) when the true model contains a highly persistent but

[1] Meghir and Pistaferri (2011) provide a survey of the empirical literature from the United States.

not quite unit root stochastic process, that is, an auto-correlation coefficient of around 0.95. Moreover, Hryshko (2012) makes no attempt to compare the performance of growth-based with levels-based GMM estimation.

In this Monte Carlo study, I systematically explore the relative performance of levels versus growth rates in GMM estimation of dynamic wage models. I also explore a third variation: using wage levels but dividing the sample into year of birth groupings rather than simply pooling all cohorts into one sample. I focus on the particularly difficult case where the true model contains a stochastic process that is almost a unit root. My results show that estimation based on wage levels divided into cohort groupings is the most accurate means of detecting the true degree of persistence in the stochastic wage process and correctly rejecting the null hypothesis of a unit root. This finding is an important contribution to the empirical wages literature as testing for a unit root is a notoriously difficult task in both time series and panel data settings (see Gustavsson and Österholm (2014)). I also find that wage levels divided into cohorts is the most effective method for correctly detecting the presence of heterogeneous growth. The intuition for why the cohort approach works best is simple: wage levels data contain valuable information that is lost when the data is first differenced. At the same time, dividing the sample into cohorts retains substantial variation in age between cells in the sample covariance matrix, which would not exist if all cohorts were pooled into a single group.[2] This information is useful in identifying the age-related trend in wages that would be generated by a HIP model (Baker and Solon (2003) and Guvenen (2009)). Overall, this study's findings provide important guidance to practitioners who wish to estimate dynamic wage models using short unbalanced panels.

2.2 Methodology

I assess the relative accuracy of GMM estimates based on three different estimation strategies: wage levels dividing the sample into four cohorts (as used in Chapter 1); wage growth rates; and wage levels pooling all cohorts together.[3] The comparison is based on a synthetic panel of 2000 working-age individuals who are each randomly allocated one of 50 birth-year cohorts. The sample window is chosen to be 16 years, consistent with a realistic panel length.[4] The log wage ($\hat{y}_{i,h}$) for individual i of age h is generated from the following encompassing model:

$$\hat{y}_{i,h} = \overbrace{\overbrace{\underbrace{\alpha_i}_{Ind.\ fixed\ effect} + \underbrace{\beta_i h}_{Ind.\ growth}}^{\text{Ind. heterogeneity}} + \overbrace{\underbrace{z_{i,h}}_{Persistent(AR(1))} + \underbrace{\nu_{i,h}}_{Transitory}}^{\text{Wage uncertainty}}}^{\text{Unobserved part of wages}} \tag{2.1}$$

[2]In an unbalanced panel, using a single sample results in almost no variation in mean age between sample covariance terms.

[3]I focus on four cohort groupings because this is probably the maximum number that would be viable for panel surveys with typical sample sizes. Any more than four would result in too few observations for calculating each cohort's sample covariance matrix.

[4]Given individuals are randomly allocated across 50 birth years, restricting the sample window to 16 years automatically generates an unbalanced panel.

where the variances of the individual-specific intercept and slopes are σ_α^2 and σ_β^2, $z_{i,h} = \rho z_{i,h-1} + \eta_{i,h}$ is the stochastic component of wages with persistence parameter ρ. The persistent and transitory shocks, $\eta_{i,h}$ are $\nu_{i,h}$, are independent draws from a mean zero normal distribution with variance σ_η^2 and σ_ν^2. The values of α_i and β_i are drawn independently from mean-zero normal distributions with variances σ_α^2 and σ_β^2.

I simulate wage profiles for $\sigma_\beta^2 = \{0, 0.0005\}$ and $\rho = \{0.95, 1\}$ giving four different combinations of parameter values, and allowing us to investigate how each estimation strategy performs when the true model is RIP or HIP, and when the true process does or does not contain a unit root. The other parameter values are the same across simulations and given by $\sigma_\alpha^2 = 0.04$, $\sigma_\nu^2 = 0.04$ and $\sigma_\eta^2 = 0.015$. The covariance between α_i and β_i (σ_{α_β}) is set to zero. Tables 2.3.1 to 2.3.4 show the means and standard deviations/errors (in parentheses) of parameter estimates over 100 simulations, and the rejection rates for $\sigma_\beta^2 = 0$ and $\rho = 1$ in each case.[5]

2.3 Results

The main findings are threefold. First, the approach of dividing the sample into cohort groupings produces estimates of ρ with the smallest bias in all four cases. This superiority is most obvious when the true wage process is governed by the RIP model, but it remains the most accurate method when the true wage process contains heterogeneous growth. Further, this method is relatively effective at correctly rejecting the null hypothesis ($\approx$ 50 per cent of the time) of a unit root even when the true value of ρ is close to, but still less than, one. This is notoriously difficult in short panel contexts (Gustavsson and Österholm, 2014). When the true model contains a unit root (i.e. $\rho = 1$), the wage levels by cohort and wage growth rate approaches both produce estimates of ρ that are downward-biased.[6][7] Nonetheless, both these methods (correctly) failed to reject the unit root the vast majority of the time ($\approx$ 85 per cent). *Overall, the results suggest that estimation using wage levels and cohort groupings is the most robust way of testing for a unit root in panel data.*

Second, estimation based on wage levels by cohort and wage growth rates both provided relatively unbiased estimates of σ_β^2. However, estimation based on wage levels divided into cohort groupings again outperforms estimation based on wage growth rates in its ability to correctly reject $\sigma_\beta^2 = 0$ when the true model contains heterogeneous growth. However, I find as Hryshko (2012) does that estimates of σ_β^2 are biased upwards when the true model contains a unit root. *Taken together, these results suggest that failing to reject $\sigma_\beta^2 = 0$ in estimation based on wage levels and cohort groupings provides strong evidence against heterogeneous growth.*

Lastly, the approach of using wage levels *without* dividing into cohorts is generally poor at recovering

[5] Rejection rules are based on a one-sided test and 5 per cent significance level.

[6] This is also a well-known problem in time series contexts. It also confirms the finding in Hryshko (2012), although that study only investigated estimation based on wage growth rates.

[7] For the unit root cases, I only compare wage levels divided by cohort and growth rates as GMM estimation using wage levels with pooled cohorts often failed to converge.

the model's parameters. It badly overstates the value of σ^2_α under both the RIP and HIP models, and overstates σ^2_β under the RIP model. Its estimates of ρ also exhibit a large downward bias.

Table 2.3.1: GMM estimates when true model is RIP ($\sigma^2_\beta = 0$) with no unit root ($\rho = 0.95$)

	Parameter estimates						Rejection rate (%)	
	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	σ^2_ν	σ^2_η	ρ	$\sigma^2_\beta = 0$	$\rho = 1$
Wage levels by cohort	0.043	0.00001	-0.00037	0.039	0.015	0.949	6	56
	(0.0092)	(0.00002)	(0.0007)	(0.0021)	(0.0016)	(0.0226)		
Wage growth		0.00010		0.040	0.016	0.918	6	19
		(0.00009)		(0.001)	(0.0019)	(0.0311)		
Wage levels pooled	0.228	0.00049	-0.00794	0.038	0.015	0.836	67	87
	(0.1568)	(0.00018)	(0.00502)	(0.0017)	(0.0025)	(0.0716)		

Table 2.3.2: GMM estimates when true model is HIP ($\sigma^2_\beta = 0.0005$) with no unit root ($\rho = 0.95$)

	Parameter estimates						Rejection rate (%)	
	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	σ^2_ν	σ^2_η	ρ	$\sigma^2_\beta = 0$	$\rho = 1$
Wage levels by cohort	0.042	0.00045	-0.00023	0.038	0.017	0.941	71	47
	(0.0282)	(0.00022)	(0.00179)	(0.0039)	(0.0036)	(0.043)		
Wage growth		0.00049		0.040	0.016	0.937	34	19
		(0.00018)		(0.001)	(0.0019)	(0.0419)		
Wage levels pooled	0.384	0.00122	-0.01361	0.037	0.022	0.854	73	55
	(0.3576)	(0.0004)	(0.01127)	(0.0022)	(0.0032)	(0.067)		

Table 2.3.3: GMM estimates when true model is RIP ($\sigma^2_\beta = 0$) with unit root ($\rho = 1$)

	Parameter estimates						Rejection rate (%)	
	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	σ^2_ν	σ^2_η	ρ	$\sigma^2_\beta = 0$	$\rho = 1$
Wage levels by cohort	0.028	0.00014	0.00112	0.037	0.017	0.966	39	15
	(0.022)	(0.00015)	(0.00154)	(0.0031)	(0.0025)	(0.0313)		
Wage growth		0.00033		0.040	0.016	0.947	20	14
		(0.00016)		(0.001)	(0.0018)	(0.0375)		

Table 2.3.4: GMM estimates when true model is HIP ($\sigma^2_\beta = 0.0005$) with unit root ($\rho = 1$)

	Parameter estimates						Rejection rate	
	σ^2_α	σ^2_β	σ_{α_β}	σ^2_ν	σ^2_η	ρ	$\sigma^2_\beta = 0$	$\rho = 1$
Wage levels by cohort	0.036	0.00064	0.00085	0.035	0.018	0.953	60	21
	(0.0378)	(0.00031)	(0.00257)	(0.0048)	(0.0047)	(0.0469)		
Wage growth		0.00083		0.040	0.016	0.950	38	15
		(0.00018)		(0.001)	(0.0018)	(0.039)		

Chapter 3

Risk aversion among Australian households

3.1 Introduction

Risk aversion is central to individual decision-making under uncertainty, a theory which underpins standard economic models of inter-temporal consumption and saving behaviour, portfolio choice and labour supply to name a few. Risk aversion has also been identified as an important factor in many other areas of economic inquiry including saving rates (Aiyagari, 1994; Carroll and Samwick, 1998), economic growth (Devereux and Smith, 1994), wage inequality (Caroli and Garcia-Penalosa, 2002) and occupational choice (Bonin et al., 2007).

The degree of relative risk aversion, for which the standard measure is the coefficient of relative risk aversion (γ), also plays a crucial role in many areas of policy analysis. For example, the value of γ is a key determinant of the optimal structure of the tax system (see for example Imrohoroglu (1998) and Conesa et al. (2009)) and is important in regulating the economy's response to changes in monetary and fiscal policy (see for example King and Rebelo (1999)).

While a wide range of international studies, mostly from the U.S., have attempted to estimate γ, this has proven to be a difficult task in practice and substantial disagreement over the value of γ remains.[1] In any event, there is little empirical evidence on the value of γ for Australia.[2] As it seems possible that Australian households might differ substantially from their U.S. counterparts in their attitude towards risk and therefore their inter-temporal consumption and portfolio choices, it is important to estimate this parameter for Australia. This will be particularly useful for researchers who build large-scale models

[1] See Guvenen (2006) for an attempt to reconcile the conflicting empirical estimates.

[2] As noted below, the only Australian estimates of γ we could find are in the cross-country study by Gandelman and Hernandez-Murillo (2015) and the recent simulation study by Iskhakov and Keane (2018a).

of the Australian economy for the purposes of policy advice and scenario analysis. This is our main contribution in this paper.

We provide Euler equation-based estimates of γ from the constant relative risk aversion (CRRA) utility function for Australia using longitudinal consumption data from the Household Income and Labour Dynamics in Australia (HILDA) survey and data on aggregate interest rates. These are the first Euler equation estimates using micro data in Australia as far as we are aware.

To preview the results, our preferred non-linear specifications that allow for the well-known and substantial problem of measurement error suggest γ lies in the range of 1.2 to 1.4 for the representative household. To put this in context, for the extreme case that $\gamma = 0$, the utility function is linear and the agent exhibits no risk aversion – a situation we soundly reject. On the other hand, our estimates of γ are not significantly higher than unity, and thus we cannot reject the logarithmic utility assumption that is often adopted in macroeconomic policy analysis. Overall, our estimates of γ therefore imply a moderate degree of risk aversion.

While there are no other Euler equation-based estimates for Australia, Gandelman and Hernandez-Murillo (2015) found a value of 1.17 for Australia in their cross-country study based on subjective measures of risk aversion. Our estimates of γ are also similar to the value of 1.5 found in the US study by Alan et al. (2009) whose approach we generally follow. In models that correct for measurement error we also estimate the household's discount factor β and find values of 0.92 and 0.96. These values are consistent with the bulk of empirical evidence internationally (including Alan et al. (2009)).

The CRRA utility function and our estimation approach embed the assumption that relative risk aversion is constant with respect to a household's level of wealth. Previous Australian studies have found mixed evidence regarding this assumption. We also test this assumption and we are unable to reject the assumption of constant relative risk aversion. We also attempt to explain the variety of findings from previous Australian studies.

In the next section, we provide some background about the research agenda around estimating risk aversion. Section 3.3 describes the data we use and Section 3.4 details our tests of the assumption that relative risk aversion is constant in wealth. Having established some empirical support for the CRRA assumption, section 3.5 presents our γ (and other parameter) estimates from various Euler equation models derived from the CRRA assumption. Section 3.6 provides some conclusions and areas where further research would be useful in the Australian context.

3.2 Background

Given the fundamental interest around risk aversion and its importance in economic modeling, an extensive body of empirical research attempts to quantify the degree of risk aversion among households.

Although there are a number of possible ways to proceed,[3] this paper follows the 'Euler equation' approach pioneered in Hall (1978). In short, this approach exploits the relationship between consumption and interest rates derived from the household's inter-temporal optimisation problem to uncover the risk aversion parameter in the household's underlying utility function.

To operationalise the Euler equation approach, a particular form of preferences must be assumed, and we opt for the constant relative risk aversion (CRRA) utility function $U(C) = \frac{C}{1-\gamma}^{1-\gamma}$, where C is consumption and γ is the coefficient of relative risk aversion. A particular advantage of CRRA preferences is that they imply an empirically tractable Euler equation, namely:

$$\mathrm{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + r_{t+1})\beta \right] = 1 \tag{3.1}$$

where r_t is the return on a generic asset at time t, β is the household's discount factor and the expectation operator E_t indicates that the household's consumption decision takes place in an environment where the value of future variables (such as earnings) is uncertain.

CRRA preferences have a number of other desirable properties, which have made them a popular specification in models of inter-temporal consumption decisions under uncertainty. This includes the convenient feature that the standard (Arrow-Pratt) measure of relative risk aversion is constant in wealth and given directly by γ,[4] while the elasticity of inter-temporal substitution (EIS) is also constant and equal to $1/\gamma$.[5] CRRA preferences also have a number of economically appealing aspects, including the presence of 'prudence', which gives rise to a precautionary (or 'buffer stock') savings motive when households face future income uncertainty.[6]

The Euler equation approach that we use was initially discredited in the early 2000s in the United States literature because it generally failed to deliver consistent and reliable estimates of γ and β despite the large number of studies (Carroll, 2001a). This disappointing performance was attributed to a number of factors including measurement error in consumption data and difficulty in finding suitable instruments for the endogenous variables, especially in the case of the log-linearised Euler equation as explained below.

More recently, lengthening panel datasets, improved availability of consumption data, and analytical refinements have allowed estimation of γ and β, using the Euler equation aproach, in a more precise way. For example, Alan et al. (2009) derive a GMM panel estimator that explicitly allows for measurement

[3] A more structural approach would be to estimate the degree of risk aversion via simulated method of moments, as done in Gourinchas and Parker (2002). Alternatively, stock returns can be used to infer the degree risk aversion, which is the approach underlying the 'equity premium puzzle' literature (see Kocherlakota (1996) for example). An entirely different approach is to estimate risk aversion based on an individual's attitude towards actual (Levy, 1994) or hypothetical (Eisenhauer and Ventura, 2003; Kimball et al., 2008) gains and losses.

[4] For a generic utility function $U(C)$, the Arrow-Pratt measure of relative risk aversion is given by -$CU''(C)/U'(C)$. For other preference specifications, the degree of relative risk aversion and the elasticity of inter-temporal substitution (EIS) are not simple mappings from preference parameters, and instead depend on the current values of consumption (or wealth).

[5] This simple inverse relationship for CRRA preferences is what allows the Euler equation to identify the agent's degree of risk aversion. However, there are more general forms of preferences in which risk aversion and the EIS are governed by separate parameters. Encouragingly, the one study to test this inverse relationship empirically found in favour of it (Yagihashi and Du, 2015).

[6] For CRRA preferences, relative prudence is equal to γ+1. In general, a utility function exhibits prudence when its third derivative with respect to consumption is positive (Kimball, 1990).

error in consumption data, while Alan et al. (2018) show that synthetic panel data based on repeated cross-sections can overcome the endogeneity problems given a long enough time dimension.[7]

Following these recent refinements in the international literature, we provide Euler equation-based estimates of γ for Australia using data from the Household Income and Labour Dynamics in Australia (HILDA) survey. Based upon empirical research and simulation studies from other countries, the HILDA panel is only now becoming long enough to support this kind of analysis, with around 14 years considered to be the minimum required panel length (Alan et al., 2009).

While CRRA preferences are a convenient assumption, they impose a restriction on the relationship between risk aversion and wealth, namely that relative risk aversion is constant with respect to a household's level of wealth. In fact, the empirical support for this form of preferences is not overwhelming, in part because there are relatively few studies that have formally tested the CRRA assumption.[8] Therefore, we undertake an initial test of whether the CRRA assumption is in fact a good one for Australian households, using an approach based on a household's share of risky assets, similar to that in Chiappori and Paiella (2011) and Tsigos and Daly (2016) (the latter being a recent Australian study). Once we account for endogeneity in risk aversion and wealth as well as measurement error – which we show is crucial in such regressions – we cannot reject the CRRA assumption.[9] While constant for a particular household over time, our analysis reveals considerable variation in relative risk aversion *between* households with a median γ of 1.8.[10] As a cross-check, we find that the distribution of this 'objective' measure of risk aversion is significantly correlated with the two subjective measures of risk aversion contained in the HILDA survey.

We next turn to a detailed description of our data and sample selection.

3.3 Data

The HILDA survey began in 2001 and provides annual data on a broad range of economic and social topics. The samples used in all components of our analysis comprise household heads of working age. A head is defined as the oldest male member of a household, or the oldest female in households without a male adult.

Our definition of household head is not important for the results presented below. We identify a 'household head' so that we can follow individuals over time using the `xwaveid` variable in HILDA. The analysis is

[7] Both studies use evidence from Monte Carlo simulations to demonstrate the reliability of their proposed estimators using plausible data generating processes and based on realistic features of the data with regards to panel length and the extent of measurement error in consumption.

[8] Some examples include Ogaki and Zhang (2001); Chiappori and Paiella (2011); Tsigos and Daly (2016); Conlin et al. (2016).

[9] In Appendix 3.C, we show that failing to take account of measurement error in wealth in a regression of risky asset share on wealth can result in a negative estimated coefficient on wealth even when the true coefficient is known to be zero.

[10] While the median value of 1.8 is somewhat above our Euler equation estimates, we note below estimates of γ using risky asset shares are very sensitive to the definition of risky asset (see Appendix 3.B) and assumptions about the market portfolio. For this reason, we consider the point estimates from our Euler equation analysis to be a better indicator of γ for the 'average' or 'typical' household.

based upon household-level data so choice of head is rather arbitrary. In our Euler equation estimates below we only use couples that stay together so choosing the female as the head of household does not alter the results. We have verified this using alternative definitions of household head and the results presented below are unchanged.

We use the sum of expenditure on groceries and meals out as a proxy for household consumption. Unlike other components of household expenditure, information on these components has been collected in almost all years of the survey.[11] Just as importantly, the data quality of these items is probably higher than other components of expenditure in the HILDA survey (Wilkins and Sun, 2010), although they still contain significant measurement error as we quantify in Appendix 3.D. These nominal series are deflated using State-specific consumer price indices for food.

Because of differing data requirements, the samples used in our risky asset share and Euler equation estimates differ in a number of key respects, as we now describe.[12]

3.3.1 Risky asset share regressions

The risky asset share regressions rely on HILDA's wealth module which is included in only four survey years: 2002, 2006, 2010 and 2014. The sample is based on an unbalanced panel of these four years consisting of all responding household heads between ages 23 and 59 inclusive.

In the base model, risky assets are defined narrowly as equities, to allow comparability with the two existing Australian studies, although we consider broader definitions of risky assets in extended models. The regressions only include households with positive risky assets (however defined) and those with information on household size and age of household head. The mean value of the risky asset share is 17 per cent based on the narrow definition of risky assets. The instrumental variable (IV) regressions also require non-missing data on the instruments, that is, household disposable income and household consumption (which we again proxy using expenditure on groceries and meals out).

3.3.2 Consumption Euler equations

Again, our Euler equation analysis uses grocery and meals out expenditure as a proxy for overall household consumption, an approach also adopted in most overseas studies.[13] The panel encompasses the period 2001 to 2016 inclusive, except for 2002 where information on expenditure was not collected in HILDA.

Our sample selection choices closely follow Alan et al. (2009) to aid comparability with that study. We restrict the sample to households consisting of couples who were in a stable relationship over the period in

[11] The expenditure information is taken from the household questionnaire for the years 2001, and 2003 to 2005, and from the self completion questionnaire for subsequent years.

[12] Further details are provided in Appendix 3.A.

[13] This includes the recent US study by Alan et al. (2009). The use of food (and meals out) expenditure as a proxy for consumption in most US studies is because food was the only component of household expenditure covered in the US Panel Survey of Income Dynamics (PSID) from the survey's inception in 1968 until 1999.

which they feature in the survey. Also, the Euler equation only holds for an interior solution, we exclude households who are liquidity constrained, which we define as having zero financial assets in any of the years they feature in the panel. This along with the age restrictions should go some way to overcoming the bias in estimates of γ generated when households are subject to borrowing constraints, a situation that is most likely to affect young households (Keane and Wolpin, 2001). In any case, this restriction removes only 36 observations. The real interest rate series is calculated as the annual-average nominal 3-month bank accepted bill rate less inflation expectations pertaining to that period.[14]

3.4 Testing the CRRA assumption using share of risky assets

3.4.1 Previous estimates

This section reviews the empirical support for CRRA preferences before we proceed to our Euler equation-based estimates. We closely follow Chiappori and Paiella (2011), a recent study that analysed whether relative risk aversion is constant using Italian panel data on risky asset shares. By solving a 2-period optimisation problem of a risk-averse investor, they show that the share of wealth invested in risky assets is proportional to investor h's relative risk aversion, γ_h:

$$\alpha^h = \frac{1}{\gamma^h}\frac{\mathrm{E}[r_m - r_f]}{\sigma_m^2} \tag{3.2}$$

where α_h is the risky asset share, r_m is the return on the risky asset, r_f is the risk free rate and σ_m^2 is the variance of the return on the risky asset.

However, they also formally demonstrate that cross-sectional wealth data alone is insufficient to test whether relative risk aversion is constant. This is because the distribution of risky asset share in the population depends on the joint distribution of wealth and risk aversion, not just the form of preferences (Chiappori and Paiella, 2011). For example, if risk aversion is heterogeneous across the population, it is reasonable to expect that less risk averse people will earn, on average, higher returns and accumulate greater wealth over time. This would lead to a negative correlation between relative risk aversion (measured by risky asset share) and wealth cross-sectionally, even when the underlying preferences were CRRA. Chiappori and Paiella show that panel data can overcome this problem, by allowing one to test whether relative risk aversion varies with wealth for a particular household.

In Australia, Tsigos and Daly (2016) test the relative risk aversion assumption using HILDA wealth data. Using a different measure of relative risk aversion (based on observed portfolio weights for each household), Tsigos and Daly find a negative correlation between risk aversion and wealth. However, they find no relationship between relative risk aversion and wealth when they instead use risky asset share

[14]We experiment with two different measures of inflation expectations: the Reserve Bank of Australia's survey of market economists and the inflation rate implied by inflation-indexed bond prices. The key results are not greatly affected by this choice.

as their measure of risk as in Chiappori and Paiella (2011). In contrast, Cardak and Wilkins (2009) find a small positive relationship between risky asset share and wealth. Their study was also based on the HILDA survey, but they only had access to a single cross-section of wealth data (for 2002) that was available at the time.

3.4.2 Our approach

Given the mixed findings in Australia, we retest the question of whether relative risk aversion varies with wealth. Compared to the previous Australian studies by Cardak and Wilkins (2009) and Tsigos and Daly (2016), we benefit from an additional 3 waves and 1 wave of data respectively. In particular, this allows us to use the panel dimension of the HILDA survey, which was unavailable to Cardak and Wilkins. We follow the measure of risk aversion used in Chiappori and Paiella and Cardak and Wilkins based on risky asset shares.

The basic equation that we estimate is:

$$\log(\alpha_t^h) = \beta_0 + \beta_1 \log(W_t^h) + \beta_2 X_t^h + u^h + \nu_t^h \tag{3.3}$$

where α_t^h is the share of risky assets for household h at time t, W_t^h is total financial wealth, u^h is a fixed effect and ν_t^h is a random error term. To control for the likely endogeneity between risky asset share and wealth, we also estimate the equation in first differences (FD) which eliminates the possible bias from u^h:

$$\Delta \log(\alpha_t^h) = \beta_0 + \beta_1 \Delta \log(W_t^h) + \beta_2 \Delta X_t^h + \Delta \nu_t^h \tag{3.4}$$

As pointed out in Chiappori and Paiella, measurement error in the wealth data can also result in biased estimates of the risk aversion-wealth relationship if not adequately dealt with. This is because total financial wealth is the denominator of the dependent variable and therefore appears on both sides of equation 3.4. We show in Appendix 3.C that for a mean risky asset share of less than 50 per cent, standard multiplicative measurement error in the wealth data will result in moderate to severe downward bias in estimates of β_1 given plausible values for the variance of the measurement error.[15] To address the problem of measurement error, we instrument wealth with disposable household income and food consumption in the levels equation. Similarly, we use growth in disposable income and growth in consumption as instruments for wealth in the first difference equation. Our final specification also includes year dummies (to control for macroeconomic effects) and controls for household composition and age.

[15] In Appendix 3.C, we assume that 20 per cent of the variance in observed wealth is measurement error. While there is very little evidence on the extent of measurement error in wealth survey data, studies that have examined measurement error in food consumption suggest that this assumption is conservative (see for example Ahmed et al. (2006); Brzozowski et al. (2017)). Further, the bias caused by measurement error is likely to be exacerbated in the first difference regression, because the process of differencing data magnifies the noise-to-signal ratio.

Table 3.4.1: Risky asset share regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS Level	IV Level	OLS FD	IV FD	IV FD	IV FD
Risky assets: numerator	E	E	E/FA	E	E+C+T	E+C+T+BE+S
denominator:	FA	FA	FA	FA	FA	FA+BE+S
Total Financial Assets*	-0.128	-0.047	-0.313	-0.173	-0.226	-0.033
	(0.0122)	(0.0452)	(0.0191)	(0.139)	(0.1355)	(0.0222)
Year dummies						
2006	-0.056	-0.08				
	(0.0405)	(0.042)				
2010	-0.229	-0.248	-0.157	-0.136	-0.099	-0.012
	(0.0437)	(0.0451)	(0.0642)	(0.0664)	(0.0658)	(0.0138)
2014	-0.363	-0.391	-0.168	-0.161	-0.166	-0.024
	(0.0463)	(0.0481)	(0.0677)	(0.069)	(0.0684)	(0.0135)
HH size	-0.205	-0.218	-0.258	-0.342	-0.211	0.035
	(0.0409)	(0.0418)	(0.0759)	(0.1018)	(0.0965)	(0.0195)
Age	0.076	0.077	0.015	0.012	0.008	0.007
	(0.0186)	(0.0189)	(0.0303)	(0.0319)	(0.0312)	(0.0056)
Age squared	-0.081	-0.085	-0.025	-0.022	-0.018	-0.01
	(0.0215)	(0.0219)	(0.0371)	(0.0388)	(0.0381)	(0.0071)
Constant	-1.415	-2.199	-0.16	-0.127	0.029	-0.055
	(0.3946)	(0.5738)	(0.6024)	(0.6331)	(0.6161)	(0.1064)
N	6952	6849	3111	3070	3245	10408

Standard errors in parentheses.
E=Equities; C= Cash investments; T=Trusts; BE=Business equity; S=Superannuation; FA=Financial assets = E+C+T+ Bank accounts.
∗ Equal to denominator used to calculate risky asset share.

3.4.3 Results

As suspected, the estimated relationship between risky asset share and wealth is sensitive both to whether we use instrumental variables to adjust for the impact of measurement error and whether the model is estimated in levels or first differences (Table 3.4.1). The basic OLS regression in levels (column 1) shows a strong and highly significant negative relationship (-0.13) between risky asset share and wealth. However, once we adjust for the impact of measurement error by instrumenting for wealth, the coefficient on wealth falls in magnitude and becomes insignificant (column 2). As we argued above, there are good reasons to think that measurement error would contribute to a negative cross-sectional relationship, even when none truly exists, if the mean risky asset share is less than 50 per cent.[16]

Column 3 re-estimates the basic model in first differences. The coefficient on wealth is again negative, but is imprecisely estimated and not significantly different from zero at standard confidence levels.

Columns 4 and 5 experiment with expanded definitions of risky assets. Column 4 adds cash investments (bonds and the like) and trusts to the measure of risky assets, while the regression in column 5 additionally includes business equity and superannuation. In the latter regression, the denominator is also adjusted to include business equity and superannuation. The coefficient on wealth remains insignificant, although it continues to be negative.

[16] The mean share of risky assets in the data is well below 50 per cent unless superannuation is included. Including superannuation as a risky asset results in a mean risky asset share of 73 per cent. If the model in column 1 is re-estimated using this broader measure of risky assets, the estimated coefficient becomes close to zero, which is what we would expect based on our simulation graph in Appendix 3.C.

Given the difference between the OLS and IV estimates, we explore whether the results are robust to alternative instruments for wealth. One obvious alternative instrument for change in wealth is the cumulative difference between disposable income and consumption in the intervening periods between observations on wealth. In other words, this utilises income and consumption data in years that did not contain a wealth module in HILDA. For example, for 2006 we subtract cumulative food consumption from cumulative disposable income for each household over the period 2002 to 2006, which yields a 'flow' measure of change in wealth. When we use this variable as the instrument in models 3 to 5, the key results are unaffected.

We also conduct a weak instruments test, and find that our chosen instruments are well correlated with financial wealth and change in financial wealth.[17] A test of over-identifying restrictions fails to reject the null hypothesis that the chosen instruments are exogenous for models 3 to 5, but does reject the null for model 2. We also try including labour force participation as a component of Z_t, and use its lag as an instrument. This addition does not affect the main results in Table 3.4.1.

Comparison with Tsigos and Daly's results

In finding no significant relationship between risky asset share and wealth, our results confirm the results from Tsigos and Daly's sensitivity analysis but contradict their main result. However, their main result – that relative risk aversion falls with wealth – is based on a different measure of risk that uses portfolio weights calculated using external information on asset returns. Tsigos and Daly do not attempt to explain why the two approaches yield different results in their paper. It is also worth noting that, compared with Tsigos and Daly, our test of the risk aversion-wealth relationship benefits from an extra year of wealth data, giving our test additional power.

While we are unable to reject the CRRA assumption, the estimated coefficient on wealth is negative in all five specifications albeit insignificant in four cases. We therefore view our results as providing tentative support for the CRRA, particularly given that the insignificant estimates only arise in the instrumental variable regressions. Additional years of HILDA data going forward will help us to work out whether risk aversion really is constant with respect to wealth or not.

Distribution of implied relative risk aversion parameter

Equation 3.2 can be used to calculate the distribution of γ_h across the population. However, as Chiappori and Paiella (2011) note, this distribution is only identified up to a scale factor, given by the ratio of the excess return to the variance of the risky portfolio $\frac{r_m - r_f}{\sigma_m^2}$. Taking the plausible (and simple) case where this ratio is equal to one, the value of γ_h is just the inverse of the risky asset ratio for investor h. Clearly, the derived value of γ_h will depend inversely on the definition of 'risky asset', and will be higher or lower for narrower and broader definitions respectively. Additionally, very small values of α_h will result in

[17] The correlation between financial assets and income is 0.44 in levels and 0.18 in growth terms, while the correaltion between wealth and consumption is 0.26 in levels and 0.08 in growth terms.

very large values of γ_h. Following Chiappori and Paiella we therefore truncate the distribution of risky asset shares to exclude values less than 6 per cent and compute the median rather than mean value. The median estimates of γ_h are 1.1, 2.2 and 2.4 across the three definitions of risky assets used in our regressions. We also find little evidence that γ_h varies by wealth. Table B.1 in Appendix 3.B contains further details.

Correlation with subjective data on risk aversion

An interesting question is whether the risk aversion measure derived from risky asset shares correlates with subjective measures of risk aversion contained in the HILDA survey. In particular, the two seperate survey questions ask respondents to indicate the degree of financial risk they are prepared to take.[18] The risky asset share is significantly correlated with both subjective measures, with a correlation coefficient of -0.25 for one subjective measure (which measures risk aversion) and 0.05 for the second measure (which measures risk appetite).

3.5 Estimates of risk aversion using a consumption Euler equation

Based upon our tentative conclusion that relative risk aversion is constant with respect to wealth, we now explore Euler equation estimation based on CRRA preferences. The Euler equation approach of recovering preference parameters began with Hall (1978), and we follow the essence of this approach. In particular, with CRRA preferences, a time discount factor β and an interest rate r, the Euler equation linking consumption in period t to consumption in period *t+1* for household h is:

$$\left(\frac{C_{t+1}^h}{C_t^h}\right)^{-\gamma}(1+r_{t+1})\beta\exp\left(\theta\Delta Z_{t+1}^h+\Delta v_{t+1}^h\right)-1=\epsilon_{t+1} \tag{3.5}$$

$$\mathrm{E}_t\left[\epsilon_{t+1}\right]=0 \tag{3.6}$$

Where the expectation operator relates to the time dimension, this implies that the mean of the expectational error, ϵ_{t+1} is equal to 0 for a given household over time rather than for the cross section.

Two other points about Equation 3.5 are worth making. First, this condition holds only for an interior solution for consumption, and will generally not hold for households subject to binding borrowing constraints. Second, Equation (3.5) is derived on the assumption that consumption and leisure are additively separable in the household's utility function. As a sensitivity test of our main results, we also consider the case where consumption and leisure are non-separable.

[18]The HILDA codes are `_firisk` and `_pntrisk` respectively. For the former variable higher values indicate higher risk aversion, whereas for the latter, higher values imply lower risk aversion.

3.5.1 Exact non-linear equation versus the linearised version

Previous studies have typically proceeded in one of two ways: estimating the exact, non-linear Euler equation as it appears in 3.5 using generalised method of moments (GMM); or taking a first- or higher-order linear approximation of the Euler equation and using OLS or linear IV estimation. Both approaches have drawbacks.

To begin with, both approaches require instruments for the model's endogenous variables – consumption growth and the interest rate. In the non-linear case, economic theory implies that any variables in the consumer's information set at time t are valid instruments, including lags of consumption growth and the interest rate. However, Carroll (2001a) and Alan et al. (2009) show that estimates based on the exact non-linear equation are biased when the consumption data contains measurement error, which is almost certainly the case in practice. These studies suggest that estimates of β will be particularly biased in the presence of measurement error, and this bias does not improve with a longer panel length (Alan et al., 2009).

The linearised approach overcomes the problem with measurement error,[19] but creates its own problems. To see this, first note that the first-order approximation is given by:

$$\begin{aligned}\Delta \log(C_{t+1}) = \frac{1}{\gamma}[&\log(\beta) + \log(1 + r_{t+1}) + \theta\Delta Z_{t+1} \\ &+\Delta v_{t+1} + \log(1 + \epsilon_{t+1}) + k_t]\end{aligned} \tag{3.7}$$

where k_t includes higher order terms remaining from the approximation process.

The problem is that the composite error term is now likely to be correlated with consumption growth and interest rates via the higher-order terms in k_t. Carroll (2001a) shows that these high-order terms are inherently endogenous in a standard model of life-cycle consumption behaviour with wage uncertainty, and that estimates of γ will be severely biased. A further disadvantage of using equation 3.7 is that β can no longer be recovered as it is subsumed into the constant along with the mean values of the measurement error and approximation errors.

Various solutions have been proposed to overcome the difficulties with each of these approaches. In the non-linear case, Ventura (1994), Chioda (2004) and Alan et al. (2009) show that explicitly taking account of measurement error in consumption can yield relatively accurate estimates of the preference parameters γ and β even in panels of only moderate length (that is, around 15 years of data). Each of these studies essentially proceeds by assuming that true consumption ($C_t^\star$) is subject to log-normal iid measurement

[19]With multiplicative measurement error, log-linearising renders the measurement error additive meaning that measurement error would no longer affect the consistency of parameter estimates.

error η_t with equal variance σ_η^2 across households, such that $C_t = C_t^\star \eta_t$. The modified Euler equation is:

$$\mathrm{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + r_{t+1}) \beta \exp\left(\theta \Delta Z_{t+1}\right) - \exp(\gamma^2 \sigma_\eta^2) \right] = 0 \tag{3.8}$$

While this single moment equation does not allow β and σ_η^2 to be separately identified, Alan et al. (2009) show that an analogous moment condition for two-period apart consumption (C_{t+2}/C_t) can be used to identify σ_η^2 (and therefore β).

Alternatively, Alan et al. (2018) show that relatively precise estimates of γ can be obtained using the linearised model if a long enough time series on consumption is available. However, their paper suggests that 'long enough' equates to perhaps 40 years of data, a requirement that few if any panel surveys satisfy. As a practical alternative, they show that a synthetic panel formed from repeated cross-sectional surveys over a long enough period can deliver reliable estimates of γ using the linearised Euler equation.

3.5.2 Our approach

With (at most) 15 years of available consumption data in the HILDA survey, we opt for GMM on the exact non-linear Euler equation, allowing for measurement error in our preferred specifications. For reference, we also obtain an estimate of γ from a linearised version of the model, noting the caveats above in relation to biased estimates in small panels such as ours.

In dealing with measurement error, we consider two approaches. As we use a single moment equation, neither approach allows us to separately identify the discount rate β and the variance of the measurement error σ_η^2 without further moment conditions or an 'external' estimate of σ_η^2.[20] In the first approach, we estimate a single parameter, which is some unknown combination of β and σ_η^2. In the second approach, we use separate estimates of the variance of the measurement error σ_η^2, which enables us to recover an estimate of β. This analysis is reported in Appendix 3.D.

The vector ΔZ_t can in principal consist of any variables that may affect the marginal utility of consumption, including endogenous variables such as labour supply choices (Attanasio and Low, 2004). While we experiment with other variables, our final specification for ΔZ_t only includes change in household size as in Alan et al. (2009). Changes in household size should have a large and direct effect on marginal utility, and this has been confirmed empirically (Attanasio and Low, 2004). We also include a 2008 year dummy in some specifications. This dummy would capture a possible structural break in the relationship between consumption growth and interest rates that may have occurred with the onset of the global financial crisis.

[20] We also consider adding a second two-period apart moment condition, which would allow β and σ_η^2 to be disentangled (Alan et al., 2009). However, we found that the addition of two-period apart moments greatly reduced the precision of the parameter estimates, resulted in implausible values for γ and often led to non-convergence in our GMM estimation routine.

Empirical model

Considering all of the above gives the following moment conditions for each period t (where the household superscript has been suppressed) whose sample counterparts are used in GMM estimation:

$$\mathrm{E}_t \left\{ \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + r_{t+1})\beta \exp\left(\theta \Delta Z_{t+1}\right) - \exp(\gamma^2 \sigma_\eta^2) \right] X_t \right\} = 0 \tag{3.9}$$

where X_t is the set of instruments at time t, as described in the next section.

As noted above, we also estimate a standard linearised Euler equation:

$$\Delta \log(C_{t+1}) = \alpha + \frac{1}{\gamma} \log(1 + r_{t+1}) + \frac{\theta}{\gamma} \Delta Z_{t+1} + e_{t+1} \tag{3.10}$$

where α now includes $\log \beta$ as well as the means of the higher-order approximation errors, and e_t now includes the household's expectational error, the measurement error and the time-varying components of the approximation error.

Choice of instruments

We follow many previous studies in using lags of the interest rate (in addition to a constant) and consumption growth as instruments in our non-linear GMM estimation. There are a number of reasons why these are likely to be good instruments. First, rational expectations suggest that lagged variables, which are known to the household at time t, will be uncorrelated with their forecast error the following period. Second, consumption and interest rates tend to be adequately correlated with lags of themselves, avoiding the problem of 'weak instruments'. Lastly, simulation exercises have shown that they are valid in a standard life-cycle model environment (Alan et al., 2009). We also use change in household size as an instrument for itself, and do likewise with the 2008 year dummy where present. We also experiment with dropping lagged consumption growth from our instrument list (see column 4 in Table D.1). This leaves the lagged interest rate and a constant to identify the two parameters of interest γ and β, which is essentially the just-identified model applied in Alan et al. (2009). Because the linearised model is a log transformation of the true model, we substitute log versions of the instruments used in the non-linear equation.[21]

All estimates of the non-linear model are based on a standard two-step GMM with robust weight matrix. Estimation based on an alternative weight matrix, such as iterative GMM, had very little impact on the results. The linearised model is also estimated using (linear) GMM.

[21] In fact, by subsuming the measurement error into the error term, the error term becomes an MA(1) process. While this means that the first lag of the interest rate is no longer a valid instrument, we find that using the second lag of the interest rate results in implausible (negative) estimates of γ.

Allowing for non-separable consumption and leisure

The Euler equations (3.5) and (3.10) were derived on the assumption of separability between consumption and leisure. However, the few international studies that have formally tested separability have generally found against this assumption.[22] As a sensitivity test, we estimate a modified Euler equation derived from non-separable Cobb-Douglass preferences $U = \frac{1}{\rho_1\rho_2}(C^{\rho_1}L^{1-\rho_1})^{\rho_2}$, where L is leisure. Here, the coefficient of relative risk aversion is a combination of both utility parameters and is given by $1 - \rho_1\rho_2$. The resulting Euler equation is:

$$\mathrm{E}_t\left[\left(\frac{C^h_{t+1}}{C^h_t}\right)^{\rho_1\rho_2-1}\left(\frac{L^h_{t+1}}{L^h_t}\right)^{\rho_2-\rho_1\rho_2}(1+r_{t+1})\beta\exp\left(\theta\Delta Z^h_{t+1}+\Delta v^h_{t+1}\right)\right]=1 \tag{3.11}$$

3.5.3 Results

The GMM point estimates of γ range from 1.18 to 1.39 across the four separable specifications tested; see Table 3.5.0. This equates to an estimated EIS of between 0.74 and 0.85. In the context of Euler equation-based studies, the GMM estimates are relatively precise with standard errors between 0.17 and 0.24. The estimate for γ of 2.4 from the linearised Euler equation is quite a bit higher than the GMM estimates, but is very imprecise. This specific finding is very similar to that in Alan et al. (2009) using PSID data. Their simulation exercise in the same study highlights that estimates of γ based on linearised models are prone to severe biases in cases where the time dimension is less than 30 years or so (see also Attanasio and Low (2004)).

As noted above, β and σ^2_η are not separately identifiable without further assumptions (or additional moment conditions). In models 4 and 5, we assume that measurement error accounts for 80 per cent of the overall variance in consumption growth, a proportion guided by our separate estimates in Table D.1. Given the variance in the Australian consumption data, this implies a measurement error variance of around 0.095 and yields estimates for β between 0.92 and 0.96 across two specifications.

The coefficient on change in household size gives an explanation for the observed hump-shaped profile of consumption over the life cycle, and is relevant for the construction of equivalence scales (see Gourinchas and Parker (2002) among others). Given our sample of households is restricted to couples who have stayed together, this coefficient basically measures the marginal utility associated with an extra child. The estimated coefficient averages about 0.3 for the GMM estimator, with a somewhat higher estimate of 0.48 using the linearised model. These estimates are in the range of those obtained in international Euler equation studies using food expenditure (for example, as Attanasio et al. (1999)), but somewhat below those in Alan et al. (2009).

In the non-separable model, we estimate a coefficient of relative risk aversion of 2.23 (column (6)).

[22] We are not aware of any Australian studies that have examined this question.

Table 3.5.0: Euler equation estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Log. Lin.	EGMM	EGMM	EGMM	EGMM	EGMMn
Coeff. of relative risk aversion (γ)	2.381	1.364	1.285	1.389	1.18	2.245
	(1.261)	(0.166)	(0.202)	(0.236)	(0.199)	(0.734)
Discount factor (β)				0.921	0.962	
				(0.061)	(0.023)	
β/Meas'nt error in consumption		0.761	0.806			0.744
		(0.065)	(0.081)			(0.226)
Household size (θ)	0.524	0.358	0.329	0.376	0.294	1.445
	(0.28)	(0.069)	(0.077)	(0.093)	(0.068)	(1.115)
2008 dummy	-0.074		-0.07			
	(0.01)		(0.05)			
Constant	0.007					
	(0.004)					
Inflation measure for real interest rate	Survey	Survey	Survey	Survey	Bond M.	Survey
Instruments**:						
Lagged interest rate	Yes	Yes	Yes	Yes	Yes	Yes
Lagged consumption growth	No	Yes	Yes	No	Yes	Yes
Households	7225	6227	6227	7225	6227	3249
Observations	44966	35254	35254	44966	35254	16430

EGMM = exact GMM; EGMMn = exact GMM with non-separable preferences.
Standard errors in parentheses.
* Measurement error variance set to 80 per cent of variance in consumption growth, consistent with our findings in Appendix 3.D.
** The set of instruments also includes a constant in all equations, as well change in HH size and the GFC dummy where present for columns (1) to (5). Lagged changes in household leisure time, and age and age-squared of the household head are also included for the model in column (6).

While somewhat higher than the separable model estimates, the much larger standard error indicates no statistically significant difference from the separable estimates at standard confidence levels.

We run a series of tests to check the sensitivity of our results to sample selection. We find that the exclusion of consumption growth outliers has very little impact on the results, nor does introducing a requirement that a respondent appears in at least five consecutive waves (as per Alan et al. (2009)). However, we find that including households whose head has begun or ended a marriage or de facto relationship during their time in the panel leads to implausibly low estimates of γ (well below 1) and instability in the GMM estimation procedure. In column 5, we experiment with an alternative real interest rate series, which is calculated using the inflation rate derived from market pricing of inflation-linked bonds. This leads to an only slightly lower estimate of γ.

Comparison with international estimates

Overall, our point estimates of γ for Australia based on the non-linear Euler equation are within the range of recent estimates obtained in the US.[23] In the most comparable US study, Alan et al. (2009) shows GMM estimates of γ ranging from 1.15 to 1.53 using data from the PSID, with a preferred point estimate of 1.45. Our estimates of γ are also within the range of estimates found in Gourinchas and Parker (2002), who use the quite different approach of simulated method of moments. However, Kocherlakota

[23] We found very few recent studies outside the US.

(1996) argue that these moderate estimates of γ conflict with the very high values needed to explain the large equity premium in the US. Finally, our estimates of 0.92 and 0.96 for β compare to a range of 0.87 (which the authors considered implausibly low) to 0.99 found in Alan et al. (2009). Our estimates of β are economically plausible, albeit they probably lie towards the lower end of the range of values used to calibrate macroeconomic models for simulation analysis.

3.6 Conclusion

Testing the CRRA assumption empirically is crucial as it underpins many theoretical and applied models of decision making under uncertainty, and our paper provides some qualified empirical support for the CRRA assumption for Australia. That said, given the difficulties associated with measurement error in the wealth data as well as the deeply endogenous relationship between risky asset share and wealth, further studies are needed to increase the level of confidence in this result. Such future studies will benefit from additional waves of panel data and may be able to approach the question from a different angle to ours, perhaps relying also on the subjective data in HILDA.

While informative, our Euler estimates of the coefficient of relative risk aversion γ are also subject to uncertainty. Other approaches, such as the more structural simulated method of moments technique established by Gourinchas and Parker (2002), should be used to confront our results given the heavy reliance on this parameter in applied policy analysis.

A recent working paper by Iskhakov and Keane (2018a) estimates a value of 0.8 for γ using simulated method of moments combined with HILDA data. This implies a much lower degree of relative risk aversion than we find.

Our estimates of γ also have interesting implications for precautionary saving that could be explored further.[24] In particular, Kimball (1990) shows that for consumers who face income (or other) uncertainty, the strength of the precautionary saving motive depends directly on their degree of prudence.[25] With CRRA preferences, our preferred estimates of γ imply a coefficient of relative prudence (given by $CU'''(C)/U''(C)$) between 2.2 and 2.4.[26] A structural life-cycle model could be used to explore what this implies for the proportion of household wealth attributable to precautionary saving. We are unaware of any Australian studies that have looked at this question.[27]

[24] There is a large empirical literature devoted to measuring the proportion of household wealth attributable to precautionary saving. Somewhat frustratingly, this literature has produced a very broad range of estimates. See Lusardi (1998) and Carroll and Samwick (1998) for examples at either extreme.

[25] Deaton (1991) and Carroll (2001b) among others point out that the prospect of a borrowing constraints can cause households to engage in precautionary saving even when the utility function does not exhibit prudence.

[26] As explained above, under CRRA preferences, the degree of relative prudence is simply equal to $\gamma + 1$.

[27] Hanel and Haisken-DeNew (2017), using HILDA data, find a large effect of uncertainty on savings (i.e. a strong precautionary saving motive), but do not quantify its contribution to overall household wealth.

Appendix

Appendix 3.A Sample selection

Table A.1: Sample selection for the risky asset share regressions

	Dropped	Remaining
Unbalanced panel of respondents		317,738
Keep years for which there are wealth data	241,025	76,713
Keep working age	38,730	37,983
Keep household heads	17,280	20,703
Keep positive risky assets	1,250	19,453

Additional observations were lost in first differencing and because instrumental variables contained some missing values.

Table A.2: Sample selection for the Euler equation analysis

	Dropped	Remaining
Unbalanced panel of respondents		317,738
Keep couples who did not split up	36,423	281,315
Keep couples	151,319	129,996
Keep HH heads	54,347	75,649
Keep non credit constrained housheolds	36	75,613
Keep working age	1,717	73,896
Keep non-missing data	19,887	55,726

Additional observations were lost because of the presence of lagged variables in the equation and the instruments.

Appendix 3.B Relative risk aversion estimates by net worth

Table B.1: Median relative risk aversion estimate based on risky asset share

Net worth quartile	Equities and shares	+Bond-like assets	+Bus. Equity and Super
Bottom	2.24	2.15	1.1
Second	2.6	2.44	1.1
Third	2.47	2.28	1.09
Top	2.43	2.01	1.08
Overall	2.47	2.18	1.09

Appendix 3.C Impact of measurement error in wealth on risky asset regressions

Research on the impact of measurement error in household panel data has mainly focused on income and expenditure data (see Bound, Brown and Mathiowetz (2001) for a discussion), but it is likely that wealth data also suffers from substantial measurement error. This appendix explores the impact on the estimated relationship between risky asset share and wealth.

We begin by assuming that the logarithm of household wealth, $\log W_h^\star$, is normally distributed with mean μ_W and variance $\sigma^2_{W^\star}$, where the star denotes the true value of the variable.[28] Each household invests a share $\alpha^\star$ in risky assets $W_p^\star$ and $(1-\alpha^\star)$ in risk-free assets $W_f^\star$.

Observed assets are subject to multiplicative measurement errors u_p and u_f, so that $W_j = W_j^\star u_j$ for $j = p, f$. We also assume that $\log(u_p)$ and $\log(u_f)$ are normally distributed with mean zero and variances σ_p^2 and σ_f^2 respectively. We further assume zero correlation between the two measurement errors.

For simplicity, we also assume that households have a common risky asset share $\alpha_h^\star = \bar{\alpha}^\star$. In this setting, cross-sectional variation in $W_p^\star$ and $W_f^\star$ are driven solely by a household's wealth level. Observed wealth, W, and the observed risky asset share, α, are then given by:

$$\begin{aligned} W &= W_p^\star u_p + W_f^\star u_f \\ &= \alpha^\star W^\star u_p + (1-\alpha^\star) W^\star u_f \\ \alpha &= \frac{\alpha^\star W^\star u_p}{\alpha^\star W^\star u_p + (1-\alpha^\star) W^\star u_f} \end{aligned} \tag{3.1}$$

We are interested in how measurement error of this kind affects OLS estimates of β_1 in the following regression:

$$\log(\alpha_h) = \beta_0 + \beta_1 \log(W_h) + \epsilon_h \tag{3.2}$$

In the absence of measurement error, OLS estimates of β_1 are unbiased, consistent and equal to zero given that the correlation between $\alpha_h^\star$ and $W_h^\star$ is zero by construction. However, in the presence of measurement error, the covariance between α_h and W_h becomes:

$$\begin{aligned} \mathrm{Cov}(\log \alpha_h, \log W_h) = \mathrm{Cov}[&\log(\alpha^\star W^\star u_p + (1-\alpha^\star) W^\star u_f), \log(\alpha^\star W^\star u_p) \\ &- \log(\alpha^\star W^\star u_p + (1-\alpha^\star) W^\star u_f)] \end{aligned} \tag{3.3}$$

This will generally not be equal to zero as measurement error affects both the dependent and independent variables. This fact also makes the direction of the bias ambiguous, as it will depend on the precise values

[28] From now on we suppress the subscript h unless it is needed to remove ambiguity.

of $\bar{\alpha}$ as well as the relative variances of the measurement errors. While there is no way of analytically quantifying this covariance expression, and therefore the direction and size of the bias, we can evaluate it numerically by simulating a large number of draws of $W^{\star}$, u_p and u_f.[29]

Figure 3.C.1 shows how the bias in OLS estimates of β_1 varies with different values of $\bar{\alpha}$ based on the numerical evaluation of equation 3.3. The simulations assume that the variances of $\log(u_p)$ and $\log(u_f)$ are equal to each other and impose a signal-to-noise ratio of four (i.e. σ_r^2 and σ_f^2 are one-quarter the variance in $\log(W^{\star})$).[30] For values of α at or below 0.5, OLS estimates of β_1 have a negative bias that becomes more severe as $\bar{\alpha}$ gets smaller. For values of $\bar{\alpha}$ above 0.5, the bias in β_1 is positive but of relatively small magnitude. In short, the most severe (negative) bias in estimates of β_1 are likely to occur when the average risky share is less than 0.5, which is the situation we face in the main text.

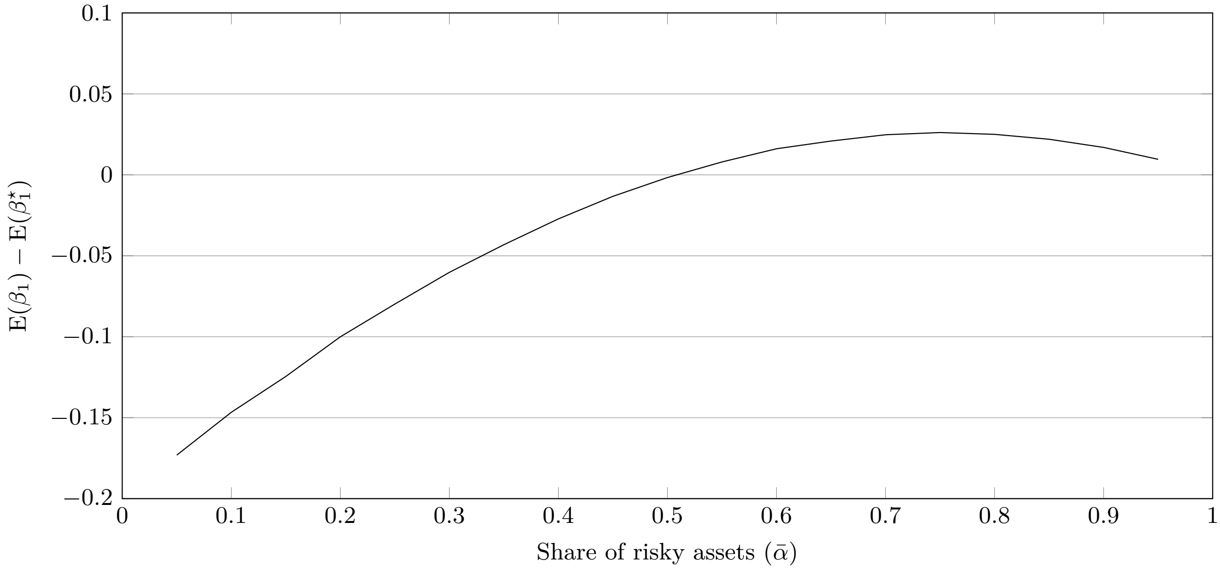

Figure 3.C.1: Bias in OLS estimates of β_1

[29]The covariance depends on non-linear transformations involving the sum of two log-normal random variables, which itself has no close form.

[30]The simulations involved 1,000,000 random draws of $W^{\star}$, u_p and u_f for each value of $\bar{\alpha}$.

Appendix 3.D Estimates of measurement error in grocery data

This appendix estimates the variance of measurement error in grocery expenditure data, which is an input into our Euler equation analysis in the main text.

We begin by making the common assumption that measurement error is multiplicative, so that $C_t^i = \eta_t^h C_t^\star$ where $C_t^\star$ is the true (unobserved) consumption for each household h.[31] We further assume that the measurement error η_t is an MA(1) process such that $\log \eta_t = \epsilon_t + \theta\epsilon_{t-1}$ where ϵ is a normally distributed iid error with with constant variance σ_ϵ^2. Consumption growth is then given by:

$$\Delta c_t = \Delta c_t^\star + \epsilon_t + (\theta - 1)\epsilon_{t-1} - \theta\epsilon_{t-2} \tag{3.1}$$

This implies the following variance and first- and second-order autocovariances for Δc_t:

$$\mathrm{Var}(\Delta c_t) = \mathrm{Var}(\Delta c_t^\star) + \sigma_\epsilon^2 + (1-\theta)^2\sigma_\epsilon^2 + \theta^2\sigma_\epsilon^2 \tag{3.2}$$

$$\mathrm{Cov}(\Delta c_t, \Delta c_{t-1}) = (\theta - 1)\sigma_\epsilon^2 - \theta(\theta-1)\sigma_\epsilon^2 \tag{3.3}$$

$$\mathrm{Cov}(\Delta c_t, \Delta c_{t-2}) = -\theta\sigma_\epsilon^2 \tag{3.4}$$

where the autocovariance in true consumption growth one or more periods apart is assumed to be zero.[32] With three equations in three unknowns, the system is identified and can be estimated using generalised method moments (GMM) given at least four years of panel data.

We fit this model using data on self-reported grocery expenditure from HILDA's self completion questionnaire (SCQ) for the response of the household head and separately for the head's partner.[33] We also provide estimates model based on the spliced grocery and meals out series used in the main text, which covers a slightly longer period, from 2004 to 2016. In the latter case, the data is averaged across multiple responses from a single household where applicable.[34] Our GMM estimation is based on the identity weight matrix.[35]

The estimates in Table D.1 imply that measurement error accounts for between 76 and 87 per cent of the variance in observed growth of grocery expenditure. This is within the range of international estimates, albeit towards the upper end. Interestingly, the process of averaging grocery responses across multiple household members, while reducing the noise, appears to reduce the signal even more so. This results in a higher estimated proportion of measurement error in this case (see column 3). The estimate of the

[31]To avoid notational clutter, we suppress the h superscript from now on.

[32]This key identification assumption is equivalent to assuming that consumption is a martingale with drift, which is a common assumption in dynamic models of household consumption (see for example Blundell et al. (2008)).

[33]We restrict the sample to couple households where both members provide a *unique* response.

[34]The expenditure section of the SCQ can be completed by anyone in the household with "any responsibility for the payment of household bills...". Wilkins and Sun (2010) show that around one quarter of households provide more than one response each year.

[35]Altonji and Segal (1996) show that the identity weight matrix is superior to the 'optimal' weight matrix (the variance of the sample moment conditions) when estimating models of covariance structures in short panels. Also, with unbalanced panel data, there is no guarantee that the variance of the sample moment conditions will be positive semi-definite matrix, a problem that we encountered occasionally.

Table D.1: GMM Estimates of measurement error in expenditure data

	(1)	(2)	(3)
Expenditure measure	Groceries	Groceries	Groc. & meals out
Data source	Head SCQ	Partner SCQ	Av. SCQ & HQ
Period	2006-2016	2006-2016	2004-2016
Var. of true cons. growth (Var($\Delta c_t^\star$))	0.0483	0.0593	0.0205
	(0.0158)	(0.0172)	(0.0018)
Variance of meas'nt error (σ_ϵ^2)	0.1001	0.0911	0.0699
	(0.0142)	(0.0158)	(0.0025)
MA(1) coefficient (θ)	0.0969	0.0356	0.0454
	(0.048)	(0.0889)	(0.015)
Proportion of Var(Δc_t) due to Msmt. Error (%)	79	76	87
Num. of obs.	5502	5502	55554

Ψ

Standard errors in parentheses.

MA(1) coefficient is significant in the first and last cases.

Chapter 4

Empirical evidence on the Frisch elasticity of labour supply for Australia

4.1 Introduction and related literature

The inter-temporal substitution (or Frisch) elasticity of labour supply is a key parameter across many areas of economics.[1] For example, it is an important factor in the optimal design of the tax system where it determines the efficiency loss from labour taxation (e.g. Conesa et al. (2009) and Peterman (2013)) and is central to the economy's business cycle dynamics (e.g. King and Rebelo (1999)).

Despite its importance, there is a dearth of empirical evidence on the value of the Frisch elasticity of labour supply for the Australian population. In this paper, I use 17 waves of data from the Household, Income and Labour Dynamics in Australia (HILDA) survey to estimate this important parameter. I find that the overall elasticity of labour supply in Australia is relatively large; our estimates range from 1.5 to 2.3 when I account for both the intensive and extensive margins and include all men and women aged 25 to 60. It is even higher at 3.2 when I include older workers aged 61 to 65. Our paper also explores how labour supply elasticity varies by partnership status and makes novel use of the HILDA survey including its annual work diary to investigate the importance of intra-year labour supply patterns, which I find is important in the case of women. Finally, I show that our main conclusions are robust to a range of alternative assumptions, including non-separable preferences.

From a methodological standpoint, our study incorporates recent insights from the international labour

[1] As I explain further below, the Frisch elasticity measures the responsiveness of current labour supply to an increase in the after tax wage *holding the marginal utility of wealth constant.* It therefore tells us to what extent individuals substitute labour supply across time in response to anticipated growth in wages over the life cycle.

supply literature.[2] Much of this recent research represents an attempt to reconcile the well-known discrepancy between micro- and macro-based estimates, whereby micro-based studies using panel data samples of prime-age men typically find the Frisch elasticity to be small (between 0 to 0.5), while estimates calibrated from aggregate time series data tend to indicate a much larger elasticity, usually in excess of 2 (Chetty et al., 2011).[3]

In particular, recent empirical and theoretical contributions suggest that the conventionally-small micro estimates probably understate the true aggregate wage elasticity of labour supply for a number of key reasons.[4] First, such estimates are typically based on samples of prime-age men, and extrapolating these estimates to the whole population is unwarranted, as there is substantial evidence that the labour supply responsiveness of women and other age groups is quite a bit higher (Keane, 2011). This is exactly what I find for Australia below. Second, most micro studies have confined their focus to changes in labour supply at the intensive margin, ignoring the additional – and potentially larger – response at the extensive margin via movements into and out of employment. Kimmel and Kniesner (1998) and more recently Fiorito and Zanella (2012) and Peterman (2016) incorporate all workers (both sexes and all working-ages) as well as both margins of response and find much larger elasticities (1.1 to 3.1) using United States panel data. Again, I demonstrate a similar pattern for Australia. Besides these two major reasons, the international literature has emphasised how improvements to the realism of the economic environment can also result in significantly higher estimates of the labour supply elasticity than conventional ones. Prime examples include allowing for the impact of fixed costs of working (Rogerson and Wallenius, 2009; Erosa et al., 2016) and human capital investment (Keane and Rogerson, 2012). Finally, there are a range of econometrics factors that have likely contributed to the conventionally-small micro estimates of the labour supply elasticity. These are mostly related to inadequate strategies for isolating the kind of variation in wages that is needed to identify the Frisch elasticity.[5] Our approach to estimation, particularly our use of a pseudo panel, endeavours to overcome these challenges.

As far as I am aware, there are no existing panel data estimates of the Frisch elasticity for the Australian population.[6] There are however numerous Australian studies that estimate labour supply responses within a static framework, a simplification which permits additional complexity to be added along other model dimensions such as the way that childcare costs and demographics affect the reservation wage (see for example Kalb (2002); Breunig et al. (2008)). While there are no Australian studies with which I can

[2]Throughout this paper, I only discuss the literature most relevant to my own study. Any attempt to summarise the overall labour supply literature is beyond the scope of our paper given its formidable size. Interested readers are directed to Keane (2011) for a detailed survey of the key issues and controversies within the labour supply literature.

[3]I use 'micro' to refer to estimates based on individual-level data and 'macro' to refer to estimates derived from cyclical fluctuations in aggregate hours based on time series data. It would be interesting to compare the results of our micro study with macro estimates for Australia. However I am not aware of any Australian study that derived a Frisch elasticity explicitly from aggregate hours fluctuations.

[4]Within this paper, I use the term 'aggregate elasticity' to refer to the combined intensive and extensive labour supply response to a wage change for the whole working-age population.

[5]As already noted and explained further below, the Frisch elasticity measures the labour supply response to wages changes that leave the agent's marginal utility of wealth unchanged. A predictable wage change falls into this category as does an unexpected wage movement that is small and temporary so that its effects on lifetime wealth is negligible. See Keane (2011) for a detailed discussion of the impact of identification strategies on elasticity estimates.

[6]Iskhakov and Keane (2018b) calibrate a life cycle model with human capital accumulation and discrete labour choice for the Australian economy using HILDA data and simulate the response of labour supply to anticipated transitory wage changes. Their results imply a Frisch elasticity that rises with age and averages about 0.5.

directly compare, our range of estimates is well above conventional micro estimates in the international literature, but is in line with the two recent US studies (i.e. Fiorito and Zanella (2012) and Peterman (2016)) that also incorporate both margins of labour supply, both sexes and a broad age range. Our estimates are intended to be used for a variety of purposes. Policy makers may find the dis-aggregated Frisch elasticity estimates by gender and marital status of particular interest across a range of policy areas, whereas economic modellers may be more interested in the traditional intensive margin estimates which provide the most direct guidance for calibrating preference parameters.[7]

The rest of the paper is structured as follows. Section 4.2 presents the basic life cycle model of labour supply upon which our estimates rest. Section 4.3 summarises some of the main difficulties faced in the labour supply literature in attempting to estimate the Frisch elasticity and discusses different ways of incorporating the extensive margin response. Section 4.4 briefly outlines the data used in our study. Section 5.6 presents the results using various estimation approaches, undertakes some sensitivity tests and discusses some extensions, caveats and implications. Section 4.6 provides a brief summary of our paper's main findings.

4.2 Standard life cycle model of labour supply

Early studies based on static models of labour supply focused on the Hicksian (utility compensanted) and Marshallian (uncompensated) elasticity of labour supply, which are analogous to the well-known concepts in standard consumer theory. However, MaCurdy and Heckman (see MaCurdy (1981) and Heckman and Macurdy (1980)) showed that a third elasticity concept arises when the labour supply model is extended to a life-cycle environment in which agents can save. This third measure, known as the Frisch elasticity, measures the labour supply response to an increase in after tax wages *holding their marginal utility of wealth (λ_t) constant*:

$$\mathcal{E}_F = \left.\frac{\partial \ln n}{\partial \ln w(1-\tau)}\right|_{\lambda_t \text{fixed}}$$

where w is the hourly wage rate, τ is the tax on wages and n is the hours per period devoted to market work. MaCurdy (1981) shows that in a standard life-cycle model the Frisch elasticity $\mathcal{E}_F$ can be estimated from a simple regression of hours worked on after tax wages using appropriate instruments to take account of the likely endogeneity in wages. This is the approach that I essentially follow, but it is worth briefly describing the underlying economic model from which it is derived.

First assume that agents maximise expected utility by choosing a sequence of consumption (c_t) and hours (n_t, which is the complement of leisure l_t) discounted at rate β^t over their lives $t = 1...T$ subject to a period budget constraint $c_t + a_{t+1} = w_t(1-\tau)n_t + a_t(1+r)$ where a_t is savings brought into period t. I also assume that utility is separable across both time and commodities with an iso-elastic period utility

[7]For reasons explained below, the aggregate Frisch elasticity will generally not have a direct relationship with the parameters in an agent's utility function and is therefore not suitable for calibrating such parameters.

function:

$$U(c,n) = \frac{c^{1+\sigma}}{1+\sigma} - \chi\frac{n^{1+\eta}}{1+\eta}$$

where $\sigma \leq 0$ governs the elasticity of inter-temporal substitution, $\eta \geq 0$ governs the elasticity of labour supply, and $\chi > 0$ is the disutility of work. Taking logs of the first-order conditions with respect to consumption and hours and the Euler condition gives:

$$\begin{aligned} \ln c_t &= \frac{1}{\sigma}\ln\lambda_t \\ \ln n_t &= \ln\chi + \frac{1}{\eta}(\ln w_t(1-\tau) + \ln\lambda_t) \\ \ln\lambda_{t+1} &= \ln\lambda_t + \ln\beta(1+r) + \zeta_t \end{aligned} \tag{4.1}$$

where ζ_t is the agent's forecast error and the last condition is a first-order approximation. Now allowing for individual variation through a vector of observables X_{it} that shift taste for consumption relative to leisure with coefficient α and an unobserved taste shock u_{it}, combining the first two conditions in Equation 4.1 provides the following labour supply estimation equation for individual i at time t:[8]

$$\ln n_{it} = \frac{1}{\eta}\ln w_{it}(1-\tau) + \frac{\sigma}{\eta}\ln c_{it} - \frac{\alpha}{\eta}X_{it} + \frac{u_{it}}{\eta} \tag{4.2}$$

In this labour supply equation, consumption is therefore a direct control for the marginal utility of wealth (λ_t). This means that the partial effect of a wage change can be interpreted as the Frisch elasticity, which can be estimated given panel data on hours worked, after tax wages and consumption. With these preferences $\mathcal{E}_F = 1/\eta$, although this direct mapping between preference parameters and the Frisch elasticity does not hold in general.[9]

Differencing equation 4.2 and using an approximation to the agent's Euler equation that links expected consumption growth to the discount factor β, the interest rate r and a forecast error ζ_{it} results in an alternative way of estimating the Frisch elasticity:

$$\Delta\ln n_{it} = \frac{1}{\eta}\Delta\ln w_{it}(1-\tau) - \frac{1}{\eta}\ln\beta(1+r_t) - \frac{\alpha}{\eta}\Delta X_{it} + \Delta\frac{u_{it}}{\eta} + \zeta_{it} \tag{4.3}$$

In practice, time dummies are typically used in place of the interest rate r to capture movements in the after tax return to saving, which is how I proceed below.

[8]See Altonji (1986) for further background on this derivation.

[9]In this case, the approximate Hicks and Marshallian elasticities are also functions of the preference parameters σ and η and are given by $\mathcal{E}_H \approx 1/(\eta-\sigma)$ and $\mathcal{E}_M \approx (1+\sigma)/(\eta-\sigma)$. These expressions are only approximate because the exact elasticities also depend on the share of labour income in total income. Given the restrictions on σ and η, this implies that the following relativity must hold: $\mathcal{E}_F > \mathcal{E}_H > \mathcal{E}_M$.

4.3 Estimation issues

4.3.1 Estimation challenges and instruments

However, the simplicity of these equations belies the numerous practical challenges that arise in estimation. Most importantly, estimating the parameter of interest $1/\eta$ in equations 4.2 or 4.3 requires variation in hourly wages to be exogenous. The problem is that much of the observed variation in wages is likely to occur endogenously with variation in hours worked. This occurs for a number of reasons. First, a change in wage, unless very small, will affect lifetime earnings and therefore have an offsetting impact on labour supply via a wealth effect, which means the wage change is not λ_t-constant. Secondly, unobserved variation in taste for work could well be correlated with wages as well as with hours worked.[10] The third and perhaps largest problem in practice is the well known 'ratio bias' that arises because hourly wages are not measured directly but are instead constructed by dividing reported earnings by hours worked. The resulting wage measure means that hours worked effectively appears on both the left-hand-side and right-and-side of equations 4.2 and 4.3. This induces a spurious negative correlation when the hours data contain any measurement error.[11]

The labour supply literature has used several methods to deal with these endogeneity concerns. I follow what I think is the most convincing approach, which is to instrument the wage in Equations 4.2 and 4.3 using the expected variation in wages over the life cycle. Such variation should be orthogonal both to λ_t and to measurement error.[12] To isolate the expected or predictable part of wage variation, I use the fitted values from a regression of the wage on an age polynomial interacted with years of schooling and father's occupational status. In the levels equation (equation 4.2), we also need to instrument c_{it}, as the marginal utility of wealth (which is what c_{it} embodies) is likely to vary endogenously with persistent wage movements via wealth effects as noted above. I use a similar set of variables to isolate the predictable component of consumption.

4.3.2 Accounting for the extensive margin

Equation 4.2 holds for an interior solution for hours worked so only defines the *intensive margin* Frisch elasticity.[13]

I now define the combined elasticity as the sum of the intensive and extensive labour supply elasticities

[10]For example, productive workers might get greater enjoyment from work than less productive workers.

[11]This means that measurement error in the hours worked data has the effect of biasing labour supply elasticity estimates towards zero.

[12]The other common approach is to use lagged values of wages as instruments for the current wage. I opt for our approach because it is less prone to the 'weak instruments' problem (Peterman, 2016). See Keane (2011) for a detailed review of alternative identification strategies in micro labour supply studies.

[13]Even for an interior solution, the direct mapping from η to $\mathcal{E}_F$ breaks down when the agent is subject to a binding borrowing constraint. This is because a household who experiences a negative wage shock may increase, rather than decrease, their labour supply as a means of smoothing consumption when running down savings or borrowing is not possible (Domeij and Flodén, 2006). Domeij and Flodén (2006) show that this is likely to be a qualitatively important source of downward bias in estimating the true preference parameter.

for the whole population:[14]

$$\frac{\partial \ln \bar{n}}{\partial \ln w} = \frac{\partial \ln P(n > 0)}{\partial \ln w} + \frac{\partial \ln \bar{n}^e}{\partial \ln w} \tag{4.4}$$

where average hours for the working age population ($\bar{n}$) is equal to the product of average hours among the employed ($\bar{n}^e$) and the proportion of the population employed (P).

To account for the labour supply elasticity at the extensive margin we need a way of incorporating the movement into and out of employment between periods in response to wage changes, and to perceived wage changes for those not employed. I explore two ways of doing this.

Pseudo panel estimation

My first way of incorporating the extensive margin is through a pseudo panel methodology. This approach involves transforming individual-level data into cohort averages for all relevant variables. For example, for each cohort c, average hours worked becomes $n_{ct} = 1/N_c \sum_{i_c}^{N_c} n_{it}\mathcal{I}_{ci}$ where $\mathcal{I}_{ci}$ is an indicator function which equals 1 if individual i belongs to cohort c and 0 otherwise. I use an individual's year of birth in defining cohorts. In general, any exogenous, time-constant variable can be used to define cohorts. Using year of birth respects these requirements and also balances the trade-off that must be made between the total number of cohort groups and the number of observations used to form each cohort (Verbeek and Vella, 2005). The resulting average hours series for each cohort incorporates people with both positive and zero hours, which allows an aggregate (intensive + extensive) elasticity to be computed.

Apart from allowing the measurement of the extensive margin response, a pseudo panel also helps to address some of the problems created by endogenous right-hand-side variables. For example, using cohort-averaged data is likely to diminish the impact of measurement error in hours. On the other hand, approximating the true cohort average of each variable using survey data introduces measurement error of a different kind, namely, the deviation of the sample cohort average from the population cohort average. While there is no exact guide, Verbeek and Vella (2005) suggest that the bias induced by this kind of measurement error should not be too great if $N_c > 100$. In this case, N_c ranges from about 180 to 5200 across 50 cohorts.

Estimation is otherwise very similar to the approach for equation 4.3 described above. In particular, predicted wages are used in place of actual wages to account for any measurement error in wages and other sources of endogeneity such as shocks to unobserved tastes for work. This is now based on cohort-averaged wages and right-hand-side variables (a quadratic in age, education and its interaction with the age quadratic, household size, and a dummy for regional status). The predicted wages for non-employed persons is imputed using a standard Heckman sample adjustment similar to that used for equation 4.6 described below. The relatively large sample size used here is likely to mitigate against the poor accuracy

[14] Note that the presence of borrowing constraints means that intensive elasticity for the whole population need not equal $1/\eta$ even if all households share the same parameter value. See previous footnote.

of predicted wages from sample selection models shown in Breunig and Mercante (2010).

Probit fixed effects estimation

I also provide separate estimates of the intensive and extensive labour supply margins along the lines of Kimmel and Kniesner (1998) using separate selectivity-adjusted fixed effects equations for hours and employment. This is done separately for both genders.

Kimmel and Kniesner (1998) measure the extensive margin by modelling the binary employment choice as a probit fixed effects function of the offer wage w_{it}:[15]

$$P(e_{it}) \equiv P(n_{it} > 0) = \Phi(\mu_i^e + \gamma \ln w_{it} + \pi^e f(t) + X_{it}^e) \tag{4.5}$$

where $P(n_{it} > 0)$ is the probability of employment e_{it} for individual i at time t, μ_i^e is an individual effect, $f(t)$ allows for nonlinear time effects, $\Phi(\cdot)$ is the cumulative normal distribution, X_{it}^e is a set of exogenous variables relevant to a person's labour supply decision including family characteristics and u_{it}^e is an independently and normally distributed error term. The extensive elasticity of labour supply is then derived as $\mathcal{E}_F^e = \gamma\phi(\cdot)/\Phi(\cdot)$.

The offer wage for both working and non-working individuals is generated from a standard fixed effects wage equation that has been adjusted for selection into employment:

$$\ln w_{it} = \mu_i^w + \pi^w f(t) + \beta^w X_{it}^w + u_{it}^w \tag{4.6}$$

where μ_i^w is an individual effect, X_{it}^w is a set of variables relevant to a person's market wage (which includes quadratics in age and time, a set of dummies for education and a marital status dummy) and u_{it}^w is an independently and normally distributed error term.[16] Finally, the intensive labour supply response is derived from a fixed effects regression of n_{it} on X_{it}^n which is restricted to have the same variable as X_{it}^e:

$$\ln \hat{n}_{it} = \mu_i^n + \pi^n f(t) + \beta^n X_{it}^n + 1/\eta \ln \hat{w}_{it} + u_{it}^n \tag{4.7}$$

In all three equations, the estimates are adjusted for selection into employment by including the inverse Mills ratio as a regressor. The ratio is derived from a preliminary fixed effects probit equation of employment status as a function of the union of X_{it}^e and X_{it}^w.

Allowing for fixed effects in all equations is designed to control for unobserved individual differences in marginal utility of wealth, taste for leisure and fixed costs of work, among other things. The latter could be particularly important in modelling women's labour supply, where individual-specific fixed costs of

[15] Heckman (1981) shows that the well-known parameter bias in probit fixed effects models is small for $T \geq 8$.

[16] Specifically, the educational dummies are defined over four groups – less than year 12, year 12, vocational training and university degree.

work are likely to explain a significant part of the variation in observed employment and hours (Kimmel and Kniesner, 1998). In the estimation below, I also include the number of dependent children and a dummy for marital status in X_{it}^e and X_{it}^n. Again this is likely to be of particular importance in explaining women's employment and hours choices. Finally, the inclusion of nonlinear time effects noted above is designed to allow for shocks in the marginal utility of wealth.

4.4 Data

The HILDA survey began in 2001 and provides annual data on a broad range of economic and social topics. The samples used in all the analysis are drawn from the first 17 waves of person-level data. For hours worked, I use the variable coded *jbhruc* which is a combination of usual hours, or if hours vary, average hours. The hourly wage is formed by dividing current weekly gross wages and salary for all jobs (coded *wscei*) by hours worked. In the absence of reliable data on marginal tax rates, I focus on the before-tax, rather than after-tax, hourly wage rate. [17] For equations that include consumption, I use the household-level value for combined spending on food, groceries and meals out. Wages and consumption are deflated using the Australian consumer price index. Years of schooling are based on the educational attainment variable *edhigh1* and father's occupational status uses the variable *fmfo6s* which is based on the father's occupation when the respondent is aged 14. The sample is restricted to people aged 26 to 60 inclusive, except in one case where I extend the sample to 65. This leaves a total of about 97,000 observations for 15,800 unique individuals, which are fairly evenly split between men and women (see Table 4.A.1 in the chapter's appendix for further details). In forming cohort averages, I weight individuals by their HILDA person weights (coded *hhwtrps).*

4.5 Results

I provide three sets of estimates:

1. An intensive margin Frisch elasticity by gender using the conventional approach of estimating equations 4.2 and 4.3. These regressions exclude individuals in borrowing constrained households and should in theory provide estimates of the preference parameter η. I define an individual as being borrowing constrained when the value of their household gross financial assets is less than one month's earnings. This results in the loss of about 70 observations in each period.[18]

2. An aggregate intensive plus extensive Frisch elasticity using a pseudo panel based on cohort-averaged data for the combined sample of men and women. Conceptually, there is no direct mapping between this

[17] This may lead to a downward bias in the elasticity estimates, but this bias is likely to be small unless a significant proportion of workers are situated at kink points on the tax schedule. The empirical evidence suggests that the extent of bunching by Australian taxpayers is relatively small (Chapman and Leigh, 2009).

[18] Financial wealth data are available only in 4 waves of the HILDA survey, so only observations in these years are excluded if they meet the criterion for being borrowing constrained. I also experiment with excluding an individual's observations from *all* 17 waves if they meet the criterion in *any* of the 4 waves where financial wealth data is available. This results in the loss of an extra 70 or so observations per year but has no significant impact on the estimates.

elasticity and the preference parameter η because we are no longer conditioning on an interior solution for either hours or saving.

3. Separate intensive and extensive Frisch elasticities using a system of selectivity-adjusted fixed effects equations. Again there is no direct mapping between these elasticities and the preference parameter η in the presence of fixed costs of work and borrowing constraints.

The first set of elasticity estimates, which capture the intensive margin only, are shown in Table 4.5.1. All estimates are significantly above zero but are also substantially less than one, ranging from 0.1 to 0.6. For men, the levels model suggests a small value of 0.12 (column I) with a somewhat higher number of 0.62 in the difference specification (column III). When I confine the sample to male household heads - the approach used in many previous studies – the estimate falls to 0.28 in the first difference specification (column V). These values are squarely within the range of estimates typically obtained in international studies based on samples of prime-aged men (Keane, 2011). For women, the estimates produced by the levels and difference approaches are broadly consistent with one-another at 0.47 and 0.50 respectively.

Estimates based on differenced data tend to be more prone to problems of weak instruments, given the typically low R-squared in the first stage regression of wages growth on the instruments (Altonji, 1986). While Table 4.5.1 confirms that the standard errors are in fact larger in relation to the first difference estimates, there does not appear to be a major problem with weak instruments, as the F-statistics from the first-stage regressions all exceed 5.[19] Conversely, levels-based estimates tend to be more prone to model mis-specification and rely crucially on the availability of reliable consumption data.[20] Indeed, the positive estimated coefficient on consumption in column I is inconsistent with the (economically sensible) restrictions imposed on η and σ in the underlying model described above. Further, I find that the levels-based estimates are relatively sensitive to the specification used in the first stage wage regression.

Table 4.5.1: Intensive margin elasticities based on MaCurdy-style regression

	I	II	III	IV	V
Dependent variable	h	h	Δh	Δh	Δh
Sample	All men	All women	All men	All women	Male heads
Wage	0.12	0.47	0.62	0.50	0.28
	(0.038)	(0.029)	(0.175)	(0.185)	(0.187)
Consumption	0.08	-0.07	-	-	-
	(0.057)	(0.051)	-	-	-
Controls					
Number of dependent children	Dummies	Dummies	Δ	Δ	Δ
Year dummies	Yes	Yes	Yes	Yes	Yes
F-stat excluded instruments	271.58	322.1	5.52	5.18	5.18

Standard errors in parentheses; regressions use HILDA person weights.

Instruments: year dummies, number of dependent children, age, age-squared, education, education interacted with age polynomial, father's occupational status.

[19]Keane (2011) points to a rule of thumb that the F-statistic should be least 5 for the results to be trusted.

[20]For example, the levels based estimates are not robust to violation in my assumption that consumption and leisure are separable within periods. In contrast, the first difference estimates are robust to this violation Altonji (1986).

The second set of elasticity estimates, contained in Table 4.5.2, measure the combined intensive and extensive response using a pseudo panel with cohort-averaged wage growth as the dependent variables. These estimates range from 1.76 to 2.36 for the 26 to 60 age range. This suggests a substantial aggregate labour supply response to wealth compensated wage changes, notwithstanding the relatively large standard errors. When I extend the age range to include people aged 61 to 65 (as in Peterman (2016)), the estimated elasticity increases to 3.29 from 2.36 (see column relative V versus column IV). This points to labour supply being particularly elastic among older workers. This is perhaps unsurprising given that people in the 60 to 65 age range are likely to have fewer constraints on their time that prevent them from adjusting their labour supply in response to wage movements (such as childcare). A similar finding is made in Peterman (2016).

Table 4.5.2: Combined elasticity estimates based on pseudo panel

	I	II	III	IV	V
Wage	2.38	1.76	2.3	2.36	3.29
	(0.593)	(0.368)	(0.607)	(0.499)	(0.556)
Δ HH size	0.43	0.137	0.461	0.447	0.381
	(0.137)	(0.1)	(0.133)	(0.135)	(0.136)
Ages	26-60	26-60	26-60	26-60	26-65
Other controls:					
Time effects	None	Dummies	None	Lin. trend	Lin. trend
Region dummy	No	No	Yes	No	No
F-stat excluded instruments	5.82	5.82	5.82	5.82	5.82

Standard errors in parentheses.

Instruments: age, age-squared, education dummies, education dummies interacted with age, regional dummy.

Again, weak instruments does not seem to be a major problem with the pseudo panel estimates. Moreover, the F-statistic from the first stage wage equation is around 6, above the commonly cited threshold of 5. To test whether the results are sensitive to the number of cohort groups, I split the sample by gender *and* year of birth which gives a total of 100, compared with 50 in the main results. This alternative definition of cohort does not greatly affect the estimates (see Table 4.B.1 the Appendix). The third set of estimates, contained in Table 4.5.3, are broken down into the extensive (employment) and intensive margin (hours worked conditional on being employed). Overall, these estimates are consistent with the combined elasticity being well above 1 as in the pseudo panel estimates. A simple average across men and women implies an elasticity for the whole population of about 1.84. However, I find women to be substantially more elastic than men along both the extensive (1.35 versus 0.25) and intensive (1.66 v 0.41) responses.

Table 4.5.3: Selectivity-adjusted extensive and intensive elasticity of labour supply

	Men	Women	Single men	Married men	Single women	Married women
Employment	0.25	1.35	0.31	0.1	0.93	1.95
	(0.168)	(0.372)	(0.197)	(0.111)	(0.408)	(0.531)
Hours worked	0.41	1.66	0.36	0.44	2.00	1.93
	(0.062)	(0.138)	(0.109)	(0.083)	(0.203)	(0.205)
Implied total	0.66	3.01	0.67	0.54	2.92	3.88
F-stat excluded instruments	116.94	404.95	21.77	276.75	36.95	533.88

Standard errors in parentheses.

Instruments: age, age-squared, education dummies, quadratic time trend, marital status.

For comparison with Kimmel and Kniesner (1998), I also provide estimates by marital status. The main finding here is that married women appear to be more responsive than single women along the extensive margin, whereas the reverse is true for men. My finding contrasts with Kimmel and Kniesner (1998), who find that single women are more elastic than married women along the extensive margin, and attribute this to “schedule co-ordination problems” between married couples that serve to enlarge the wife’s reservation hours and reduce their wage elasticity. However, my finding that married women are more elastic to wage changes than single women is in line with the idea that single women may operate more like men in their labour supply behaviour. Given the changing nature of relationships and family structures over time, I also explore a broader definition of partnership that includes *de facto* relationships as well as legal marriages.[21] This alternative definition had very little impact on the results as shown in Table 4.B.2 in the Appendix. I also split the sample into households with and without children under 5 years of age. In this case, I found that women in households with children under 5 had a significantly larger elasticity along the intensive margin than other women.

I find no evidence of weak instruments, consistent with the greater variation that comes from level rather than differenced variables. This finding accords with the general experience in the labour supply literature (e.g. Altonji (1986), Kimmel and Kniesner (1998), and Domeij and Flodén (2006)). In particular, the test of excluded instruments – which is based on the explanatory power of the assumed exogenous variables in the wage equation – has a large F-statistic of more than 30 across all specifications investigated. As a sensitivity test, I experiment with alternative ways of allowing for non-linear time effects in labour supply by interacting the education dummies with the quadratic time trend. While the estimated coefficients are generally somewhat lower, the broad conclusions are unchanged (see details in Table 4.B.3 of the Appendix).

[21] In the HILDA survey, partners are either married or *de facto* and include same sex couples.

4.5.1 Further exploration

Incorporating information from HILDA's time use calendar

The idea underlying Equation 4.4 is that total labour supply can be decomposed into an average hours for those already working (the intensive margin) and share of people employed (the extensive margin at the population level). But as Blundell et al. (2013) note, the definitions of the intensive and extensive margins really depend on the reference period being considered. When the reference period is a year (as is the case with data used in this paper), they argue that it makes more sense to define the extensive margin as the proportion of days during the year that an individual worked, rather than the conventional approach of defining a binary variable for employment.

The HILDA survey includes information on the proportion of the financial year each individual spends in market work (as well as other activities such as education), which is summarised in the variable *capj*. One simple way is to expand the labour supply measure used in the first set of regressions to at least partly take account of the extensive margin by simply multiplying usual hours worked by *capj*. When I re-run the regressions in columns IV and V of Table 4.5.1 using this expanded definition of labour supply, the elasticities rise to 0.70 (from 0.62) for men and to 1.36 (from 0.50) for women. For women, this indicates that the proportion of the year spent working is an important labour supply margin in its own right.

Alternative measures of 'age'

So far I have used a person's age as a proxy for their level of work experience, which is a common approach in the life cycle literature. In relation to employed men, the distinction between age and experience may be of little empirical importance because of this group's strong attachment to the labour force once they begin their career after completing full time studies. However, for women the choice of work experience variable may make a material difference to the results, given the greater prevalence of career breaks. Conveniently, the HILDA survey includes the variable *ehtjb* which is a more direct measure of work experience. Specifically, the variable *ehtjb* provides an estimate of each individual's total number of years in paid work. I can therefore investigate whether the results underlying Table 4.5.1 are sensitive to using the variable *ehtjb* in place of age. After re-running the regressions in columns IV and V of Table 4.5.1, the new estimates are 0.69 (up from 0.62) for men and 0.97 for women (up from 0.50). As anticipated, the estimated elasticity for women is somewhat sensitive to this alternative way of defining experience, while the estimate for men is largely unaffected.

Non-separable consumption and labour

I have made the strong assumption throughout the previous analysis that the period utility function is separable in consumption and labour. If this assumption is wrong, consumption can no longer act as a proxy for the marginal utility of wealth λ_t, as movements in consumption shift the supply of leisure, which alters the marginal utility of wealth. This means that the coefficient on the wage term in Equation 4.2 will no longer equate to the Frisch elasticity, and the model will be mis-specified. The appropriate modification of the estimation equation depends on the exact form of non-separability being assumed. While it is notoriously difficult to empirically test for the existence of non-separability, the general consensus is that there is a positive but small complementarity between consumption and labour (Chetty, 2006). To check the robustness of my results to this possibility, I first run a regression of consumption growth on the growth in wages, which I instrument using the same variables in the MaCurdy style regressions, along with age and year dummies. In such a regression, the coefficient on wage growth should provide a measure of the degree of complementarity between consumption and labour (Altonji, 1986). The estimates yield a value of 0.22 but with a standard error of 0.2, consistent with there being a small, perhaps zero, degree of complementarity. As an alternative check, I instead assume that agents have Cobb-Douglass preferences $[c^{\rho}(1-n)^{1-\rho}]^{1-s}/(1-s)$ thereby allowing for complementarity between consumption and labour.[22] I then re-run the MaCurdy style first difference-based estimates[23] and find implied intensive Frisch elasticity estimates of about 0.4 for both men and women.[24] These estimates are similar to those in Columns III and IV of Table 4.5.1 that are based upon separable preferences.

4.5.2 Omissions and caveats

The estimates have taken into account a number of important factors including the multiple margins of labour supply, the existence of fixed costs of work and differences by gender, family structure and age. However there are also numerous, possibly important, factors that have been ignored. One such factor is the impact of human capital accumulation on labour supply choices. For example, Imai and Keane (2004) argue that there are sizable human capital benefits to work experience, especially at younger ages, in addition to the monetary wage received by the worker. This results in a much flatter total return on labour supply (which the authors call the 'shadow wage') over the life cycle, which in turns implies a much larger Frisch labour supply elasticity given the empirical profile of hours worked.[25] In summary, ignoring the impact of on-the-job human capital accumulation suggests that the estimates in this paper could understate the true elasticity of labour supply.

From an econometric viewpoint, all my estimates of the Frisch elasticity ultimately hinge on a small set of instrumental variables to identify λ_t-constant wages changes. The choice of instruments assumes that the common life-cycle evolution in wages is both expected by the agent and exogenous to shocks in their taste for work. While the first of these assumptions seems plausible, the second is more contentious and is difficult to test. In an ideal world, the results would therefore be confronted with results from an entirely different approach, for example, by using a quasi-experiment method. However, quasi-experimental

[22] Consumption and labour are complements if $s > 1$, which is the more likely case.

[23] Altonji (1986) claims that the first difference estimates in Equation 4.3 should be unaffected by non-separability, but it is not obvious why this would be the case (Domeij and Flodén, 2006).

[24] Proceeding with Cobb-Douglass preferences produces a *leisure* supply equation and an estimate of the wage elasticity of leisure. This must be multiplied by $(1 - \bar{n})/\bar{n}$ to derive the labour supply elasticity. I assume a time endowment of 112 hours per week, noting that the resulting estimates will be sensitive to the value used for the time endowment.

[25] Interestingly, their model also implies a Hicks elasticity (in response to permanent after tax wage change) that is an order of magnitude larger than conventional estimates (Keane, 2015).

scenarios are often hard to find.[26]

Finally, I have also ignored the issue of infra-household labour supply, which may be another source of bias in my estimates. In addition, estimating a joint model of couples' labour supply would shed light on how one member of a household responds to changes in his/her partner's wage, which appears to be an empirically important channel by which households react to wage shocks (Attanasio et al., 2005; Blundell et al., 2016).

4.5.3 Implication for calibrating preference parameters and policy experiments

It is worth reiterating that only the first of the three estimation approaches employed in this paper - the 'conventional approach' – can provide a direct estimate of the preference parameter η.[27] These estimates suggest that a value of $\eta = 2$ would be a reasonable calibration of the utility function, consistent with a borrowing-unconstrained, intensive-margin Frisch elasticity of 0.5.

The aggregate elasticity being measured in the second and third approaches – which includes the extensive margin response as well as households who face binding borrowing constraints – may have only a very loose connection to η. Does this mean these estimates are less useful for the purposes of policy experiments? As Chetty et al. (2011) point out, such aggregate elasticity estimates, while no longer equating to a deep structural parameter, are still very valuable for considering policy experiments in cases where the interest is on aggregate labour supply responses taking all these factors as given.

4.6 Conclusion

This paper has used a variety of estimation approaches to explore the responsiveness of labour supply to changes in the after tax wage in Australia. Overall, my findings point to a relatively high Frisch elasticity: my point estimates range from 1.7 to 2.3 for the population aged 26 to 60, rising to 3.3 when the age range is expanded to 26 to 65. These large aggregate elasticity estimates reflect the combination of incorporating the extensive margin (which I find especially important for women) and a more elastic supply response from women and older workers in general. Both are ignored in much of the labour supply literature.

Equally, this paper's estimates appear to confirm that the hours worked (or intensive) response by employed men is relatively modest. The point estimates for men aged 26 to 60 range from 0.12 to 0.62 across the various specifications I have considered, and is 0.28 when I confine the sample to only household heads, broadly consistent with the international literature. Accounting for the additional effect of

[26] A classic example of a quasi-experimental labour supply study is Bianchi et al. (2001), which examines the impact of Iceland's 1987 tax holiday.

[27] Of course, this mapping also requires that the assumed model was an accurate representation of the underlying data generating process and that the equation was not otherwise mis-specified.

infra-year work intensity had little impact on the estimates for men but resulted in a substantially higher estimate for women.

My findings may have important implications for policy design to the extent that existing policies are based on the assumption that labour supply is uniformly inelastic across the population. For example, changes in income tax rates may have a greater impact on aggregate labour supply than currently assumed, via an increase in women's participation.

Appendix

Appendix 4.A Sample selection

Table 4.A.1: Sample selection

Removal criteria	Observations dropped	Observations remaining
Initial unbalanced panel		341,170
Missing education data	88,125	253,045
Outside age range 26 to 60	104,508	148,537
MaCurdy style regressions		
Zero or missing hours	44,013	104,524
Missing instruments or controls	26,127	78,397

Appendix 4.B Sensitivity checks

Table 4.B.1: Combined elasticity estimates based on pseudo panel - extra cells by gender

	I	II	III	IV	V
Wage	2.29	1.72	2.18	2.26	3.15
	(0.602)	(0.37)	(0.617)	(0.506)	(0.565)
Δ HH size	0.432	0.137	0.463	0.449	0.385
	(0.137)	(0.1)	(0.134)	(0.136)	(0.137)
Ages	26-60	26-60	26-60	26-60	26-65
Time effects	None	Dummies	None	Lin. trend	Lin. trend
Region dummy	No	No	Yes	No	No
F-stat excluded instruments	5.73	5.73	5.73	5.73	5.73

Standard errors in parentheses.

Table 4.B.2: Extensive and intensive elasticity of labour supply - broader definition of partnered

	Men	Women	Single men	Married men	Single women	Married women
Employment	0.24	1.38	0.24	0.17	1.03	1.61
	(0.161)	(0.381)	(0.169)	(0.154)	(0.522)	(0.45)
Hours worked	0.45	1.59	0.33	0.51	1.83	1.64
	(0.06)	(0.133)	(0.156)	(0.068)	(0.253)	(0.166)
Implied total	0.68	2.97	0.56	0.68	2.86	3.25

Standard errors in parentheses.

Table 4.B.3: Extensive and intensive elasticity of labour supply - including education interacted with time trend

	Men	Women	Single men	Married men	Single women	Married women
Employment	0.15	1.07	0.22	0.06	0.68	1.56
	(0.102)	(0.295)	(0.142)	(0.074)	(0.304)	(0.424)
Hours worked	0.27	1.05	0.34	0.19	1.21	1.19
	(0.045)	(0.109)	(0.085)	(0.059)	(0.161)	(0.16)
Implied total	0.42	2.11	0.56	0.25	1.89	2.75
	116.94	110.87	123.93	196.08	38.45	636.55

Standard errors in parentheses.

Chapter 5

Optimal income taxation, marital risk and family labour supply

5.1 Introduction

The issue of optimal taxation has occupied economists and policymakers since at least the time of Adam Smith yet central questions remain unresolved. Chief among these is whether capital should be taxed in the long run. On the one hand, the traditional public finance literature (e.g. Atkinson and Stiglitz (1976)) suggests that the optimal tax on capital is zero. This follows from the uniform commodity taxation principle being applied to consumption (and leisure) in different periods of time. Chamley (1986) and Judd (1985) extend this result to a dynamic environment with a single infinitely-lived household.

However, subsequent studies have challenged this classical result. Within an infinitely-lived setting, Aiyagari (1995) finds that the introduction of un-insurable idiosyncratic risk and borrowing constraints causes precautionary saving among households that leads to excessive capital accumulation in aggregate. This tendency towards dynamic inefficiency makes it optimal for the government to tax capital. Likewise, a parallel line of studies based on a class of overlapping generations (OLG) models with elastic labour supply (but lacking idiosyncratic risk) have also found that a positive tax on capital is generally optimal. A key reason for this is that a household's consumption and leisure are generally not constant over the life cycle, so that uniform commodity taxation no longer applies. Instead, the typically increasing life cycle profiles for consumption and leisure make it efficient to tax these goods more heavily as the household gets older. Taxing capital achieves this objective indirectly by raising the relative price of future consumption and leisure, as Erosa and Gervais (2002) show in their analytical study. Subsequent numerical-based OLG studies seem to suggest that the result in Erosa and Gervais (2002) may generalise to more complex and realistic model environments such as those with idiosyncratic risk, fixed differences between workers' productivity, an explicit retirement period, a social security system and the capacity to tax labour income progressively. Most notable is Conesa et al. (2009) who find an optimal long-run tax on capital of between 21 and 36 per cent depending on the form of preferences. Fehr and Kindermann (2015) and Peterman (2013) broadly confirm the result in Conesa et al. (2009).[1]

But how robust are these studies? In this paper, I re-examine the question of optimal income taxation in an OLG model that features a more realistic representation of the household sector. In particular, I extend the 'lifelong bachelor' model used in Conesa et al. (2009) by including two-earner married (in addition to single) households, and by allowing each individual's marital status to evolve scholastically over their lifetime. I calibrate the model economy to the United States using an approach and target values similar to those used by Conesa et al. (2009). This aids comparability with their study.

The main finding is that the introduction of these two aspects – a dual labour supply margin and marital risk – results in an optimal tax on capital of zero. This contrasts with the optimal tax of 30 per cent I find using a standard 'bachelor' version of my model (which roughly corresponds to the model used in Conesa et al. (2009)). Secondly, I find that the optimal tax on labour income is a flat tax of 30 per cent (compared with 26 per cent in the bachelor model) with a deduction equivalent to 16 per cent of average income (compared with 22 per cent). The main results are robust to whether the interest rate is endogenous or exogenous, and whether the tax system involves separate or joint filing by married couples. What drives the zero tax on capital? As I demonstrate analytically in a simple 2-period model with wage certainty and inelastic labour, the introduction of marital risk has two important competing effects on saving. For married households, the possibility of divorce next period induces a form of precautionary saving, which leads to higher saving than otherwise. Conversely, for single households the possibility of becoming married next period reduces saving via a type of free rider problem, which leads them to save less than they would otherwise. This latter effect is magnified by the economies of scale in consumption available to married households. I show that the net effect of these two competing forces is to reduce overall saving for plausible parameter values, an important insight for understanding the mechanisms at play in the full OLG model. In particular, my analytical results imply that a relatively low tax on capital may be optimal in the presence of marital risk as a way of mitigating the tendency towards under-saving.

In the full model, the interaction of wage uncertainty and dual labour supply re-enforces the benefits of taxing capital lightly. Married households have an enhanced capacity to exploit the stochastic variation in productivity so long as partners' wages are not perfectly correlated. For married households, the environment starts to look more like the complete markets model analysed in Pijoan-Mas (2006). While precautionary saving can now be undertaken more cheaply, it is less valuable (given the ability to adjust

[1] That said, both Fehr and Kindermann (2015) and Peterman (2013) are more qualified in their conclusions than Conesa et al. (2009) in the following respects. The former study highlights that the optimal long-run tax rate is sensitive to a number of alternative model assumptions, including the life-cycle decline in the elasticity of labour supply generated by the particular form of preferences assumed in Conesa et al. (2009) and the inclusion of accidental bequests in the tax base for capital income. In the absence of compelling empirical evidence to the contrary, I retain the preference form that is most sympathetic to a positive tax on capital as used in Conesa et al. (2009). In contrast with Conesa et al. (2009), I tax bequests separately to other capital income to prevent the taxation of bequests being a driver of my main result (as argued in Peterman (2013)) but I also experiment with the Conesa et al. (2009) approach as a sensitivity check. Meanwhile, Fehr and Kindermann (2015) show that the optimal tax on capital is non-zero only in the long run, and that it falls to zero when the welfare of transitional generations is taken into account.

work effort along two dimensions) which makes saving more elastic to the interest rate.[2] I illustrate this channel in an extensive sensitivity analysis by examining the extreme case of zero correlation in partners wage shocks. This allows an even larger capacity for self insurance through dual labour supply adjustments, which results in an even higher interest rate elasticity of saving. At the other extreme – when partners' wage shocks are perfectly correlated – the sensitivity analysis shows that a large positive tax on capital is instead optimal, reflecting a much lower interest rate elasticity of saving in this version of the model. This finding confirms the well established theoretical link between the elasticity of saving and the optimal tax on capital (e.g. Summers (1981); Engen and Gale (1997) and İmrohoroğlu (1998)). *Therefore, the zero tax on capital result is driven by the presence of marital risk and family labour supply.* Given the empirical relevance of both these features, my study represents an important contribution to the optimal taxation literature.

In addition, my paper relates to the study of 'partial insurance' – the mechanisms by which households buffer themselves against income shocks over their life cycle in the absence of formal insurance markets. Contributions to this field include Blundell et al. (2008) and Kaplan and Violante (2010) who look at the transmission of inequality from earnings to consumption over the life cycle, and Blundell et al. (2016) who extend these studies by allowing for endogenous family labour supply. My paper is novel insofar as it analyses partial insurance within a general equilibrium setting, allowing issues of inequality transmission to be examined in conjunction with questions of optimal taxation. By incorporating two-person households and marital risk, this paper also adds to broader efforts at modelling households in a more realistic way than the conventional bachelor representation used in standard macroeconomic models. In addition to tax policy, this realism is especially relevant in the area of social security and retirement income analysis where policies often depend on the specific structure and marital status of the household.[3] It is worth highlighting that the optimal tax system I consider in the main policy experiment involves a shift from joint to separate filing for married households. In this respect, my study closely relates to Guner et al. (2012a) who also explore a shift to separate filing, as well as the welfare effects of moving to proportional income taxation. However, in their study, this is carried out in two standalone shocks and there is no attempt to search for the *jointly optimal* tax structure.[4] Moreover, compared with the model I use in this paper, their model lacks idiosyncratic earnings risk and changes in marital status. This arguably makes my model more suitable for identifying the optimal income tax mix, as idiosyncratic risk has been identified as an important determinant of optimal progressivity (Conesa et al., 2009), while the absence of transitions in marital status is likely to provide misleading overall welfare changes. This is because an empirically-consistent life cycle evolution of marital status creates an interplay in the welfare functions of married and single households. A newborn generation can *ex ante* expect to spend time as both a single and married household over their lifetime, and should rationally incorporate both these

[2] The dual labour supply margin also means that the zero borrowing constraint is less likely to bind than for single households, which all else equal reduces the precautionary saving motive.

[3] Recent examples in this general area include Nishiyama (2019) who examines social security reform with dual-income households, Kaygusuz (2010) who investigate the impact of United States tax reform in the 1980s on women's labour market participation and Guner et al. (2012b) who look at taxing married men and women at different rates.

[4] Interestingly, the welfare gains in my experiment turn out to be roughly half-way between those in Guner et al. (2012a) for their two shocks.

future possibilities into their optimisation decisions.[5]

The rest of the paper is organised as follows. Section 5.2 develops some intuition for how the introduction of marital risk and dual-earner households will affect saving and therefore optimal taxation of capital in the full OLG model. This section begins by presenting some analytical results from a simple 2-period model to isolate the impact of marital risk on the household saving of both single and married households. This is followed by a qualitative discussion of the impact of dual labour supply on saving. Section 5.3 outlines the model and the recursive equilibrium. Section 5.4 discusses calibration and presents some features of the baseline economy. Section 5.5 explains the policy experiment, while Section 5.6 presents the main results, uses a sensitivity analysis to isolate the key factors behind the zero capital tax result and discusses the connection between the optimal tax mix and partial insurance. Section 5.7 briefly outlines some areas where further research is warranted.

5.2 Impact of adding marital risk and dual earner households

5.2.1 A two-period model with marital risk

There are relatively few studies that examine the channels by which marital risk affects household saving in a multi period setting, either numerically or analytically. Exceptions include the numerical studies by Cubeddu and Rós-Rull (2003) and Voena (2015) and the analytical results in Doepke and Tertilt (2016).[6] This section therefore presents a simple two period partial equilibrium model in order to derive a number of useful analytical results about the impact of marital risk on the saving of married and single households. To focus on the impact of marital risk and to facilitate analytical expressions, I assume labour supply is exogenous and abstract from wage uncertainty in both periods.

Problem for a couple household

For a married household (denoted with a superscript c), I assume that the couple's consumption (denoted c^c) is a public good among partners with economies of scale (or conversely congestion factor) governed by $1 \leq \theta \leq 2$ and the man's Pareto weight (or relative bargaining power) given by λ. In the first period, the married household faces an exogenous probability D of becoming single in the second period. The married household receives $w_1^m + w_1^f = w_1$ in period 1 and $w_2^m + w_2^f$ in period 2 if marriage continues, where the superscripts denote male (m) and female (f). If divorce occurs, the divorced man and woman separately receive wages w_2^m and w_2^f respectively in period 2, while the man receives α share of the assets and the woman $1 - \alpha$. If β is the time discount factor, the couple household's maximisation problem is

[5] On the other hand, Guner et al. (2012a) include childcare costs and fixed heterogeneity (via education differences) in their model, consistent with their central focus on how tax policy changes could affect female labour supply at the extensive margin. While my model does not allow for fixed costs of working, the choice of preferences nonetheless results in a non-trivial extensive margin among married households.

[6] The life cycle study by Cubeddu and Rós-Rull (2003) analyses these competing effects using numerical simulations.

therefore:

$$\max_{c_1, a^c} \left\{ u(c_1^c/\theta) + \beta[(1-D)v^c(a^c) + D\lambda v^m(\alpha a^c) + D(1-\lambda)v^f((1-\alpha)a^c)] \right\}$$

subject to

$$c_1^c + a^c = w_1^m + w_1^f$$

where a^c denotes saving, and the value functions for continued marriage and divorce for each spouse are given by $v^c(a^c)$ and $v_g^s(a^c)$ $(g = m, g)$ respectively. A gross return of $1 + r$ is paid on a^c, and period 2 income is fully consumed in either state, married or single.

The first order condition for a is given by:

$$\phi = \beta(1-D)\frac{1}{\theta}u'(c_2^c/\theta)(1+r) + \beta D\left[\lambda u'(c_2^m)(1+r)\alpha + (1-\lambda)u'(c_2^f)(1+r)(1-\alpha)\right]$$

where ϕ is the Lagrange multiplier on the budget constraint, and c_2^c, c_2^m and c_2^f are second period consumption for married, single male and single female households respectively. Using the budget constraint for each period to express in terms of a^c and substituting $\phi = u'(c_1^c)$ gives:

$$\begin{aligned}\frac{1}{\theta}u'(\frac{w_1 - a^c}{\theta}) =& \beta(1-D)\frac{1}{\theta}u'(\frac{a^c(1+r) + w_2^m + w_2^f}{\theta})(1+r) + \\ & \beta D\left[\lambda u'(\alpha a^c(1+r) + w_2^m)(1+r)\alpha + (1-\lambda)u'((1-\alpha)a^c(1+r) + w_2^f)(1+r)(1-\alpha)\right]\end{aligned} \tag{5.1}$$

Assuming standard iso-elastic preferences $u = c^{1-\sigma}/(1-\sigma)$, optimal saving in the absence of divorce risk $(D = 0)$ is given by:

$$\underline{a}^c = \frac{[\beta(1+r)]^{1/\sigma}w_1 - (w_2^m + w_2^f)}{(1+r) + [\beta(1+r)]^{1/\sigma}}$$

which is notably independent of the economies of scale parameter θ.

When $D > 0$, a closed form expression for a is generally not available. However, even in the presence of divorce risk, there are a number of special cases when the optimal saving will correspond with $a^c = \underline{a}^c$. This will occur if $w_2^m = w_2^f$, the asset distribution is equal $(\alpha = \frac{1}{2})$ AND:

- $\theta = 2$, i.e., there are no economies of scale; or

- $\theta \leq 2$ and $\sigma = 1$, i.e., there are economies of scale and preferences are logarithmic.[7]

[7] In the case of $u(c) = \log(\frac{c}{\theta})$, θ does not affect the marginal utility of consumption and therefore has no effect on optimal saving behaviour even in the presence of marital risk. The related analysis in Doepke and Tertilt (2016) is confined to this special case.

These results can be seen by first substituting $\alpha = \frac{1}{2}$ into Equation 5.1 and re-arranging:

$$\frac{1}{\theta}u'(\frac{w_1 - a^c}{\theta}) = \beta(1+r)\Bigg\{\frac{1}{\theta}u'(\frac{a^c(1+r)+2w_2}{\theta}) + D\left[\lambda \cdot \frac{1}{2}u'(\frac{a^c(1+r)+2w_2^m}{2}) + (1-\lambda)\cdot\frac{1}{2}u'(\frac{a^c(1+r)+2w_2^f}{2}) - \frac{1}{\theta}u'(\frac{a^c(1+r)+w_2^m+w_2^f}{\theta})\right] \tag{5.2}$$

which implies that the term in square brackets in Equation 5.2 must be zero in order for $a^c = \underline{a}^c$ if $D > 0$. In the first case ($\theta = 2$) and assuming that $\sigma \neq 1$, this term becomes:

$$D\left[\lambda\cdot(\frac{a^c(1+r)+2w_2^m}{2})^{-\sigma} + (1-\lambda)\cdot(\frac{a^c(1+r)+2w_2^f}{2})^{-\sigma} - (\frac{a^c(1+r)+w_2^m+w_2^f}{2})^{-\sigma}\right] \tag{5.3}$$

which is zero if $w_2^m = w_2^f = \bar{w}$. The second case ($\theta < 2$ and $\sigma = 1$) follows similar logic.

However, the saving choice in the presence of divorce risk will generally differ from $\underline{a}^c$. Notably, this will occur when there are economies of scale in consumption ($\theta < 2$) and $\sigma \neq 1$ even if $w_2^m = w_2^f$ and $\alpha = \frac{1}{2}$. In the more interesting situation *when $w_2^m \neq w_2^f$, saving is at least as high as $\underline{a}'$ if we assume that the distribution of assets and bargaining weights are equal across partners* ($\alpha = \lambda = \frac{1}{2}$), which are the assumptions used in the full life cycle model described below.[8] This follows from the convexity of $u'(\cdot)$, which guarantees that $\frac{1}{2}u'(w_2^m) + \frac{1}{2}u'(w_2^f) \geq u'(\frac{1}{2}w_2^m + \frac{1}{2}w_2^f)$ by Jensen's inequality. Thus Equation 5.3 is non-negative when $\theta = 2$ (no economies of scale), which by Equation 5.2 requires $a^c \geq \underline{a}^c$. When there are economies of scale ($\theta < 2$), by similar logic, the term in square brackets is now strictly positive so Equation 5.2 implies $a^c > \underline{a}^c$. *Thus economies of scale in consumption reinforces the demand for precautionary saving in the presence of divorce risk.*

Problem for a single household

In contrast, the possibility of becoming married next period tends to reduce saving by single households. Without loss of generality, this can be seen by considering the problem of a male household that is single in period 1:

$$\max_{c_1^m, a}\left\{u(c_1^m) + \beta[(1-M)v^m(a) + Mv^c(a^m + a^f)]\right\}$$

subject to

$$c_1^m + a = w_1^m$$

[8]For other combinations of λ and α, divorce risk can increase or decrease saving compared with $\underline{a}^c$. Doepke and Tertilt (2016) show for log preferences that divorce risk increases saving if the asset share exceeds the bargaining weight of the spouse who is made worse off by divorce.

where M is the probability of becoming married next period and a^f is the saving of the prospective female partner. Given a^f, the first order condition for a^m is:

$$u'(w_1^m - a^m) = \beta(1-M)u'(a^m(1+r)+w_2^m)(1+r) + \beta M \frac{1}{\theta} u'(\frac{a^m(1+r)+a^f(1+r)+w_2^m+w_2^f}{\theta})(1+r) \tag{5.4}$$

For $M = 0$ the expression for optimal saving $\underline{a}^m$ is almost identical to the one for the married household. For $M > 0$, the right hand side of Equation 5.4 can be re-arranged to give:

$$\begin{aligned}\beta u'(a^m(1+r)+w_2^m)(1+r) + \beta M[\theta^{\sigma-1}u'(a^m(1+r)+a^f(1+r)+w_2^m \\ +w_2^f) - u'(a^m(1+r)+w_2^m)](1+r)\end{aligned} \tag{5.5}$$

Whether saving is greater or less than $\underline{a}^m$ again depends on the sign of the term in square brackets, which in turn depends on the prospective partner's saving choice a^f. Assuming a symmetric equilibrium (with shared preferences and wages), the single male's saving choice will always be lower than without marital risk. More generally, the single male's optimal saving is decreasing in their prospective partner's wage and the economies of scale in married partners' consumption. Notably, when economies of scale are highest ($\theta = 1$), the single male's saving will be less than $\underline{a}$ as long as $w_2^f > -a^f(1+r)$, which is a relatively weak requirement. *This tendency towards lower saving is magnified when a single household has a lower wage than their prospective partner ($w_2^m < \bar{w}_2$). In summary, the possibility of marriage next period tends to reduce saving, which is essentially a free rider problem.*

Net effect of saving

The effect of introducing marital risk on combined household saving ($a^c + a^m + a^f$) is ambiguous and depends on all parameter values and the fraction of single and married households in the population. However, if we confine our attention to the most empirically-plausible parameter values and assume an equal proportion of people are married and single in period 1, numerical results indicate that marital risk tends to reduce overall saving when wages are not too different between genders. This is shown in Table 5.2.1, which tabulates overall saving for different combinations of marital risk[9] and relative wages between men and women. Saving in each cell of the table is expressed as proportion of saving in an otherwise similar economy without marital risk. Numbers less than 1 therefore indicate that marital risk has the effect of reducing saving. I also assume that the economies of scale factor is given by $\theta = \sqrt{2}$ (as per the McClements scale for a two-adult household) and that $\sigma = \frac{3}{2}$ as in the calibrated life cycle model below.[10] Saving in the absence of marital risk is the numeraire. As can be seen, total saving is lower than the corresponding equilibrium with no marital risk except where men and women's wages are very unequal.[11]

[9] In doing so, I constrain the marriage rate to be twice the divorce rate. If anything this is an understatement – age-specific marriage rates tend to be more than twice the divorce rate according to CPS data.

[10] The Pareto weight λ and the asset distribution share α are both assumed to be $\frac{1}{2}$.

[11] I am also ignoring any general equilibrium effects that saving has on the interest rate.

Table 5.2.1: Combined married and single saving

	Divorce rate (D =1/2 M)									
w^m/w^f	0.00	0.03	0.07	0.10	0.13	0.17	0.20	0.23	0.27	0.30
0.67	1.00	0.95	0.89	0.83	0.77	0.70	0.62	0.55	0.46	0.37
0.83	1.00	0.92	0.83	0.74	0.65	0.55	0.44	0.33	0.21	0.09
1.00	1.00	0.91	0.81	0.71	0.61	0.50	0.38	0.26	0.13	-0.01
1.17	1.00	0.92	0.83	0.74	0.65	0.55	0.44	0.33	0.21	0.09
1.33	1.00	0.95	0.89	0.83	0.77	0.70	0.62	0.55	0.46	0.37

Values are expressed as a fraction of saving in the no marital risk equilibrium; the probability of marriage is fixed at two-times the probability of divorce.

5.2.2 Dual labour supply and partial insurance

There is a growing literature exploring the mechanisms by which households are insured against wage and other income shocks in the absence of formal insurance markets. These mechanisms include family labour supply, social security and the progressivity of the tax system in addition to the much-researched mechanism of precautionary saving (see Blundell (2014) for a survey).[12] By introducing two-person married households, this paper focuses on the impact of a dual labour supply margin on the dynamics of household saving and consumption, and what this means for optimal taxation. In order to understand the impact that dual labour supply might have on optimal taxation, it is worth emphasising some key aspects of the optimal tax literature. First, early life cycle studies that ignored wage uncertainty and elastic labour supply tended to find large welfare gains from eliminating capital taxation. The most prominent example is Summers (1981) who shows that the interest elasticity of aggregate saving in such models is relatively large meaning that there are substantial welfare gains from eliminating capital taxation. However, subsequent studies have shown that the magnitude of this elasticity is very sensitive to key model parameters. For example, Evans (1983) finds that using different values for key parameters such as the EIS and the rate of time preference can substantially reduce the interest elasticity of saving and reverse the main conclusion in Summers (1981). The result in Summers (1981) is also sensitive to various features of the model environment. This includes the presence of idiosyncratic wage uncertainty, which creates a precautionary saving motive in addition to the purely life-cycle saving motive present in Summers's deterministic OLG model. Whereas life-cycle saving is by its nature sensitivity to changes in the inter-temporal price of consumption, Engen and Gale (1997) and Cagetti (2001) show that the presence of precautionary saving reduces the interest rate elasticity of saving. In both cases, the large welfare gains from eliminating capital taxation found by Summers (1981) disappear or are at least greatly diminished.

However, the picture is more complicated again when households can freely adjust their labour supply. Low (2005), Flodén (2006) and Pijoan-Mas (2006) show that, in the presence of wage uncertainty, precautionary motives also give rise to longer work hours, especially when households are young with low assets. Higher hours worked *before* shocks are realised provides households with more income, which reduces the cost of precautionary saving. At the same time, Low (2005) shows that the ability to adjust

[12]Blundell (2014) surveys the 'partial insurance' models – the ways in which households insure themselves against idiosyncratic risk. I use the term *dual* labour supply throughout this paper as a specific case of the 'family' labour supply channel used in Blundell et al. (2016).

hours worked *after* shocks have been realised makes precautionary saving less valuable, meaning the overall impact on precautionary saving is ambiguous. However, it seems reasonable to conjecture that elastic labour supply will make saving more elastic to interest rate changes than if labour supply were fixed. Further, the introduction of a *dual* labour supply margin should make married households' saving even more interest rate elastic, a proposition that I demonstrate numerically in Section 5.6.[13] In essence, I show that the additional capacity to self-insure via labour supply adjustments makes an economy with married households look more like a complete markets setting. In such a setting, households can focus on substituting their hours worked towards favourable productivity states (Pijoan-Mas, 2006). I demonstrate this important channel by computing the within-age correlation between productivity and hours worked, highlighting the much greater correlation for married households.

5.2.3 Empirical evidence on saving by marital status

The results from the model presented in Section 5.2.1 suggest that married households will save proportionally more of their income than single households. This reflects that being married is a better state than being single because of the economies of scale in consumption. This leads to precautionary saving during marriage by married households and a free rider problem for single households which lowers their incentive to save. Moving from the two period model to the full OLG model (presented in the next section) is likely to strengthen this tendency. This is because the addition of a dual labour supply margin for married households increases the desirability of being married, a point that I elaborate upon in 5.6.3. The stochastic process for marital status in the OLG model means that being single and married are transient states over the life cycle. The desire of households to smooth their consumption between these divergent income states induces higher saving while married and lower saving while single.

These theoretical predictions are supported by the small number of empirical studies that have examined these questions. Firstly, household level data show that married households do in fact save at more than twice the rate of single households, on average. In the United States, this is shown by Lupton and Smith (1999) using the United States Panel Survey of Income Dynamics, while similar patterns have been found using other data sets and in other countries (e.g., see Zagorsky (2005)). Perhaps more importantly, there is also empirical support for the specific economic mechanisms behind these differences in saving rates. Kureishi and Wakabayashi (2013) show that the saving rates of single households with a high probability of being married within three years are much lower than for single households with a low probability. This difference is attributed to both precautionary and retirement saving motives. In relation to the impact of divorce risk on saving, multiple studies have shown that this risk leads to higher saving by married households. This includes Finke and Pierce (2006) who use the Panel Survey of Income Dynamics (PSID) to show that spouses who end up getting divorced have a substantially higher saving rate in prior years than spouses that remain married. Also using PSID data, Voena (2015) finds that the introduction of unilateral divorce laws with equal division of assets led couples in those State jurisdictions to save

[13] However, as noted above, the introduction of divorce risk provides a distinct motivation for precautionary saving by married households.

more than couples in other States. Similarly, González and Özcan (2013) use a difference-in-difference methodology to examine the effect of divorce law changes in Ireland and find that the increased divorce risk resulted in higher (precautionary) saving by married households. Finally, the Italian panel study by Pericoli and Ventura (2012) develops a probabilistic measure of divorce and shows that this variable is significantly correlated with higher rates of saving among married households because of precautionary motives.

While there is some evidence on the partial effect of marriage and divorce risk on saving, there is little evidence on the *net* effect of both sources of risk. This is an inherently difficult thing to empirically quantify, as it depends on the saving response of both single and married households to both sources of marital risk *and* the proportion of married and single households in the population. In reality, age-specific probabilities of both divorce and marriage have exhibited large changes in recent decades, which has flowed through to trend movements in the shares of marital status in the population. These sources of non-stationarity compound the difficulty in isolating the general equilibrium effect on saving of higher marriage risk.

Nonetheless, Figure 5.4.1 shows that the baseline OLG model is at least able to endogenously generate the main empirical regularity noted above, namely, that married households save more than single households on a per adult basis.[14]

5.3 The model

Demographics and household transitions

The model is populated by an equal mass of males ($g = m$) and females ($g = f$) who enter the economy at $j = 1$ and live a maximum of J years. While alive, both genders face an age-specific probability ψ_j of surviving from age j to $j + 1$.

At each age, a fraction of each gender π_j is in a marriage with marital status $ms_j = c$ while the remaining fraction is single with marital status $ms_1 = s$. A fraction π_1 of individuals begin life married and $1 - \pi_1$ begin life single, and thereafter each individual's marital status evolves exogenously over their lifetime as a first-order Markov process $Q(ms'|ms)$: single households are subject to a potentially age-varying marriage rate M_j and similarly married households face a divorce rate D_j. When a marriage ends in divorce, assets are divided among the divorcees with $\alpha \in (0, 1)$ share going to the man and $1 - \alpha$ to the woman.[15] New marriages take place randomly among newlyweds both of age j drawn from their gender's asset distribution for that age. This process is understood by households and incorporated rationally into their decision rules.

[14] Also of note, the model is able to capture the empirical observation that assets are an increasing function of marriage duration (Lupton and Smith, 1999).

[15] In other words, I assume an 'equitable division' rather than a title-based regime. See Doepke and Tertilt (2016) for a survey of the area of macroeconomics and families.

Mortality operates at the household level, so that both members of a married couple die together. This restriction is imposed for tractability and means that the population contains no widows. The assets a of those who die (that is, accidental bequests) are redistributed as a lump-sum B to all currently living individuals.

Individual labour productivity

The average productivity of workers aged j is governed by an efficiency index e_j, which is allowed to vary by a scale factor $\gamma_{g,ms}$ for each gender-marital status combination. Around these deterministic age-trends, an individual's productivity evolves as an auto-regressive process z_j in which shocks are independently and identically distributed among workers. Pre-tax earnings y_j are then given by the product of these factors, along with the wage (per efficiency unit) w and the individual's labour supply n_j:

$$y_j = \gamma_{g,ms} e_j z w n_j \tag{5.6}$$

All agents retire at j_r, and retirement is an absorbing state. Each individual is therefore summarised by the state vector (a, j, z, g, ms) in the joint state space $\mathcal{A} \times \mathcal{J} \times \mathcal{Z} \times \mathcal{G} \times \mathcal{MS}$. I denote $\phi_g^{ms}(a, j, z)$ as the measure of individuals for each combination of gender ($g = m, f$) and marital status ($ms = s, c$).

Preferences

Individuals are endowed with one unit of time per period, which they allocate between labour n and leisure l during their working age years. Married couples maximise joint utility, which is separable in consumption c and both partners' leisure. Within the household, consumption is assumed to be a public good with possible congestion, where θ regulates the degree of congestion. Lastly, the Pareto weight λ determines the relative importance of each partner's leisure in joint utility, which gives the following functional form for a married couple's utility:

$$u^c = \frac{(c/\theta)^{1-\sigma}}{1-\sigma} + \lambda \chi_m^c \frac{l_m^{1-\eta}}{1-\eta} + (1-\lambda)\chi_f^c \frac{l_f^{1-\eta}}{1-\eta} \tag{5.7}$$

Analogously, a single household's momentary utility function for each gender g is given by:

$$u_g^s = \frac{c^{1-\sigma}}{1-\sigma} + \chi_g^s \frac{l_g^{1-\eta}}{1-\eta} \tag{5.8}$$

While the empirical support for separable preferences is mixed (e.g. Blundell et al. (2016) and Chetty (2006)), the decision to use this form of preferences is made partly to ensure that the computational exercise remains tractable. Moreover, most of the related literature (including Peterman (2013) and Conesa et al. (2009)) have used this functional form in either their main model or sensitivity analysis, which permits a comparison of results across studies.

Government

The government administers a compulsory social security system into which workers pay τ_{SS} fraction of their annual earnings, and retirees receive a gender- and marital status-specific pension $P_{g,ms}$ based on a replacement rate Ψ.[16] The tax rate τ_{SS} adjusts to ensure that the social security system is always in balance. The government undertakes unproductive consumption spending G which it finances by taxing adjusted-earnings $\tilde{y}_j = y_j(1-\tau_{SS}/2)$ according to a possibly progressive tax function $T(\cdot)$, and through a flat tax τ_c on household consumption and a flat tax τ_k on capital income ar. In the main policy experiment below, $T(\cdot)$ is applied separately to each member of the married household's labour income. In other words, the main experiment entails a shift from joint to separate filing, although I also consider a joint filing regime in the sensitivity section. Finally, there is a separate tax τ_B levied on accidental bequests.

Technology

The production side of the economy is standard, comprising a representative firm operating with Cobb-Douglass technology:

$$Y_t = A_t K_t^{\alpha} N_t^{1-\alpha} \tag{5.9}$$

where A_t, N_t and K_t represent aggregate productivity (set to 1 as the numeraire), labour supply and capital stock respectively. The latter is subject to a depreciation rate δ.

Market structure

While households do not have access to formal insurance markets, they can self-insure themselves against wage and marital risk through their holdings, a, in 1-period bond, which pays an interest rate r. However, this insurance is limited by the requirement that a be non-negative at all times.

Competitive equilibrium

With this form of preferences, the household's problem can be presented recursively. For a single household the problem is:

$$v_g^s(a,z,j) = \max_{c,n,a'} \left\{ u_g^s(c,l) + (1-M_j)\psi_j \beta \mathrm{E}_{z,\phi_{g*}^s(a,j)} v_g^s(a',z',j+1) + M_j \psi_j \beta \mathrm{E}_z v^c(a',z',j+1) \right\}$$

[16] Consistent with Conesa et al. (2009) and most other related studies, an individual's pension is unrelated to his or her unique earnings history. This assumption facilitates tractability but it would be interesting to investigate the effect of more realistic set ups.

subject to

$$
\begin{aligned}
c(1+\tau_c)+a' &= (1+r(1-\tau_k))a + y_{g,s}(1-\tau_{SS}) + B - T(\tilde{y}_{g,s}), \text{ if } \quad j <= j_r \\
c(1+\tau_c)+a' &= (1+r(1-\tau_k))a + B + P_{g,s}, \text{ if } \quad j > j_r \\
\tilde{y}_{g,s} &= (1-\tau_{SS})y_{g,s} = (1-\tau_{SS})\gamma_{g,s} ezwn_g \\
a \geq 0,\ n \geq 0,\ g = m, f
\end{aligned}
$$

and the expectation operator $\mathrm{E}_{z,\phi^s_{g*}(a,j)}$ indicates that a single household's saving decision takes into account the distribution of assets among the opposite gender $g*$, which is the pool of potential spouses in period $j+1$. The married household's problem is:

$$
\begin{aligned}
v^c(a,z,j) = \max_{c,n_m,n_f,a'} \Big\{ & u^c(c,l_m,l_f) + (1-D_j)\psi_j \beta \mathrm{E}_z v^c(a',z',j+1) + \\
& D_j \psi_j \beta \mathrm{E}_z \left[\lambda v^s_m(\frac{a'}{2}, z', j+1) + (1-\lambda) v^s_f(\frac{a'}{2}, z', j+1) \right] \Big\}
\end{aligned}
$$

subject to

$$
\begin{aligned}
c(1+\tau_c)+a' &= (1+r(1-\tau_k))a + y_m(1-\tau_{SS}) + y_f(1-\tau_{SS}) + 2B - T(\tilde{y}_m) - T(\tilde{y}_f), \text{ if } \quad j <= j_r \\
c(1+\tau_c)+a' &= (1+r(1-\tau_k))a + B + P_{m,c} + P_{f,c}, \text{ if } \quad j > j_r \\
\tilde{y}_{g,c} &= (1-\tau_{SS})y_{g,c} = (1-\tau_{SS})\gamma_{g,c} ezwn_g \\
a \geq 0,\ n \geq 0,\ g = m, f
\end{aligned}
$$

2. The factor prices w and r satisfy:

$$
r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \quad \text{and} \quad w = (1-\alpha)\left(\frac{N}{K}\right)^{\alpha}
$$

3. The social security system is in balance:

$$
\tau_{ss} wN = \sum_g \sum_{ms} P_{g,ms} \int \phi^{ms}_g (da \times dz \times dj)
$$

4. Transfers equal after-tax accidental bequests:

$$
B = (1-\tau_B) \sum_g \sum_{ms} \int (1-\psi_j) a'(a,z,j,g,ms) \phi^{ms}_g (da \times dz \times dj)
$$

5. The Government's budget is in balance:

$$
\begin{aligned}
G = & \tau_B \sum_g \sum_{ms} \int (1-\psi_j) a'(a,z,j,g,ms) \phi_g^{ms}(da \times dz \times dj) + \sum_g \sum_{ms} \int \tau_k a r \phi_g^{ms}(da \times dz \times dj) \\
& + \sum_g \sum_{ms} \int T(\tilde{y}) \phi_g^{ms}(da \times dz \times dj) + \sum_g \sum_{ms} \int \tau_c c(a,z,j,g,ms) \phi_g^{ms}(da \times dz \times dj)
\end{aligned}
$$

6. Markets clear:

$$
\begin{aligned}
K = & \sum_g \left[\int \frac{a}{2} \phi_g^c(da \times dz \times dj) + \int a \phi_g^s(da \times dz \times dj) \right] \\
N = & \sum_g \sum_{ms} \int \gamma_{g,ms} e_j z n(a,z,j,g,ms) \phi_g^{ms}(da \times dz \times dj)
\end{aligned}
$$

$$
\begin{aligned}
& AK^\alpha L^{1-\alpha} + (1-\delta)K \\
& \quad = G + \sum_g \left[\int \frac{c}{2} \phi_g^c(da \times dz \times dj) + \int c \phi_g^s(da \times dz \times dj) \right] \\
& \qquad + \sum_g \left[\int \frac{a'(a,z,j,g,c)}{2} \phi_g^c(da \times dz \times dj) + \int a'(a,z,j,g,s) \phi_g^s(da \times dz \times dj) \right]
\end{aligned}
$$

7. The share of single and married households in the population is consistent with the survival-adjusted measure of individuals in the state space $\mathcal{A} \times \mathcal{Z}$ and is symmetric by gender for all j:

$$
\begin{aligned}
\pi_{j,g} &= \int \phi_g^c(da \times dz \times \{j\}) / \prod_{q=1}^{j} \psi_q \\
1 - \pi_{j,g} &= \int \phi_g^s(da \times dz \times \{j\}) / \prod_{q=1}^{j} \psi_q \\
\pi_{j,m} &= \pi_{j,f}
\end{aligned}
$$

8. Law of motion:

$$
\phi_g^{ms} = \mathcal{P} \phi_g^{ms}
$$

where $\mathcal{P}$ is a one-period recursive operator that incorporates the joint processes for marital status Q, idiosyncratic productivity z and mortality.

5.4 Calibration

The key calibrated parameters are contained in Table 5.4.1.

Demographics

Individuals enter the economy at a real life age of 20 ($j = 1$) and live to a maximum age of 99 ($J = 80$). The survival rate ψ_j is based on the mean male and female death rates. Population growth n is set to zero. There is no data readily available for annual marriage and divorce events by age. Therefore I rely on marital transition probabilities derived in a previous study (Cubeddu and Rios-Rull, 1997), while also maintaining consistency with the shares of married and single households by age contained in the Current Population Survey (CPS).[17]

Preferences

The parameter governing relative risk aversion, σ is set to 1.5. This implies an elasticity of inter-temporal substitution (EIS) of $\frac{2}{3}$, which is slightly higher than the $\frac{1}{2}$ assumed in Conesa et al. (2009) but is arguably more consistent with recent panel data evidence (Alan et al., 2009) and with the plausible upper bound on relative risk aversion derived by Chetty (2006). The parameter η is chosen to achieve an unconstrained intensive Frisch labour supply elasticity of $\frac{2}{3}$. This is towards the upper end of conventional estimates of the Frisch elasticity based on micro data samples of prime-age men, but is arguably consistent with more recent studies that include women (see Keane (2011) for a review of the empirical labour supply literature).[18] For married couples, the congestion factor for consumption θ is set to 1.42 consistent with the McClements scale for two-adult households. The values for $\chi_{g,ms}$ are chosen to be consistent with the relativity in work hours between men and women in the CPS. In particular, I target the work hours of single women to be 93 per cent of single men, and the work hours of married women to be 86 per cent of married men. I also target an overall mean work time ($\bar{n}$) of $\frac{1}{3}$, the value used in the vast majority of past studies.

Government policy

In the baseline economy, government consumption is equivalent to 17 per cent of GDP, and is funded through a progressive tax on the sum of labour and capital income, which is levied at the household level consistent with joint filing. This is done according to the Gouveia-Strauss tax function:

$$T(\tilde{y}, \kappa_0, \kappa_1, \kappa_2) = \kappa_0(-(\tilde{y}^{\kappa_1} + \kappa_2)^{-1/\kappa_1}) \tag{5.10}$$

where the values for κ_0 and κ_1 are based on the estimates from their 1994 paper (Gouveia and Strauss, 1994). The remaining parameter, κ_2, is adjusted to balance the government's budget. The flat tax on household consumption is set $\tau_c = 0.05$, consistent with available data and comparable studies. Redistributed bequests are not included in the capital income tax base and are instead taxed separately at a rate $\tau_B = 0.5$ to avoid the confounding results of taxing them together demonstrated by Peterman (2013).[19] The replacement rate for social security pensions is set at 50 per cent of average earnings for each of the four household types.

[17] I fit a simple log function to the age profile in Cubeddu and Rios-Rull (1997), so that marriage rate falls log-linearly from 0.25 at age 23 to 0.1 at age 99. Similarly, the divorce rate begins at 0.15 and falls to 0.01 by age 99.

[18] With these preferences, the intensive Frisch elasticity depends in general on the value of n. The elasticity of $\frac{2}{3}$ is based on a mean value of n_j of $\frac{1}{3}$.

[19] Peterman (2013) shows that taxing the lump-sum B is non-distorting, so the 'optimal' tax on the combined $ar + B$ is always higher than the optimal rate on ar alone.

Individual productivity

I use PSID data on average hourly wages by age, gender and marital status as targets for calibrating $\gamma_{g,ms}$. Similarly I use CPS data on mean annual hours by gender and marital status to calibrate the preference parameters $\chi_{g,ms}$.

I approximate the stochastic process for individual productivity using a first-order Markov process. The underlying $AR(1)$ process that I seek to mimic contains a number of parameters that need to be chosen for each household type: the autoregressive parameter $\rho_{g,ms}$, the variance of the shock $\sigma^2_{\eta,g,ms}$, and the initial variances $\text{Var}(z_{1,g,ms})$. For married households I also need to set the value of ρ_{mf} in the bi-variate $AR(1)$ process, which governs the correlation between the husband's and wife's productivity processes.

My choices attempt to match a number of empirical moments in the wages data. Firstly, I target the observed life-cycle pattern of wage inequality described in Storesletten et al. (2004). However, to achieve this I must take account of the impact of marriage transitions on the life-cycle evolution of wages for each household type. I do this by simulating the processes for marriage and z jointly then targeting the resulting life-cycle wage profiles by varying $\sigma^2_{\eta,g,c}$, $\text{Var}(z_{1,g,ms})$ and ρ_{mf}. In particular, I target an initial (at age 20) variance of 0.15 for log wages, and a total increase over the working-age period of around 0.6 log points. For married households, I target an average correlation in partners' wages of 0.4 to 0.5 over the working age years consistent with Hyslop (2001).[20] By adjusting the initial conditions of the AR(1) process I am also able to incorporate an upward trend in this correlation over the life cycle, consistent with the empirical evidence on the intra-family covariance in wage shocks (see for example Shore (2015) and Krueger and Wu (2018)).

Key features of the baseline economy

Before proceeding to the policy experiment, I highlight some features of the baseline economy in order to provide context for the optimal tax results below. First, the discount rate ($\beta = 0.995$) needed to achieve the target capital-to-output ratio of 2.92 is higher than in a otherwise similarly-calibrated bachelor model ($\beta = 0.986$). This largely reflects the presence of marital risk, which reduces the motivation for single households to save because of a free rider problem.[21]

The different saving incentives between single and married households can be seen in the first two panels of Figure 5.4.1, which show the mean life cycle profiles for assets and consumption. Married households accumulate greater assets than single households and this is accompanied by a consumption profile that peaks at a later age. More generally, the cost of precautionary saving, and therefore the overall profile for life cycle assets, will depend on the extent to which married households use labour supply to self insure early in life. This in part depends on the correlation between partners' wages in different states of the

[20] As Hyslop (2001) notes, part of the overall correlation in partners' wages reflects assortative matching. My calibration is therefore a simple way of incorporating the effects of assortative matching without modelling it explicitly.

[21] As I show below, the disincentive to save from the free rider effect is quantitatively more important than the precautionary saving motive induced by wage uncertainty.

Table 5.4.1: Calibration

	Value	**Target**	**Source**
Demographics			
Lifespan (real age) (J)	79 (99)	Assumed	
Retirement age (real age) (j_r)	45(65)	Assumed	
Survival probabilities (ψ_j)	-	Data	Bell and Miller (2002)
Marriage and divorce rate (M_j,D_j)	-	Data: marital status by age	CPS
AR(1) productivity process			
Married			
Correlation (ρ_{mf})	0.04	Partner wage corr.=0.4-0.5	Hyslop (2001)
Women wage ratio ($\gamma_{c,f}/\gamma_{c,m}$)	0.6	Data	CPS
Single			
Persistence parameter (ρ_s)	0.98	Data: earnings by age	Storesletten et al. (2004)
Variance of AR(1) process (η_s)	0.025	Data: earnings by age	Storesletten et al. (2004)
Women wage ratio ($\gamma_{s,f}/\gamma_{s,m}$)	0.8	Data	CPS
Household preferences			
Discount factor (β)	0.995	K/Y=2.92	NIPA
Relative risk aversion (σ)	1.5	Assumed	Micro evidence
Paramater governing Frisch (η)	3	Frisch ELS=2/3	Micro evidence
Consumption congestion factor (θ)	1.4023		McClements scale
Disutility of labour			
Male married (χ_m^m)	2.45	$\bar{n}_{f,c}/\bar{n}_{m,c} = 0.86$	CPS
Female married (χ_m^f)	1.75	$\bar{n}_{f,c}/\bar{n}_{m,c} = 0.86$	CPS
Male single (χ_s^m)	3.3	$\bar{n}_{f,s}/\bar{n}_{m,s} = 0.93$	CPS
Female single (χ_s^f)	3.6	$\bar{n}_{f,s}/\bar{n}_{m,s} = 0.93$	CPS
Technology			
Capital share (α)	0.36	Data	NIPA
Depreciation (δ)	0.0833	I/Y=0.255	NIPA
Government			
Replacement rate (Ψ)	0.5	Assumed	Past studies
Social Security Tax Rate (τ_{ss})	0.1651		Endogenous
Government consumption (G)	G/Y=0.18	Data	NIPA
Flat consumption tax (τ_c)	0.05		Mendoza et al (1994)
Flat capital tax on capital (τ_k)	0 in baseline		
Income tax function			
Marginal tax parameter (κ_0)	0.258	Data: tax payments	Gouveia and Strauss (1994)
Tax progressivity parameter (κ_1)	0.768	Data: tax payments	Gouveia and Strauss (1994)

world, which I calibrate to around 0.4 in the baseline. Below I explore the two extreme cases of zero and perfectly correlated wage shocks, and show that the optimal tax system is very sensitive to this single parameter.

Another important aspect of the baseline economy is that hours worked decline monotonically for all household types over the life cycle (see Figure 5.4.1), implying that the optimal age-dependent labour tax (if it could be implemented) would tend to have a similar profile (Erosa and Gervais, 2002). However, the mean marginal income tax rate in the baseline tax system does not do a very good job of mimicking this theoretical optimum as I elaborate upon below, instead having a very flat profile over the life cycle. This flatness reflects that the progressive tax function in the baseline economy is based on a household's combined labour and capital income, and the joint filing requirement for married households. A similar observation is made in Gervais (2012).

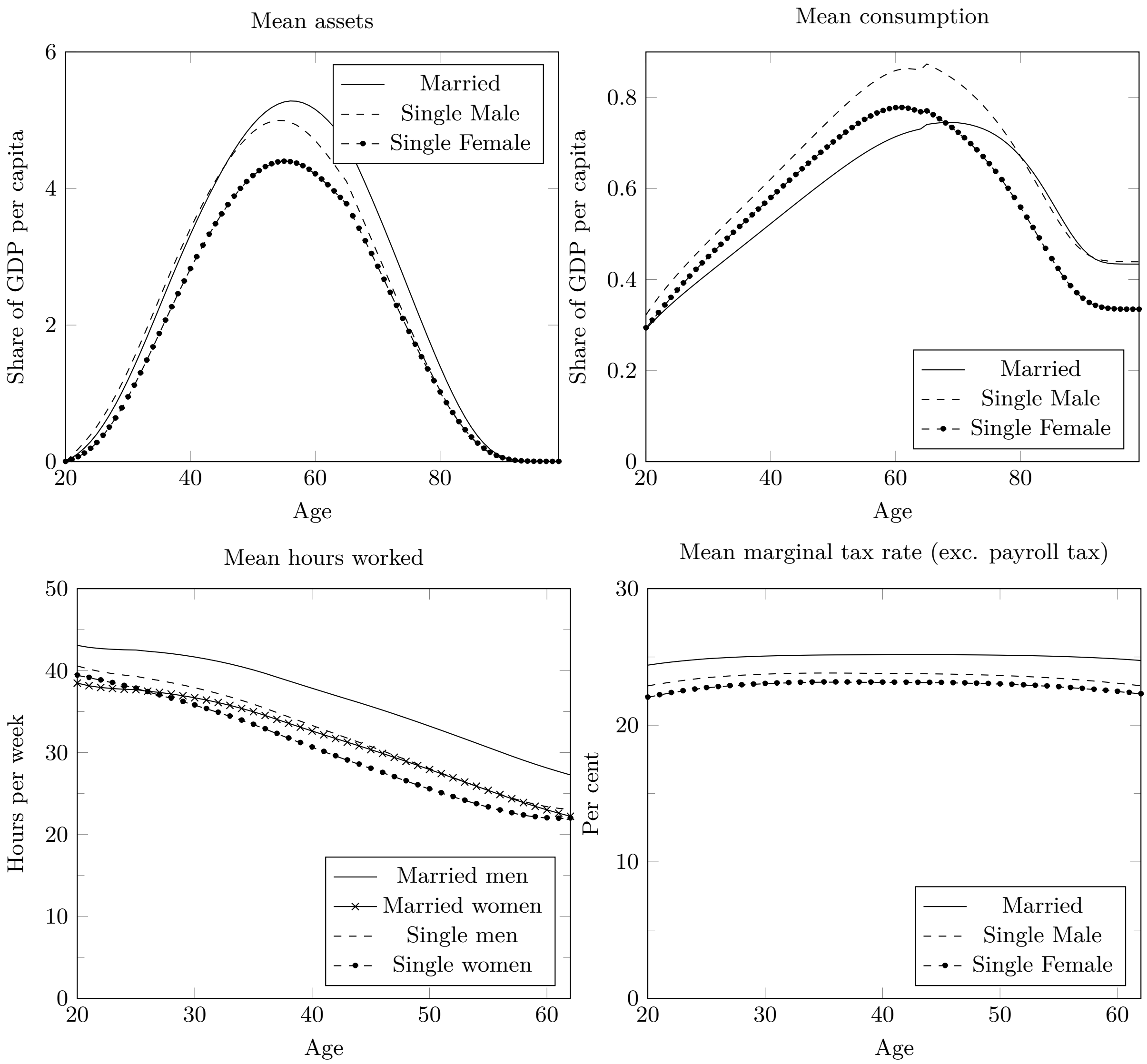

Figure 5.4.1: Mean life cycle profiles for assets, consumption, hours worked and marginal tax rates

5.5 The policy experiment

Following the related literature, the main experiment consists of switching from the current tax system (in which the sum of labour and capital income are taxed progressively) to a system that features a (possibly progressive) tax on labour income and a separate flat tax on capital τ_k. The computatishiional exercise is therefore to choose a combination $\{\kappa_0^\star, \kappa_1^\star, \tau_k^\star\}$ that maximises social welfare $SW(\kappa_0, \kappa_1, \tau_k)$ subject to parameters and government consumption being fixed at their baseline values. The parameter κ_2 is used to balance the budget and I rule out tax systems where no such κ_2 exists. In the case of a flat tax ($\kappa_1 \to \infty$) with deduction, κ_2 determines the deduction consistent with a balanced budget. I further assume anonymity of the tax code, rule out age-dependent and negative tax rates, and restrict the tax function's parameters to be the same across gender and marital status. In the main experiment, I consider a tax system based on separate filing although I later relax this restriction in the sensitivity analysis. SW is given by the expected lifetime utility of a newborn individual weighted by the distribution of the initial productivity shock and the proportion of households of each type.

5.6 Results and discussion

5.6.1 The optimal tax system in the benchmark economy

I find an optimal tax mix $\{\kappa_0^\star, \kappa_1^\star, \tau_k^\star\} = \{0.31, \infty, 0\}$. Thus I find that *the optimal long-run tax on capital is zero* as per the classic Chamley-Judd result but contrary to the finding of recent life cycle studies. In relation to the taxation of labour income, the optimal parameters imply a flat tax of 31 per cent combined with a deduction equivalent to 16 per cent of mean income.

As a reference point, I also compute the optimal tax mix in a 'bachelor' version of the economy, a specification that basically corresponds to the models used in Conesa et al. (2009) and Fehr and Kindermann (2015). Broadly consistent with these studies, I find that the optimal tax parameters in a bachelor economy are $\{\kappa_0^\star, \kappa_1^\star, \tau_k^\star\} = \{0.26, \infty, 0.31\}$ with a deduction on labour earnings equivalent to 20 per cent of average income.[22] Therefore, compared with the case of a bachelor economy, the optimal tax structure in this paper's benchmark model involves a higher tax burden on labour income via both a higher tax rate and smaller deduction, which makes up for the much lower (zero) tax on capital.

Moving to the optimal tax system would yield an *ex ante* expected lifetime welfare gain for a newborn household of 1.2 per cent measured in consumption equivalent variation (CEV). Conversely, assuming the 'wrong' model and imposing the optimal tax parameters from the bachelor economy would *reduce* welfare by about 0.3 per cent. As can be seen in Table 5.6.1, the welfare gain from moving to the optimal tax system reflects a fall in mean hours worked, balanced by a fall in mean consumption. As discussed further below, part of the welfare gain stems from reduced life cycle consumption inequality, made possible by better targeting of progressivity in the taxation of labour income and the shift to separate filing which

[22] For separable preferences, Conesa et al. (2009) find a tax on capital of 21 per cent along with a flat tax on labour income of 34 per cent and a deduction of about 20 per cent of GDP per person. One reason for the slight differences with the bachelor results could be the surprisingly large and persistent earnings shock assumed in Conesa et al. (2009) for the separable preferences case, namely $\sigma_\eta^2 = 0.0841$ and $\rho = 0.995$. These parameters differ from those used in the main (non-separable preferences) section of their paper and the parameters I use in my benchmark model, both of which are based on Storesletten et al. (2004).

allows married households to use their dual labour supply margin to exploit the stochastic variation in their joint productivity realisations.

Table 5.6.1: Optimal tax results

	$\bar{n}$	N	K	Y	K/Y	C	CEV
Baseline (level)	33.12	0.35	1.86	0.64	2.92	0.37	-
Optimal (% change)	-0.82	-2.51	-0.05	-1.63	1.61	-2.61	1.150
Bachelor optimum (% change)	1.66	-0.90	-7.48	-3.32	-4.30	-2.33	-0.267

5.6.2 Sensitivity

To further explore why the optimal tax system in the benchmark model differs from that in the standard bachelor economy I now undertake an extensive sensitivity check. This involves shutting down key elements of the benchmark model and re-computing the optimal tax structure in each of these versions of the model. In each case, the baseline economy is re-calibrated (for relevant variables) to achieve the same target values in the benchmark economy. The results are shown in Table 5.6.2 where the first two columns show the results already discussed for the benchmark and bachelor versions of the economy, while the remaining columns show the features and optimal tax results for the alternative models A1 to A7.

Table 5.6.2: Optimal tax sensitivity analysis

	Bench.	Bach.	A1	A2	A3	A4	A5	A6	A7
Dual-earners	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Marital risk	Yes	-	No	Yes	Yes	Yes	Yes	No	Yes
Wage uncertainty	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Tax filing in optimal	Single	Single	Single	Single	Single	Single	Single	Single	Joint
Endog. interest rate	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Fixed prod. diff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Corr. partners' wages	0.4	-	0.4	0.4	0	1	-	-	0.4
Discount factor (β)	0.995	0.983	0.9868	0.995	0.999	0.972	1.023	0.995	0.995
Interest rate ($r(1-\tau_K)$)	3.8	3.18	3.1	4.01	3.78	2.91	3.34	3.09	4
Flat tax - labour (τ_L)	31	26	27	33	31	22	19	16	35
Deduction (% GDP/head)	16	20	24	18	17	14	0	0	26
Flat tax - capital (τ_K)	0	31	32	0	0	32	0	24	0

This exercise provides a number of key insights. First, as I argued above, the presence of marital risk is crucial to the zero capital tax result. Model A1 turns off this feature and keeps all other aspects of the model unchanged, so that the model still includes both married households and single households *but marital status is now fixed over an individual's lifetime.*[23] The share of married and single households is re-calibrated to match the aggregate share across all ages. This produces an optimal tax on capital that is roughly the same as in the bachelor model, namely, about 32 per cent. Next Model A2 shows that the benchmark result of a zero capital tax carries over to an 'open economy' where the interest rate is now held constant at its baseline value.

[23] Having a fixed proportion of married and single households greatly reduces the computational challenge. Many previous studies that have incorporated family structure have therefore opted to ignore marital risk. The recent study by Nishiyama (2019) is a case in point.

Model A3 shows that the presence of marital risk interacts with the dual labour supply margin, the two novel features in this paper. In the extreme case when partners' wages are uncorrelated (rather than the coefficient of correlation of 0.4 used in the benchmark model), the benefits of family labour supply are significantly larger. This makes being married an even more desirable future state for currently single households, further reducing their saving and necessitating an even higher calibrated discount factor to achieve the target capital-to-output ratio. Married households now have a greater willingness to exploit variation in productivity across states. As noted in the following sub-sections, this makes precautionary saving – and therefore aggregate saving – more interest rate elastic (see Table 5.6.3) as well as inducing a stronger within-age correlation between married households' wages and hours worked. These factors strengthen the zero capital tax result. At the other extreme, Model A4 imposes a perfect correlation in partners wages (i.e. a correlation of unity), which results in the optimal tax on capital reverting back to the bachelor optimum of around 32 per cent. This reflects a reversal of the factors at play in A3 with a much lower interest rate elasiticty of saving (see Table 5.6.3) and reduced within-age correlation between wages and hours. This necessitates a much lower discount factor to hit the capital-to-output target ratio, which provides greater welfare gains from taxing saving and inducing flatter consumption profiles. Model A5 eliminates wage uncertainty, which has no effect on the zero capital tax result if marital risk is retained. This model is essentially an extension of the 2-period model studied in Section 5.2.1, with the presence of marital risk reducing the tendency for saving. Without a motive to undertake precautionary saving to buffer against wage risk, this means an even higher discount factor is needed to reach the target capital-to-output ratio. Also, in the absence of wage risk, the optimal taxation of labour income is no longer progressive, which is unsurprising. Model A6 has neither marital nor wage risk, which results in an optimal tax on capital that is about 24 per cent. Again, this highlights the importance marital risk and its interaction with dual labour supply in contributing to the zero tax on capital result. Model A7 retains joint filing among married couples instead of imposing a shift to separate filing as in the previous experiments. The optimal tax on capital continues to be zero, indicating that this paper's main result is robust to the way in which married couples file their tax returns. In a final sensitivity check (not shown in the table), I experiment with taxing the return on bequests at the same rate as capital income τ_k but keep all other features the same in the benchmark economy. The optimal tax on capital remains zero, although the optimal flat tax on labour income is slightly higher and the deduction slightly smaller.

5.6.3 Dual labour supply and insurance

So what impact does dual labour supply have in this model? First and foremost, the presence of dual labour supply allows married households to exploit the stochastic variation in wages, which is welfare enhancing. This mechanism is similar to the one noted in Pijoan-Mas (2006), namely, married households can readily adjust their hours worked in line with variations in their joint productivity realisations. This is in contrast to asset-poor single households who must work even when their productivity is low for precautionary reasons. This can be seen in Figure 5.6.1 which plots the correlation between within-age productivity and hours worked under the optimal benchmark economy.[24] The labour supply of married women is most sensitive to variations in wages while single men are least sensitive. Clearly, the value to married households of dual labour supply is a function of the (lack of) correlation between partners' wages. In the benchmark model, which is calibrated to the empirically observed correlation, the results show that the presence of dual labour supply is a very valuable aspect of being married. It adds to the economies of scale benefit of being married (identified in Section 5.2.1), accentuating the tendency towards lower saving by single households who anticipate the possibility of becoming married in the future. This is highlighted in the extreme cases of zero correlation. In this case, the fraction of the capital stock held by single households is just 22 per cent compared with 26 per cent when there is perfect correlation in partners' wages (see Models A3 and A4).

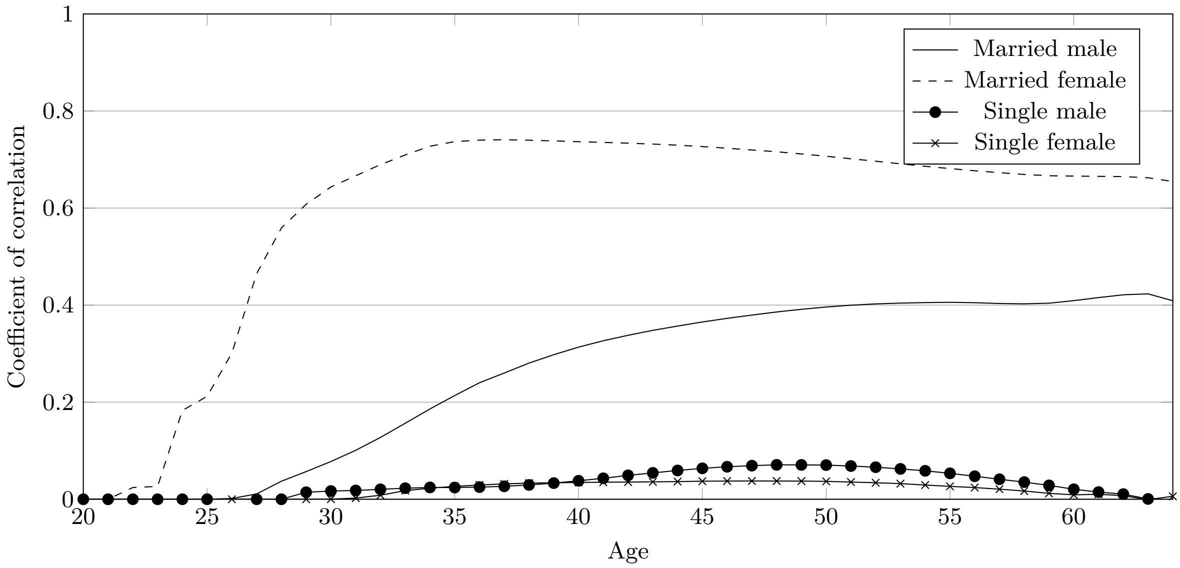

Figure 5.6.1: Correlation between hours worked and productivity realisation

Secondly, dual labour supply affects the degree to which wage inequality maps into consumption inequality. As discussed earlier, there is a growing focus in macroeconomics on the various means by which households can insure themselves against idiosyncratic wages shocks in an incomplete markets setting. This research agenda has involved attempts to both measure these partial insurance mechanisms empirically (e.g. Blundell et al. (2008, 2016)) and to investigate whether calibrated OLG models are capable of generating empirically-consistent amounts of partial insurance (e.g. Kaplan and Violante (2010); Krueger and Wu (2018)). Figure 5.6.2 shows that the increase in life cycle consumption inequality (measured as the variance of log consumption) that is endogenously-generated by the model. The overall life cycle increase lines up well with the actual data presented in Storesletten et al. (2004). This consistency provides some comfort that the model does a good job of capturing consumption-saving dynamics over the life cycle. Figure 5.6.2 also shows that the transmission of wage to consumption inequality is more muted for married households than for single households, consistent with the capacity for couples to risk share. It is also worth reiterating that the optimal tax system provides more insurance than the tax system in

[24]Note that the plotted correlation has only a very weak connection with the Frisch (wealth compensated) elasticity, in part because variations in productivity in this model are very persistent and therefore carry significant wealth effects.

the baseline economy as indicated by lower consumption inequality. This is driven by the significant tax free threshold in the optimal tax system (similar to the finding in Conesa et al. (2009)) along with the shift to separate filing which de-couples marginal tax rates among spouses.

Related to the previous two points, the dual labour supply margin means saving by married households is more elastic to the after tax return on savings. This increases the overall elasticity of saving in the benchmark economy (1.22) compared with a bachelor economy (1.09), which plays a well known role in determining the optimal tax on capital as discussed earlier in relation to Summers (1981) and Engen and Gale (1997) among others. The relationship between dual labour supply and the interest rate elasticity of saving is highlighted by comparing models where the correlation in partners' wages is zero (A3) and unity (A4). In the first case, the elasticity is 1.4, while in the second it falls to 0.54. These results are shown in Table 5.6.3. Finally, the interest rate elasticity of saving is likely to depend closely on the elasticity of inter-temporal substitution (EIS), given by the reciprocal of the preference parameter σ. The final row shows the result of reducing the EIS to 1/2 from the value of 2/3 used in the benchmark economy. Unsurprisingly, the interest elasticity falls (from 1.22 to 0.92) and the associated *optimal tax on capital becomes positive, at 9.6 per cent.* Clearly the EIS is a critical assumption in determining the optimal structure of the tax system, consistent with the findings of past optimal income tax studies including Conesa et al. (2009).

Table 5.6.3: The interest rate elasticity of saving

	Discount factor (β)	After tax return ($r(1-\tau_k)$)	Capital stock (K)	Elasticity of saving
Benchmark	0.995	3.81	1.87	1.22
Bachelor	0.983	3.18	1.68	1.09
Zero correlation (A3)	0.999	3.78	1.96	1.42
Perfect correlation (A4)	0.972	2.91	1.63	0.54
Lower EIS ($1/\sigma$=1/2)	0.991	3.65	1.75	0.92

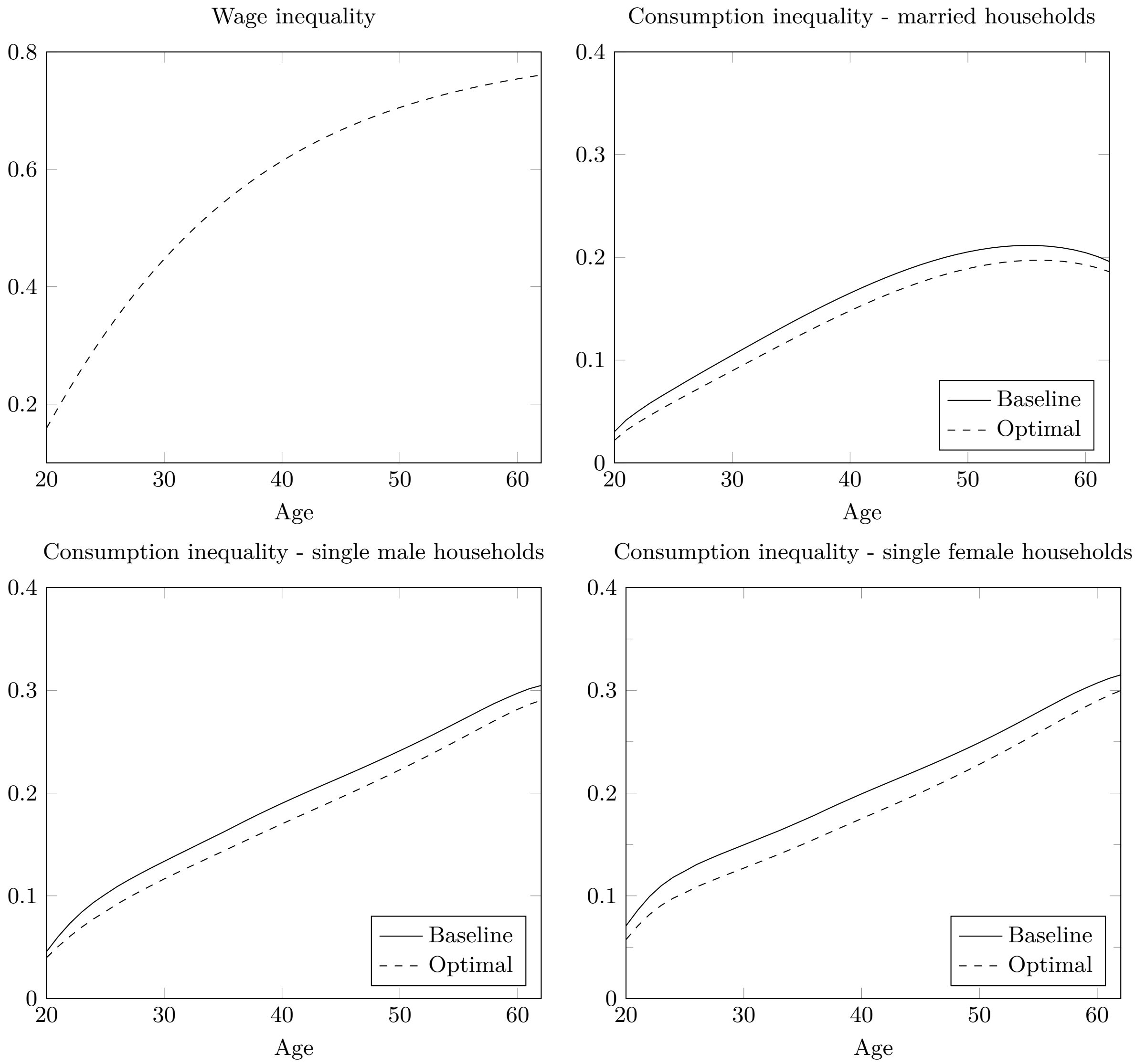

Figure 5.6.2: Inequality transmission over the life cycle

5.6.4 Relationship to Erosa and Gervais (2002) and Gervais (2012)

The study by Erosa and Gervais (2002) uses Ramsey taxation principles to establish some important optimal taxation results in a deterministic OLG model with elastic labour. In particular, Erosa and Gervais (2002) show within this framework that it is almost always optimal to tax labour and consumption at different rates over the life cycle. When labour supply falls over the life-cycle – as it does in this study's baseline economy (see Figure 5.4.1) – Ramsey principles mean it is optimal to tax it labour earnings more heavily at younger ages. When the labour tax code cannot be conditioned on age, they show that this can be achieved (imperfectly) by a positive tax on capital.

However, there are a number of reasons why the result in Erosa and Gervais (2002) does not directly apply here. First, the models considered in Erosa and Gervais (2002) (as well as in the related study by Gervais (2012)) are deterministic and therefore ignore the efficiency implications of having within-age variation in productivity and hours worked. While a tax on capital may be able to mimic the variation in hours *over the life cycle*, such a tax cannot easily mimic variation in hours *by age* across productivity states. This is not very important in a bachelor model because the substitution effect from wage variations is dominated by the wealth effect given the persistence of the productivity shock, which limits the within-age variability in hours worked. However, the picture is very different for married households whose dual labour supply margin allows a much larger substitution effect. This in turn results in a much stronger within-age correlation in hours worked and productivity for workers in married households (see Figure 5.6.1). A tax on capital is therefore no longer an effective way of taxing labour more heavily when it is high (and therefore inelastic), which is the mechanism underlying the results in Erosa and Gervais (2002) and Conesa et al. (2009).

Second, the results derived in Erosa and Gervais (2002) consider only proportional taxation. But in a closely related study, Gervais (2012) demonstrates that a progressive tax on labour can mimic optimal age dependent taxation. In this sense, Gervais (2012) argues that progressive taxation can be justified purely on efficiency grounds. It turns out that the average marginal tax rate in the benchmark economy does a good job of mimicking optimal age taxation, especially compared with the tax system in the baseline economy which taxes capital and labour income together. Figure 5.6.3 shows the mean marginal tax rates by age for each household type. As can be seen, the marginal tax rate on labour income tends to fall over the working life, consistent with the trend in hours worked which is similar to that shown in Figure 5.4.1. Notably, this progressivity is achieved entirely through having a tax free threshold.

Third, Gervais (2012) notes that a tax system can confer welfare gains by inducing flatter life cycle profiles for consumption and leisure. Thus a tax on capital (or saving) can have welfare implications simply by encouraging flatter consumption and leisure profiles. However, the magnitude of this welfare gain depends on the calibrated discount factor used to target the capital-to-output ratio. All else equal, a higher discount factor (more patient households) reduces the welfare benefit from capital taxation. As discussed above, the tendency for marital risk to reduce saving by single households means that a relatively high discount rate is needed to achieve the target capital-to-output ratio in the baseline economy. This is another factor contributing to the relatively low optimal tax on capital in the benchmark economy.

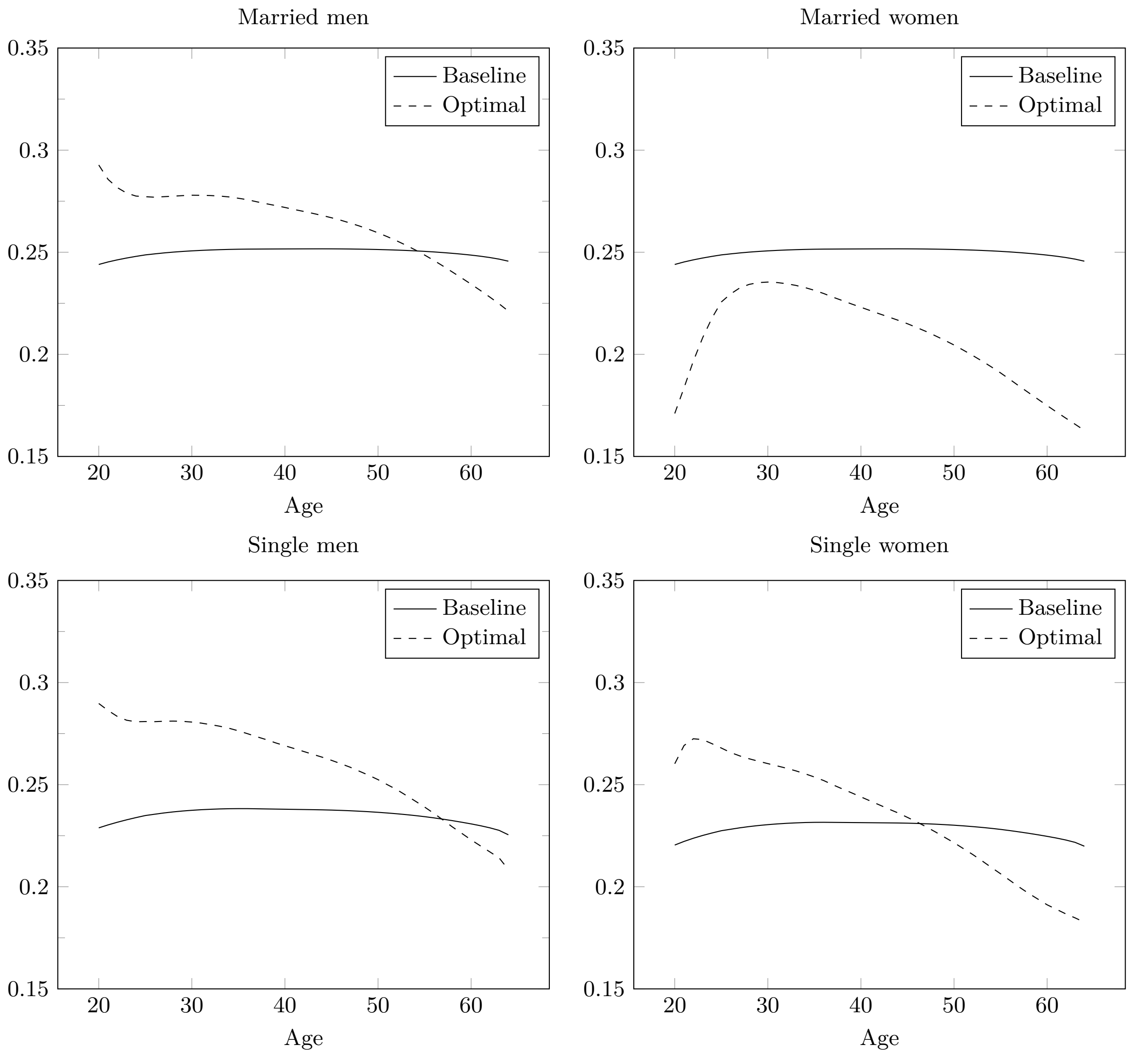

Figure 5.6.3: Marginal tax rates faced by average earner

5.7 Further work

This paper has demonstrated that the optimal long-run structure of the tax system depends on the presence of both marital risk and family labour supply, both of which are empirically important features. While my results show relatively modest welfare gains from moving to the optimal tax system of zero capital taxation – especially compared to the huge gains found in Summers (1981)[25] – my paper at the very least calls into question the robustness of recent OLG studies such as Conesa et al. (2009) based on bachelor economies.

However, further work is needed to understand the precise factors behind this paper's contrary result. While a number of main channels have been uncovered including the importance of the EIS and interest elasticity of saving, further work is needed to see whether the results are robust to alternative assumptions about other key parameters and the assumption of separable preferences. This includes the divorce regime, the bargaining weights among spouses and the stochastic properties governing marriage and divorce. In addition, a natural extension to the model presented in this paper would involve introducing further realism into the household structure by capturing the presence of children (via an age-varying economies of scale term θ), which has been identified as an important factor behind observed patterns in life-cycle consumption. Similarly, further investigation of the impact on the optimal tax structure of social security settings and alternative ways of modelling of bequests will be important given their interplay with the insurance role played by private saving and dual labour supply. While the OLG model studied in this paper already incorporates an extensive margin of labour supply adjustment by married households, adding fixed costs of work could provide a larger and more empirically-consistent degree of extensive adjustment.

Further analytical studies are also needed to understand the channels through which the dual labour supply margin affects consumption and saving, and therefore optimal taxation, over the life cycle. This could involve extending a study like Erosa and Gervais (2002) to allow for dual labour supply and wage uncertainty, or at the very least exploring these features in a simplified two period model.

[25] As discussed in Section 5.6.3, Summers (1981) was based on a highly stylised model and its main findings have largely been discredited by subsequent OLG studies.

Appendix

Appendix 5.A Computational appendix

5.A.1 Household decision rules and marginal value functions

I solve the household decision rules and value functions using an extension of Chris Carroll's method of endogenous grid points (Carroll, 2006). The decision rules for married households of working age (which is the highest dimensional case) are defined over the joint state space:

$$(a, j, z, g, ms) = [0, \infty] \times \{1..., J\} \times \{z_1, ...z_{36}\} \times \{m, f\} \times \{c, s\} \tag{5.11}$$

where the couple's two dimensional 6×6 productivity space is vectorised into a 36 point grid and the continuous asset space a is discretised into $\hat{a}_i$ using 215 grid points from $a_1 = 0$ to $a_{215} = 50$. The solution operates via backward induction from the last period of life in which the agent dies with certainty and leaves no bequest.

1. In $j = J$, the agent consumes all their income, which allows the state-contingent consumption rule and marginal value function, denoted $v_a(\hat{a}, j, z, g, ms)$, to be approximated at each grid point $\hat{a}_i$ for *beginning period assets.*

2. With period J marginal value function in hand, the consumption rule in period $J-1$ can be calculated on the basis of the assumed vector of *next* period's assets (saving) $\hat{a'_i}$ by inverting next period's marginal value function $v_a(\hat{a'}, j, z, g, ms)$ obtained in step 1, which is weighted by the probabilities of marital status transition matrix . For single households, the marginal value function is also weighted by the distribution of assets of prospective partners conditional on becoming married next period.

3. With consumption calculated, the implied cash in hand (m) can also be calculated at each grid point for $\hat{a'_i}$ using the identity $m = c + \hat{a'_i}$.

4. Using the expressions for household income shown in Section 5.3, income is calculated at each point $\hat{a}_i$.

5. Using the decision rule from 2 and the associated values of m, the consumption decision rule in terms of income derived in step 4 is generated using linear interpolation. This gives the consumption rule in terms of on *beginning period assets* $\hat{a}_i$.

6. Income from step 4 and the consumption rule from step 5 are combined to generate the saving rule, while the current period marginal value function is calculated at each $\hat{a}_i$ using the consumption decision rule and the envelope condition.

We can then iterate backwards until period j_r. Subsequent iterations need to incorporate the agent's continuous choice of hours. The steps are similar for the retired agent except that the decision rule for hours n is now found for each value of the consumption rule generated in step 2 using the first order conditions for consumption and hours. This is more complicated than models featuring proportional taxation because the non-linear tax function makes a closed form solution unavailable. Further, the expected marginal value function in step 2 must now incorporate the additional state variable associated with the joint productivity process. Next, the hours rule can be used to calculated tax payable at each point in next period asset grid $\hat{a_i'}$. The income and tax payable values for each point on this grid is used to reverse engineer beginning period assets using the budget constraint shown in Section 5.3. This generates a relation between $\hat{a_i'}$ and $\hat{a}_i$ that can be inverted using linear interpolation to generate the saving rule in terms of $\hat{a}_i$. Next, this saving rule is combined with next period's marginal value function to infer consumption and then hours rules as a function of current assets $\hat{a}_i$. The consistent values for income, tax payments and marginal tax rates can then also be derived. Finally, the envelope condition and consumption rule is used to generate the current period's marginal value function $v_a(\hat{a}, j, z, g, ms)$ and the process continues back until the first period of life. At each period, the value function itself is also evaluated and saved for later use in welfare calculations.

The task of finding the household decision rules in the baseline economy is somewhat more difficult because the progressive tax function is applied to combined capital and labour income consistent with the current US tax system. This requires a slightly different algorithm. To ensure acceptable run time, we use a binary search routine to find beginning period of assets conditional on next period assets by exploiting the montonicity in the decision rule for hours worked. This exploits the pre-computated values of the inverted progressive tax function, which considerably shortens the run time.

5.A.2 Calculation of stationary asset distribution and optimal tax mix

The following steps describe the procedure for finding the economy's stationary equilibrium in the post-reform world.

1. With the decision rules in hand, the asset distribution is calculated by iterating on the saving rules and joint law of motion described in Section 5.3. To achieve an acceptable degree of precision, the asset space is finer than the one used to calculate the decision rules. This produces the aggregate supply of

saving conditional on an assumed set of factor prices, tax rates and pension benefits.

2. Because of the presence of marriage risk, we must also iterate on agents' expectations of the distribution of prospective partners' assets, which was also taken as given in the first step, until it converges to the actual distribution, as explained in Section 5.3. In practice this can take place very efficiently by including this iteration as part of the outer loop that finds the economy's general equilibrium.

3. In the outer loop, we iterate on the factor prices until the supply of saving from 2 matches demand for capital.

4. The search for the optimal tax system is undertaken across the tax space $\{\kappa_0, \kappa_1, \tau_k\}$ which is discretised into a three dimensional grid with the value of the remaining tax parameter κ_2 implied by the balanced budget restriction. We restrict the grid points such that the optimal marginal tax rate for labour (κ_0) and capital (τ_k) lie in the interval $[0, 0.5]$. This was never found to be binding. In practice, all optimal tax systems entailed either a very small or very large value of κ_1 corresponding to purely proportional taxation of labour income or a flat tax with deduction. This simplified the search in many cases. All computation was undertaken in MATLAB, with the run time in finding the optimal tax system greatly reduced by utilising parallel processing.

Bibliography

Abowd, John and David Card, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 1989, *57* (2), 411–45.

Ahmed, Naeem, Matthew Brzozowski, and Thomas Crossley, "Measurement Errors in Recall Food Consumption Data," *IFS Working Papers, Institute for Fiscal Studies (IFS)*, November 2006.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving*," *The Quarterly Journal of Economics*, 08 1994, *109* (3), 659–684.

_ , "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 1995, *103* (6), 1158–1175.

Alan, Sule, Kadir Atalay, and Thomas F. Crossley, "Euler Equation Estimation on Micro Data," *Macroeconomic Dynamics*, 2018, p. 1–26.

_ , Orazio Attanasio, and Martin Browning, "Estimating Euler equations with noisy data: two exact GMM estimators," *Journal of Applied Econometrics*, 2009, *24* (2), 309–324.

Altonji, Joseph, "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy*, 1986, *94* (3), S176–S215.

Altonji, Joseph G. and Lewis M. Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 1996, *14* (3), 353–366.

Andrews, Donald W. K., "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, 1993, *61* (1), 139–165.

Atkinson, Anthony and Joseph Stiglitz, "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 1976, *6* (1-2), 55–75.

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos, "Female Labor Supply As Insurance Against Idiosyncratic Risk," *Journal of the European Economic Association*, 04/05 2005, *3* (2-3), 755–764.

Attanasio, Orazio P. and Hamish Low, "Estimating Euler Equations," *Review of Economic Dynamics*, April 2004, *7* (2), 405–435.

_ , James Banks, Costas Meghir, and Guglielmo Weber, "Humps and Bumps in Lifetime Consumption," *Journal of Business and Economic Statistics*, 1999, *17* (1), 22–35.

Baker, Michael, "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, April 1997, *15* (2), 338–75.

_ and Gary Solon, "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics*, April 2003, *21* (2), 267–288.

Barrett, Garry F, "The Return to Cognitive Skills in the Australian Labour Market*," *Economic Record*, 2012, *88* (280), 1–17.

Barrett, Garry F., Thomas F. Crossley, and Christopher Worswick, "Consumption and Income Inequality in Australia," *Economic Record*, 2000, *76* (233), 116–138.

Bianchi, Marco, Björn R. Gudmundsson, and Gylfi Zoega, "Iceland's Natural Experiment in Supply-Side Economics," *The American Economic Review*, 2001, *91* (5), 1564–1579.

Blundell, Richard, "Income Dynamics and Life-cycle Inequality: Mechanisms and Controversies," *The Economic Journal*, 2014, *124* (576), 289–318.

_ **and Ben Etheridge**, "Consumption, income and earnings inequality in Britain," *Review of Economic Dynamics*, 2010, *13* (1), 76 – 102. Special issue: Cross-Sectional Facts for Macroeconomists.

_ **, Antoine Bozio, and Guy Laroque**, "Extensive and Intensive Margins of Labour Supply: Work and Working Hours in the US, the UK and France*," *Fiscal Studies*, 2013, *34* (1), 1–29.

_ **, Luigi Pistaferri, and Ian Preston**, "Consumption Inequality and Partial Insurance," *American Economic Review*, December 2008, *98* (5), 1887–1921.

_ **,** _ **, and Itay Saporta-Eksten**, "Consumption Inequality and Family Labor Supply," *American Economic Review*, February 2016, *106* (2), 387–435.

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde, "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes," *Labour Economics*, 2007, *14* (6), 926 – 937. Education and Risk.

Borland, Jeff and Michael Coelli, "Labour Market Inequality in Australia," *Economic Record*, 2016, *92* (299), 517–547.

Bound, John, Charles Brown, and Nancy Mathiowetz, "Chapter 59 - Measurement Error in Survey Data," in James J. Heckman and Edward Leamer, eds., *James J. Heckman and Edward Leamer, eds.*, Vol. 5 of *Handbook of Econometrics*, Elsevier, 2001, pp. 3705 – 3843.

Breunig, Robert and Joseph Mercante, "The Accuracy of Predicted Wages of the Non-Employed and Implications for Policy Simulations from Structural Labour Supply Models*," *Economic Record*, 2010, *86* (272), 49–70.

_ **, Deborah A. Cobb-Clark, and Xiaodong Gong**, "Improving the Modelling of Couples' Labour Supply*," *Economic Record*, 2008, *84* (267), 466–485.

Browning, Martin, Mette Ejrnjes, and Javier Alvarez, "Modelling Income Processes with Lots of Heterogeneity," *The Review of Economic Studies*, 2010, *77* (4), 1353–1381.

Brzozowski, Matthew, Thomas F. Crossley, and Joachim K. Winter, "A comparison of recall and diary food expenditure data," *Food Policy*, 2017, *72*, 53 – 61. Special Issue: Food counts.

Cagetti, Marco, "Interest Elasticity in a Life-Cycle Model with Precautionary Savings," *The American Economic Review*, 2001, *91* (2), 418–421.

Cardak, Buly and Roger Wilkins, "The determinants of household risky asset holdings: Australian evidence on background risk and other factors," *Journal of Banking and Finance*, 2009, *33* (5), 850–860.

Caroli, Eve and Cecilia Garcia-Penalosa, "Risk aversion and rising wage inequality," *Economics Letters*, 2002, *77* (1), 21 – 26.

Carroll, Christopher, "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)," *The B.E. Journal of Macroeconomics*, 2001, *1* (1), 1–38.

_ , "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 2006, *91* (3), 312–320.

Carroll, Christopher D., "A Theory of the Consumption Function, with and without Liquidity Constraints," *The Journal of Economic Perspectives*, 2001, *15* (3), 23–45.

_ **and Andrew A. Samwick**, "How Important is Precautionary Saving?," *The Review of Economics and Statistics*, 1998, *80* (3), 410–419.

Chamberlain, Gary, "Panel data," in Z. Griliches† and M. D. Intriligator, eds., *Handbook of Econometrics*, 1 ed., Vol. 2, Elsevier, 1984, chapter 22, pp. 1247–1318.

Chamley, Christophe, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, May 1986, *54* (3), 607–622.

Chapman, Bruce and Andrew Leigh, "Do Very High Tax Rates Induce Bunching? Implications for the Design of Income Contingent Loan Schemes*," *Economic Record*, 2009, *85* (270), 276–289.

Chatterjee, Arpita, Aarti Singh, and Tahlee Stone, "Understanding Wage Inequality in Australia," *Economic Record*, 2016, *92* (298), 348–360.

Chetty, Raj, "A New Method of Estimating Risk Aversion," *The American Economic Review*, 2006, *96* (5), 1821–1834.

_ , Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, May 2011, *101* (3), 471–75.

Chiappori, Pierre-Andre and Monica Paiella, "Relative Risk Aversion Is Constant: Evidence From Panel Data," *Journal of the European Economic Association*, December 2011, *9* (6), 1021–1052.

Chioda, Laura, "Estimating Euler Equations with Measurement Error: A Nonparametric Approach," Available: https://pdfs.semanticscholar.org/f977/9ee90802d4b60fe76a466188602f6a737d73.pdf [15 November 2018], 2004.

Coelli, Michael and Roger Wilkins, "Credential Changes and Education Earnings Premia in Australia," *Economic Record*, 2009, *85* (270), 239–259.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, March 2009, *99* (1), 25–48.

Conlin, Andrew, Jouko Miettunen, Jukka Perttunen, Mikko Puhakka, and Rauli Svento, "Risk Aversion, Wealth and Personality," 2016. Unpublished working paper. Available at `https://www.univaasa.fi/en/research/groups/frg/researchseminars/financeseminars/file/`.

Cubeddu, Luis and José-V´ctor R´os-Rull, "Families as shocks," *Journal of the European Economic Association*, 2003, *1* (2-3), 671–682.

_ and Jose-Victor Rios-Rull, "Marital risk and capital accumulation," Technical Report 1997.

Dabla-Norris, Era, Kalpana Kochhar, Nujin Suphaphiphat, Frantisek Ricka, and Evridiki Tsounta, "Causes and Consequences of Income Inequality; A Global Perspective," IMF Staff Discussion Notes 15/13, International Monetary Fund June 2015.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59* (5), 1221–1248.

Devereux, Michael B. and Gregor W. Smith, "International Risk Sharing and Economic Growth," *International Economic Review*, 1994, *35* (3), 535–550.

Dmytro, Hryshko, "Labor income profiles are not heterogeneous: Evidence from income growth rates," *Quantitative Economics*, 2012, *3* (2), 177–209.

Doepke, M. and M. Tertilt, *Families in Macroeconomics*, Vol. 2 of *Handbook of Macroeconomics*, Elsevier, July

Dollman, Rosetta, Greg Kaplan, Gianni La Cava, and Tahlee Stone, "Household Economic Inequality in Australia," *Reserve Bank of Australia Research Discussion Papers*, 2015, *2015* (15).

Domeij, David and Martin Flodén, "The labor-supply elasticity and borrowing constraints: Why estimates are biased," *Review of Economic Dynamics*, 2006, *9* (2), 242 – 262.

Doris, Aedin, Donal O'Neill, and Olive Sweetman, "Identification of the Covariance Structure of Earnings Using the GMM Estimator," IZA Discussion Papers 4952, Institute for the Study of Labor (IZA) May 2010.

Eisenhauer, Joseph G. and Luigi Ventura, "Survey measures of risk aversion and prudence," *Applied Economics*, 2003, *35* (13), 1477–1484.

Engen, Eric M. and William G. Gale, "Consumption Taxes and Saving: The Role of Uncertainty in Tax Reform," *The American Economic Review*, 1997, *87* (2), 114–119.

Erosa, Andrés and Martin Gervais, "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 2002, *105* (2), 338 – 369.

_ , Luisa Fuster, and Gueorgui Kambourov, "Towards a Micro-Founded Theory of Aggregate Labour Supply," *Review of Economic Studies*, 2016, *83* (3), 1001–1039.

Evans, Owen J., "Tax Policy, the Interest Elasticity of Saving, and Capital Accumulation: Numerical Analysis of Theoretical Models," *The American Economic Review*, 1983, *73* (3), 398–410.

Fehr, Hans and Fabian Kindermann, "Taxing capital along the transition—Not a bad idea after all?," *Journal of Economic Dynamics and Control*, 2015, *51* (C), 64–77.

Finke, Michael S. and Nancy L. Pierce, "Precautionary Savings Behavior of Maritally Stressed Couples," *Family and Consumer Sciences Research Journal*, 2006, *34* (3), 223–240.

Fiorito, Riccardo and Giulio Zanella, "The Anatomy of the Aggregate Labor Supply Elasticity," *Review of Economic Dynamics*, 2012, *15* (2), 171–187.

Flodén, Martin, "Labour Supply and Saving Under Uncertainty*," *The Economic Journal*, 2006, *116* (513), 721–737.

Gandelman, Nestor and Ruben Hernandez-Murillo, "Risk Aversion at the Country Level," *Review*, 2015, *97* (1), 53–66.

Gervais, Martin, "On the optimality of age-dependent taxes and the progressive U.S. tax system," *Journal of Economic Dynamics and Control*, 2012, *36* (4), 682 – 691.

González, Libertad and Berkay Özcan, "The Risk of Divorce and Household Saving Behavior," *Journal of Human Resources*, 2013, *48* (2), 404–434.

Gourinchas, Pierre-Olivier and Jonathan A. Parker, "Consumption Over the Life Cycle," *Econometrica*, 2002, *70* (1).

Gouveia, Miguel and Robert P. Strauss, "Effective Federal Individual Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, 1994, *47* (2), 317–39.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura, "Taxation and Household Labour Supply," *Review of Economic Studies*, 2012, *79* (3), 1113–1149.

_ , _ , and _ , "Taxing women: A macroeconomic analysis," *Journal of Monetary Economics*, 2012, *59* (1), 111 – 128. Carnegie-NYU-Rochester Conference Series on Public Policy at New York University on April 15-16, 2011.

Gustavsson, Magnus and Pär Österholm, "Does the labor-income process contain a unit root? Evidence from individual-specific time series," *Journal of Economic Dynamics and Control*, 2014, *47*, 152 – 167.

Guvenen, Fatih, "Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective," *Journal of Monetary Economics*, 2006, *53* (7), 1451 – 1472.

_ , "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, January 2009, *12* (1), 58–79.

_ , "Macroeconomics With Heterogeneity: A Practical Guide," NBER Working Papers 17622, National Bureau of Economic Research, Inc November 2011.

Hahn, Youjin, Stephen Matteo Miller, and Hee-Seung Yang, "Inequality, Risk-Sharing and the Crisis: A View From Australia," Monash Economics Working Papers 15-16, Monash University, Department of Economics 2016.

Hall, Robert E., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 1978, *86* (6), 971–987.

Hanel, Barbara and John Haisken-DeNew, "Keep Calm and Consume? Subjective Uncertainty and Precautionary Savings," July 2017. Melbourne Institute Working Paper No. 18/17.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante, "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, January 2010, *13* (1), 15–51.

_ , Kjetil Storesletten, and Giovanni L. Violante, "Two Views of Inequality Over the Life Cycle," *Journal of the European Economic Association*, 04/05 2005, *3* (2-3), 765–775.

_ , _ , and _ , "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 2009, *1* (1), 319–354.

Heckman, J. J., "The incidental parameter problem and the problem of initial conditions in estimating a discrete time-discrete data," *Structural Analysis of Discrete Data with Econometric Applications*, 1981, pp. 179–195.

Heckman, James J. and Thomas E. Macurdy, "A Life Cycle Model of Female Labour Supply," *The Review of Economic Studies*, 1980, *47* (1), 47–74.

Hoffmann, Florian, "HIP, RIP and the Robustness of Empirical Earnings Processes," *Unpublished manuscript*, 2013.

Hryshko, Dmytro, "Labor income profiles are not heterogeneous: Evidence from income growth rates," *Quantitative Economics*, 2012, *3* (2), 177–209.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, December 2011, *101* (7), 2923–54.

Hyslop, Dean R., "Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings," *The American Economic Review*, 2001, *91* (4), 755–777.

Imai, Susumu and Michael P. Keane, "Intertemporal Labor Supply and Human Capital Accumulation*," *International Economic Review*, 2004, *45* (2), 601–641.

Imrohoroglu, Selahattin, "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, May 1998, *39* (2), 307–328.

Iskhakov, Fedor and Michael Keane, "Effects of Taxes and Safety Net Pensions on life-cycle Labor Supply, Savings and Human Capital: the Case of Australia," ANU Working Papers in Economics and Econometrics 2018-661, Australian National University, College of Business and Economics, School of Economics May 2018.

_ and _ , "Effects of Taxes and Safety Net Pensions on life-cycle Labor Supply, Savings and Human Capital: the Case of Australia," Discussion Papers 2018-09, School of Economics, The University of New South Wales May 2018.

Judd, Kenneth, "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics*, 1985, *28* (1), 59–83.

Kalb, Guyonne, "Estimation of Labour Supply Models for Four Separate Groups in the Australian Population," Melbourne Institute Working Paper Series wp2002n24, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne November 2002.

Kaplan, Greg and Giovanni L Violante, "How much consumption insurance beyond self-insurance?," *American Economic Journal: Macroeconomics*, 2010.

Kaygusuz, Remzi, "Taxes and Female Labor Supply," *Review of Economic Dynamics*, October 2010, *13* (4), 725–741.

Keane, Michael and Richard Rogerson, "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom," *Journal of Economic Literature*, June 2012, *50* (2), 464–76.

Keane, Michael P., "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 2011, *49* (4), 961–1075.

— , "Effects of permanent and transitory tax changes in a life-cycle labor supply model with human capital," *International Economic Review*, 2015, *56* (2), 485–503.

— **and Kenneth I. Wolpin**, "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 2001, *42* (4), 1051–1103.

Kimball, Miles S., "Precautionary Saving in the Small and in the Large," *Econometrica*, 1990, *58* (1), 53–73.

Kimball, Miles S, Claudia R Sahm, and Matthew D Shapiro, "Imputing Risk Tolerance From Survey Responses," *Journal of the American Statistical Association*, 2008, *103* (483), 1028–1038. PMID: 20407599.

Kimmel, Jean and Thomas J. Kniesner, "New evidence on labor supply:: Employment versus hours elasticities by sex and marital status," *Journal of Monetary Economics*, 1998, *42* (2), 289 – 301.

King, Robert G. and Sergio T. Rebelo, "Chapter 14 Resuscitating real business cycles," in "in," Vol. 1 of *Handbook of Macroeconomics*, Elsevier, 1999, pp. 927 – 1007.

Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 1996, *34* (1), 42–71.

Krueger, Dirk and Chunzan Wu, "How Much Consumption Insurance in Bewley Models with Endogenous Family Labor Supply?," CEPR Discussion Papers 12828, C.E.P.R. Discussion Papers March 2018.

— **, Fabrizio Perri, Luigi Pistaferri, and Giovanni L. Violante**, "Cross-sectional facts for macroeconomists," *Review of Economic Dynamics*, 2010, *13* (1), 1 – 14. Special issue: Cross-Sectional Facts for Macroeconomists.

Kureishi, Wataru and Midori Wakabayashi, "What motivates single women to save? the case of Japan," *Review of Economics of the Household*, December 2013, *11* (4), 681–704.

Levy, Haim, "Absolute and relative risk aversion: An experimental study," *Journal of Risk and Uncertainty*, May 1994, *8* (3), 289–307.

Lillard, Lee and Yoram Weiss, "Components of Variation in Panel Earnings Data: American Scientists, 1960-70," *Econometrica*, 1979, *47* (2), 437–54.

Low, Hamish W., "Self-insurance in a life-cycle model of labour supply and savings," *Review of Economic Dynamics*, 2005, *8* (4), 945 – 975.

Lupton, J. and J.P. Smith, "Marriage, Assets, and Savings," Technical Report 1999.

Lusardi, Annamaria, "On the Importance of the Precautionary Saving Motive," *The American Economic Review*, 1998, *88* (2), 449–453.

MaCurdy, Thomas E., "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 1981, *89* (6), 1059–1085.

— , "The use of time series processes to model the error structure of earnings in a longitudinal data analysis," *Journal of Econometrics*, January 1982, *18* (1), 83–114.

Manovskii, Iourii, Dmytro Hryshko, and Moira Daly, "Reconciling Estimates of Earnings Processes in Growth Rates and Levels," Technical Report 2015.

Meghir, Costas and Luigi Pistaferri, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 2004, *72* (1), 1–32.

— **and** — , *Earnings, Consumption and Life Cycle Choices*, Vol. 4 of *Handbook of Labor Economics*, Elsevier,
ΩİImrohoroğlu

İmrohoroğlu, Selahattin, "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, 1998, *39* (2), 307–328.

Nishiyama, Shinichi, "The joint labor supply decision of married couples and the U.S. Social Security pension system," *Review of Economic Dynamics*, 2019, *31*, 277 – 304.

OECD, "OECD/IAB Employment Protection Database, 2013 update," *www.oecd.org/employment/protection*, 2013.

Ogaki, Masao and Qiang Zhang, "Decreasing Relative Risk Aversion and Tests of Risk Sharing," *Econometrica*, 2001, *69* (2), 515–526.

Pericoli, Filippo and Luigi Ventura, "Family dissolution and precautionary savings: an empirical analysis," *Review of Economics of the Household*, Dec 2012, *10* (4), 573–595.

Peterman, William, "Determining the motives for a positive optimal tax on capital," *Journal of Economic Dynamics and Control*, 2013, *37* (1), 265–295.

Peterman, William B., "Reconciling micro and macro estimates of the Frisch labor supply elasticity," *Economic Inquiry*, 2016, *54* (1), 100–120.

Pijoan-Mas, Josep, "Precautionary savings or working longer hours?," *Review of Economic Dynamics*, 2006, *9* (2), 326 – 352.

Rogerson, Richard and Johanna Wallenius, "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic Theory*, 2009, *144* (6), 2277 – 2292. Dynamic General Equilibrium.

Schwab, Klaus and Xavier Sala i Martin, "Global Competitiveness Report 2017-2018," World Economic Forum 09/2017, World Economic Forum 2017.

Shore, Stephen H., "The co-movement of couples' incomes," *Review of Economics of the Household*, Sep 2015, *13* (3), 569–588.

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron, "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, April 2004, *51* (3), 609–633.

Summers, Lawrence H., "Capital Taxation and Accumulation in a Life Cycle Growth Model," *The American Economic Review*, 1981, *71* (4), 533–544.

Tran, Chung and Alan Woodland, "Trade-Offs in Means Tested Pension Design," Working Papers 201110, ARC Centre of Excellence in Population Ageing Research (CEPAR), Australian School of Business, University of New South Wales 2011.

Tsigos, Stamatios and Kevin Daly, ""Fair go" for all? Wealth and Risk Aversion of Australian Households*," *Australian Economic Papers*, 2016, *55* (3), 274–300.

Ventura, Eva, "A note on measurement error and Euler equations," *Economics Letters*, 1994, *45* (3), 305 – 308.

Verbeek, Marno and Francis Vella, "Estimating dynamic models from repeated cross-sections," *Journal of Econometrics*, 2005, *127* (1), 83 – 102.

Voena, Alessandra, "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?," *American Economic Review*, August 2015, *105* (8), 2295–2332.

Whiteford, Peter, "Inequality and Its Socioeconomic Impacts," *Australian Economic Review*, 2015, *48* (1), 83–92.

Wilkins, R., *The Household Income and Labour Dynamics in Australia Survey : selected findings from Waves 1 to 14* 2016.

_ **and C. Sun**, "Assessing the quality of the expenditure data collected in the self-completion questionnaire," HILDA Project Discussion Paper Series no. 1/10, Melbourne Institute of Applied Economic and Social Research 2010.

Wilkins, Roger, "Measuring Income Inequality in Australia," *Australian Economic Review*, 2015, *48* (1), 93–102.

Yagihashi, Takeshi and Juan Du, "Intertemporal elasticity of substitution and risk aversion: are they related empirically?," *Applied Economics*, March 2015, *47* (15), 1588–1605.

Zagorsky, Jay L., "Marriage and divorce's impact on wealth," *Journal of Sociology*, 2005, *41* (4), 406–424.

Disclaimer

The Household, Income and Labour Dynamics in Australia (HILDA) Survey was initiated and is funded by the Australian Government Department of Social Services (DSS), and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views based on these data should not be attributed to either DSS or the Melbourne Institute.