University of New South Wales
School of Economics

Honours Thesis

Life-Cycle Inequality and Wage Dynamics in Australia

Author: Tahlee Stone
Student ID: 3309014

Supervisor: Arpita Chatterjee

Bachelor of Economics (Financial Economics/Economics) (Honours)

21st November, 2013
Declaration

I declare that this thesis is my own work and that, to the best of my knowledge, it contains no material that has been published or written by another person(s) except where due acknowledgement has been made. This thesis has not been submitted for award of any other degree or diploma at the University of New South Wales or at any other educational institution. I declare that the intellectual content of this thesis is the product of my own work except to the extent that assistance from others is acknowledged.

.................................
Tahlee Stone
15 November 2013
Disclaimer

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA), and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper; however, are those of the author and should not be attributed to either FaHCSIA or the Melbourne Institute.
Acknowledgements

Most importantly, I would like to thank my supervisor, Dr. Arpita Chatterjee for your time, encouragement, support and especially your gentle spirit. I have learnt so much from you this year and would not have been able to produce this thesis without your guidance.

Next, I must thank Professor Denzil Fiebig for your patience with my never-ending questions, Dr. Pauline Grosjean, for your kindness and Dr. Chris Bidner for all the back-up support you have provided during the honours program. Thanks must also be given to Assoc. Professor Valentyn Panchenko, for your assistance with administrative issues and to other members of the University of New South Wales Economics Faculty – I have been in constant awe of you all.

I would like to also thank the Economics Honours Class of 2013. Never have I been surrounded by such a ‘heterogeneous’ group of likeminded people. I owe particular thanks to Jason Qu and Yasaman Naghiloo for not only assisting me in editing this thesis but also for all the amazing help you both have provided throughout the entire year. I must additionally thank Shahin Elliin, who by all means has been an honorary, honours student.

To Christopher and Madeline, I thank you particularly for proof-reading this thesis but I also thank you generally for being particularly awesome siblings.

Finally, to my incredible parents, Greg and Jenny – I send my greatest thanks. I am extremely lucky for the love and opportunities you have always provided me. Thank you for making this all possible.
# Contents

The document contains the following sections:

- **Table of Contents**
- **Abstract**
- **1 Introduction**
- **2 Literature Review**
  - 2.1 Incomplete Markets Literature
  - 2.2 Empirical Life-Cycle Inequality Profiles
  - 2.3 Estimating the Stochastic Wage Process
  - 2.4 Australian Inequality Literature
- **3 Data and Measurement**
  - 3.1 Data description
  - 3.2 Sampling Strategy
  - 3.3 Key Variables of Interest
    - 3.3.1 Income Measures
    - 3.3.2 Education
    - 3.3.3 Control Variables
  - 3.4 Descriptive Statistics
    - 3.4.1 Descriptive Statistics by Education Group
- **4 Inequality over Time**
  - 4.1 Cross-Sectional Wage Inequality Trends
  - 4.2 Dispersion of Hours, Wages and Earnings
  - 4.3 Between and Within-Group Inequality
- **5 Dynamics of Inequality**
  - 5.1 Statistical Model
  - 5.2 Results
- **6 Life-Cycle Inequality**
  - 6.1 Model Specification and Identification
  - 6.2 Results
- **7 Simulated Life-Cycle Profiles**
7.1 Implications of Permanent and Transitory Parameters over the Life-Cycle. ................................................................. 39

8 The Role of Education ................................................................. 42
  8.1 Wage Inequality Trends by Education Group .......................... 42
  8.2 Within-Group Inequality by Education Group ....................... 44

9 Summary and Conclusion .......................................................... 46

A Data ............................................................................................ 48
  A.1 Sampling Strategy: Male Full-Time Sample ......................... 48
  A.2 Education Groups: Trends in Average Weekly Earnings ....... 48
  A.3 Description of Key Demographic Variables ...................... 49
  A.4 Descriptive Statistics for Education Subsamples .................. 49

B Life-Cycle Inequality .................................................................... 50
  B.1 Life-Cycle Profiles by Education Group .............................. 50

C Simulated Life-Cycle Inequality Profiles .................................... 51
  C.1 Sampling Strategy: Simulated Life-Cycle Inequality Profiles ................................................................. 51
  C.2 Summary: Steps for Simulating Life-Cycle Profiles ............. 51
  C.3 Simulated vs. Actual Life-Cycle Inequality Profiles by Education Group ................................................................. 53
List of Figures

3.1 Average Labour and Pre-Tax Income: 2000-01 to 2010-11 .......................... 17
3.2 Average Wages and Hours: 2000-01 to 2010-11 ................................. 18
3.3 Average Wages and Hours by Education: 2000-01 to 2010-11 .................. 19

4.1 Wage Inequality Trends: 2000-01 to 2010-11 ........................................ 20
4.2 Average Earnings, Wage and Hours Growth: 2000-01 to 2010-11 ............. 22
4.3 Between-Group and Residual Wage Inequality: 2000-01 to 2010-11 .............. 24

5.1 Estimates of Variances of Permanent and Transitory Components ......... 30

6.1 Life-Cycle Wage Inequality: Time & Cohort Views ............................... 35

7.1 Actual vs. Simulated Age Profile: Variance of Log Hourly Wages .......... 39
7.2 Actual vs. Simulated Age Profile: Gini Coefficient ............................... 40

8.1 Wage Inequality Trends by Education: 2000-01 to 2010-11 ....................... 42
8.2 Estimates of Variances of Permanent and Transitory Components by Education .................................................. 44

B.1 Life-Cycle Wage Inequality by Education: Time & Cohort Views .............. 50

C.1 Actual vs. Simulated Life-Cycle Inequality Profiles by Education ............ 53
## List of Tables

3.1 Descriptive Statistics for Full-time Male Sample .......................... 18

5.1 Estimates of Wage Dynamics: Cross-Country Comparison .................. 31

7.1 MSE estimates: Simulated vs Actual Variance Profiles .................... 40

7.2 MSE estimates: Simulated vs. Actual Gini Profiles ........................ 41

8.1 Estimates of Wage Dynamics by Education Group ............................ 45

A.1 Sampling Strategy: Full-time Male Sample .................................. 48

A.2 Description of Key Demographic Variables .................................. 49

A.3 Descriptive Statistics by Education Group .................................. 49

C.1 Sampling Strategy: Simulated Life-Cycle Profiles ............................ 51
Abstract

Quantifying welfare implications of rising income inequality is an important research agenda. This thesis documents empirical evidence to inform future quantitative macroeconomic analysis using longitudinal income data from Australia. I use the Household, Income and Labour Dynamics in Australia (HILDA) survey to characterise wage inequality patterns across both time and life-cycle dimensions. A substantial rise in wage inequality is observed between 2001 and 2011 and I find robust evidence for an increasing inequality age-profile in the data. I also employ an unobservable components model and the panel structure of HILDA to analyse wage dynamics. This model decomposes residual inequality into its permanent and transitory elements. I employ two estimation procedures – a ‘level’ and ‘difference’ approach – for this task and the two approaches deliver substantially different parameters. The disparity in my findings supports the view in the literature that this conventional model may be misspecified. I further investigate this empirical puzzle by simulating life-cycle inequality profiles using my two sets of permanent-transitory moments and comparing them to their actual counterparts. My results indicate a better fit for the first-difference approach in Australian data. In addition to these main results, I also find evidence for a stable education wage premium and provide education-specific estimates of permanent and transitory parameters for Australia.


1 Introduction

Income inequality has risen in Australia over the last three decades. This trend has also been observed for other advanced economies including the United States, Canada and Great Britain. For Australia, inequality growth has been concentrated to the past decade. In the 2000s, Australia’s top income decile experienced the largest increase in income of any other country in the OECD and the level of income inequality in Australia rose above the OECD average for the first time in 2005. These recent changes have raised two important questions: How does rising inequality impact the decisions and life-time utility of individuals? And what does this imply for the effective targeting of government-funded social insurance programs in Australia?

These questions have already been addressed in international literature, mainly in the United States, within the framework of incomplete-markets heterogeneous-agent life-cycle models. These general equilibrium models are used for quantitative macroeconomic analysis into how inequality impacts upon individual resource-allocation decisions and what affect this has on the entire distribution of a variety of economic outcomes. This framework has not yet been employed within the Australian literature; existing empirical studies instead focus largely on addressing the sources of rising inequality without considering whether these changes have important welfare implications. In this thesis, I use longitudinal Australian data to provide the empirical evidence necessary for informing the future implementation of such models.

This paper makes two contributions to existing studies. First, I document the wage inequality trends required to calibrate these quantitative macro models. I characterise wage inequality by dissecting it across both time and life-cycle dimensions and reporting both its magnitude and evolution. I examine cross-sectional trends of wage inequality across the period 2001 to 2011, using conventional indices of inequality. I confirm a substantial rise in inequality over this period, particularly for the top half of the wage distribution. I then analyse both between-group and within-group inequality and find that increased inequality is

---

2 OECD (2011) ‘Divided We Stand Report’.
predominately driven by unobservable characteristics or within-group inequality.

Second, I estimate a stochastic wage process by employing an unobservable components model to further investigate the role of residual (within-group) inequality. I model the dynamics of wages in a similar fashion to that of existing studies including Gottschalk and Moffit (1994), Heathcote et al. (2010) and Huggett, Ventura, and Yaron (2011). I use the panel structure of the Household and Income Dynamics in Australia (HILDA) survey to differentiate permanent (long-run) from transitory (short-term) residual inequality.

This variance decomposition is important as it provides key inputs to these general equilibrium life-cycle models. Both uninsurable permanent and insurable transitory wage shocks are fed through the model and agents respond by adjusting their hours, earnings and consumption. These responses provide the key predictions of these models for welfare and enable the analysis of a wide range of policy issues.

Following Heathcote et al. (2010), I employ two conventional approaches from within the literature to estimate this model. I first estimate moments based on log wage growth rates. This represents the ‘difference’ approach which has been well established by the labour economics literature (e.g. Abowd and Card (1989)). Second, I adopt the ‘level’, a method favoured by macroeconomists (e.g. Storesletten, Telmer, and Yaron (2004b)), which estimates moments based on log wage levels. Both methods are standard approaches and theoretically should provide similar estimates of the moments of interest.

However, like Heathcote et al. (2010), I find that the two methods deliver significantly different estimates of permanent and transitory inequality. This finding is actually an empirical regularity which has been found to hold in other countries including Canada, Sweden and Germany. Therefore, the results presented in this paper provide Australian evidence supporting the argument within the literature that this conventional model may be misspecified. This empirical puzzle represents an ongoing area of research. Economists are actively trying to resolve this issue to ensure the sustained feasibility of the incomplete-markets framework.

Given the life-cycle structure of the majority of these structural models, it is also necessary to evaluate how inequality behaves over the life-cycle. I achieve this by adopting an empirical approach first developed by Deaton and Paxson (1994). I construct age-profiles of the statistics of interest and find evidence for increasing inequality over the life-cycle which is robust to different measures of inequality and

---

4See Brzozowski et al. (2010) for Canada; Flodén and Domeij (2010) for Sweden; and Fuchs-Schündeln, Krueger, and Sommer (2010) for Germany.
alternative identification strategies commonly used by the literature. I then evaluate my two sets of level and difference estimates using these life-cycle inequality profiles. I simulate inequality age-profiles using the diverging set of moments taken from both approaches and then compare these simulated profiles to the actual profiles estimated from data. This task involves simulating the stochastic wage process for each individual over the sample period using an unbalanced dynamic panel. This analysis presents results useful for informing future research into reconciling the two approaches. For Australian data, I find evidence supporting the first-difference approach.

Finally, additional to my main results, I also ask the question: Are higher levels of education associated with higher levels of both between and within-group inequality? I find a stable education earnings premium. I also estimate education-specific permanent and transitory parameters. I find that these results of both between and within-education group inequality align with the findings of both Australian and international empirical studies.

The rest of the paper is structured as follows. Section 2 discusses the related literature and Section 3 provides brief overview of the data. Section 4 documents cross-sectional inequality trends between 2001 and 2011. Section 5 uses the panel dimension of HILDA to decompose permanent and transitory components of wage dynamics. In Section 6, I then analyse the behavior of inequality over the life-cycle. Section 7 evaluates the diverging sets of level and difference parameters estimated in Section 5 by simulating life-cycle inequality profiles and comparing these to their actual counterparts estimated by the procedure developed in Section 6. Finally, Section 8 assesses between and within education group inequality. Section 9 concludes.

5The two identification strategies used in this analysis place separate restrictions on cohort and time effects in order to capture age effects. Heathcote, Storesletten, and Violante (2005) discuss both approaches to estimating life-cycle inequality in greater detail.
2 Literature Review

In this chapter I provide an overview of the increasingly important role of the incomplete-markets literature. I then discuss the relevant literature that pertains to the empirical strategies used to calibrate/estimate incomplete-markets models. Finally, I survey the related areas of research in the Australian inequality literature.

2.1 Incomplete Markets Literature

In the last decade there has been a departure in international literature, away from identifying drivers of dramatic changes in the wage distribution, towards the examination of welfare implications related to rising inequality. This new area of research within the quantitative macroeconomic literature has been facilitated by the development of standard incomplete-markets heterogeneous-agent models. These models have been able to capture the joint distribution of economic outcomes across individuals as opposed to an aggregate outcome implied by one agent. Compared to the earlier representative-agent framework, heterogeneous-agent models therefore allow for a more realistic analysis of policy issues.

Early studies in this area (see Imrohoroglu (1989); Huggett (1993); Aiyagari (1994) and Rios-Rull (1995)) established the flexible theoretical foundation of this structural approach, while recent literature has extended the standard incomplete-markets framework, to mainly investigate the following questions: Firstly, what are the different sources and policy implications of inequality and heterogeneity? Secondly, to what extent do insurance channels mitigate the transmission of income to consumption inequality?

The first question has been addressed in several life-cycle versions of this model (Keane and Wolpin (1997), Storesletten, Telmer, and Yaron (2004a) and Huggett et al. (2011), for example). These studies have explored whether life-cycle inequality is driven by differences in the initial conditions of individuals such as ability, financial wealth and human capital or by differences in exogenous shocks to income received over the working lifetime (e.g. luck or cyclical fluctuations). Among this research, Keane and Wolpin (1997) use U.S. National Longitudinal Study of Youth (NLSY)(1979-1991) data and find that approximately 90% of overall life-cycle inequality can be attributed to differences in initial endowments across men. Similarly, Huggett et al. (2011) employ a risky-human capital life-cycle model to show that initial levels of human capital are the greatest source of lifetime variation.
In contrast to these findings, Storesletten et al. (2004a) analyse Panel Study of Income Dynamics (PSID)(1969-1992) income data in conjunction with Consumer Expenditure Survey (CEX)(1980-1990) consumption data to find that income shocks during the working years are more significant. Furthermore, Storesletten et al. (2004a) argue that without incorporating idiosyncratic shocks received over the working lifetime, it is not possible to account for the co-movement of income and consumption inequality across age.

Storesletten et al. (2004b) is one of several papers that use the joint evolution of income and consumption inequality to address the second key question asked by this literature. Motivated by the predictions of Friedman’s 1957 permanent income hypothesis, studies including Guvenen (2007), Kaplan and Violante (2010), Kaplan (2012) and Heathcote, Storesletten, and Violante (2013) analyse the degree of insurance individuals and households have against income shocks of varying persistence. These papers conclude that a portion of income risk is insurable and do not translate to consumption. This evidence of partial insurance is used to reconcile the empirical relationship between income and consumption inequality first identified by Deaton and Paxson (1994).

Deaton and Paxson (1994) use repeated cross-sectional data from the United States, Britain and Taiwan to construct life-cycle profiles of inequality. Their results show that as uninsurable permanent income shocks accumulate over the life-cycle, both within-cohort income and consumption inequality increase linearly with age; although consumption rises to a less extent. This key finding represents a salient feature of several data sources and has been widely used to inform the theoretic-empiric analyses of incomplete-markets economies.

While much of the literature has focused on risk-sharing and the transmission of income to consumption inequality (a more direct measure of wellbeing), these incomplete-markets models have a much broader scope for analysis. Researchers have not only been able to investigate the direct implications of income risk for consumption and wealth inequality but have also been able to relate this work to savings and labour supply decisions, social insurance and progressive taxation policies, human capital accumulation and within-family risk-sharing. Other studies have also incorporated different types of risks including unemployment, family and health shocks.

1 Other studies including Heathcote, Storesletten, and Violante (2010) and Krueger and Perri (2006) for the U.S. and Blundell and Preston (1998) for the U.K., analyse the joint behaviour of consumption and income inequality over the time dimension.

2 See Heathcote, Storesletten, and Violante (2009) for a detailed survey of recent developments made in the literature relating to these models.
This modelling framework therefore represents a burgeoning area of research. As computational capacity and microeconomic data sources develop, researchers will be able to extend the standard framework further by incorporating different types of risk, insurance mechanisms, heterogeneity and endogenous decisions.

It is not surprising then, that there has been a substantial movement within the literature to gather and document empirical facts that can be used as building blocks for quantitative macroeconomists. A special issue of *The Review of Economic Dynamics Journal (2010)* is one such study that conducts an extensive analysis of economic inequality across nine different countries to provide stylised empirical facts useful for quantitative analysis.

A key recommendation that the curators of this *R.E.D (2010)* special issue, (Krueger, Perri, Pistaferri, and Violante (2010)), make is that more studies of this nature on other countries and new data sources are required. In evaluating other potential countries for which this type of analysis might be both relevant and feasible, it is important to note that the majority of countries included in the *R.E.D (2010)* special issue belong to the Cross National Equivalent File.

Out of the countries that belong to the CNEF, Australia was one of the only members not included in the *R.E.D (2010)* cross-country analysis. This observation, together with Australia’s recently recorded growth in income inequality, make Australia a good candidate for future research in this area.

I therefore follow the contributors to the *R.E.D (2010)* special issue, specifically Heathcote et al. (2010) and take the first steps towards informing these structural models using Australian data. To achieve this, I draw upon literature that has developed the standard methods used to calibrate and provide inputs to these models.

### 2.2 Empirical Life-Cycle Inequality Profiles.

To enable reliable policy analysis using these heterogeneous-agent models, microlevel individual and household data is standardly used to capture important empirical facts. For cross-sectional analysis, standard inequality indices are commonly used to document key features of inequality over time. For models that have an explicit life-cycle structure, empirical evidence portraying the life-cycle patterns of inequality is an additional requirement.

---

3 The CNEF contains and crosslinks longitudinal household data from eight different countries. It is administered by the Department of Policy Analysis and Management at Cornell University, USA.

4 CNEF members Switzerland and Korea were also not included.
The Deaton and Paxson (1994) approach to estimating empirical age-inequality profiles is the most commonly cited strategy for informing life-cycle versions of these models. I adopt a similar method to investigate life-cycle inequality in this paper. Therefore, it is important to note, that when estimating age-profiles the interdependency of age, time and cohort effects is an identification issue that has long been acknowledged in the literature. As it is not possible to simultaneously control for both cohort and time effects, it has become a standard procedure to place an identification restriction on only one. There is currently no consensus among researchers as to whether it is better to restrict age or cohort effects.

Heathcote et al. (2010) weigh in on this ‘time’ versus ‘cohort’ debate. Using PSID data (1967-1996) to estimate life-cycle profiles from both views, they successively control for time and cohort effects. Heathcote et al. (2005) contend that the time view (estimation of age effects in the presence of year fixed effects but not cohort effects) is the most consistent approach. They base this finding on evidence from the 1980s during which the cohort view (which normalizes time fixed effects to zero) overestimates inequality profiles by not taking into account the time effects of the dramatic rise in inequality observed during this period.

Although this evidence supports the time approach, researchers have continued to provide age-profiles estimated from both views. I follow this strategy in my analysis and provide estimated life-cycle inequality profiles that have been separately conditioned on time and cohort effects.

2.3 Estimating the Stochastic Wage Process.

The key inputs to the incomplete market framework are permanent and transitory shocks to wages. These shocks act as the primitive forces driving the mechanisms of these models. Permanent and transitory changes in wages (inputs) feed through intermediary choices and institutions and generate predictions for the distribution of labour supply, earnings, consumption and savings (outputs) across individuals.

The estimation of the stochastic wage process is therefore the foundation for this type of analysis. To empirically identify individual wage shocks in the data, macroeconomists have adopted a statistical model that has emerged largely from within the labour economics literature.

Early contributions to this field of research traditionally estimate the wage process based on first-difference or wage growth rates, mainly using Panel Study of

---

5 Agents can insure against transitory shocks via the mechanism of borrowing/lending a risk-free bond. Permanent shocks are generally uninsurable or only partially insurable depending on the set-up of the model.
Income Dynamics Survey (PSID) data. Wage dynamics were originally modelled by assuming an autoregressive of order one, AR(1), transitory component and individual specific component for the wage residual. This form was used by such studies as Lillard and Willis (1978), Lillard and Weiss (1979) and Hause (1980) and was characterised by modest transitory persistence and heterogeneity in individual earnings profiles. This earlier version of the model was later classified by Guvenen (2009) as the heterogeneous income profile (HIP) class of these models.

In a related study, MaCurdy (1982) challenged this earlier model and allowed for a more flexible error scheme by adopting an autogressive moving-average (ARMA) process. He tested a variety of specifications using PSID data and found an ARMA(1,1) process to be the most suitable. Macurdy’s findings also provided evidence that supported the elimination of the individual specific error component and the presence of a unit root in the error scheme. This evidence led to the emergence of the restricted income profile (RIP) category of the income process which featured a random walk process and homogeneity across individuals. Based on the findings of MaCurdy (1982), more recent work such as Abowd and Card (1989) and Meghir and Pistaferri (2004) have incorporated a similar unit-root specification into the error term.

Gottschalk and Moffit (1994) were the first to recognise the relevance of this model for the decomposition of inequality trends. They used this model and longitudinal analysis of US PSID data to show that the variance of temporary shocks accounted for approximately one-third to a half of observed inequality trends in the US during the 1970s and 1980s. The remainder was then attributed to the variance of permanent shocks.

Since Gottschalk and Moffit (1994), permanent-transitory models have become a standard approach within empirical inequality literature. An important contribution of Gottschalk and Moffit (1994) was the particularly parsimonious version of the restricted income profile (RIP) process that they adopted. They assumed that the residual of income comprised of an AR(1) permanent (unit root) component and found evidence for an uncorrelated transitory shock, even when they imposed an ARMA(p,q) structure for this term.

This simple specification of the wage process has gained particular traction among macroeconomists. They favour the parsimony of this model as it generates the minimum number of empirical inputs required for their incomplete-market models.

---

6 Although he did not rule out the possibility of an ARMA(1,2) specification either.
7 Guvenen (2009) provides a thorough discussion of the two families of models that exist within the literature.
Researchers in this field generally estimate the wage process based on log wage levels (Storesletten et al. (2004b) and Heathcote et al. (2010), for example.).

The statistical model I adopt in this analysis is also based on Gottschalk and Moffit (1994)’s parsimonious error scheme. The estimation procedure I employ is closest to that of Heathcote et al. (2010). Heathcote et al. (2010) analyze income dynamics in the United States using biannual PSID data (1967-2000). Their model, like that of Gottschalk and Moffit (1994), features a unit root AR(1) permanent component and a serially uncorrelated transitory component.

An important feature of Heathcote et al. (2010)’s approach is that they draw upon both the labor economic and macroeconomic traditions and estimate two versions of the permanent-transitory model on the residual of log wages. They first follow an approach standard in the labor economics literature (Abowd and Card 1989; Meghir and Pistaferri 2004 and Blundell, Pistaferri, and Preston 2008, for example) and estimate the model using conditions based on log wage growth rates. They then adopt a method favored by macroeconomists and use conditions based on log wage levels (e.g. Storesletten et al. (2004b); Guvenen 2007; Heathcote et al. (2010)).

Heathcote et al. (2010) observe a substantial difference in their results when covariance conditions of the first-difference (labor) approach and of the level (macroeconomic) approach are separately used. The variance of the transitory shock is larger when estimated in the levels form whereas the variance of the permanent shock is larger when estimated using the difference method. Heathcote et al. (2010) suggest that this finding provides evidence of an ‘empirical puzzle’ within the literature. They draw supporting evidence from the results of Brzozowski et al. (2010) in Canada; Fuchs-Schündeln et al. (2010) in Germany and Flodén and Domeij (2010) for Sweden.

In Canada, Brzozowski et al. (2010) use Survey of Labour Dynamics (SLID) (1994-2005) data to also find that the permanent variance of wages is on average higher in the difference approach, whilst the transitory variance is larger in the level approach. Moreover, Fuchs-Schündeln et al. (2010) and Flodén and Domeij (2010) use data from German Socio-Economic Panel Study (GSOEP) (1984-2002) and Longitudinal Individual Base (LINDA) (1978-2004) respectively to find that the same contradictory pattern seen in Canada and the US persists for Germany and Sweden. These cross-country findings indicate that this conventional permanent-transitory model is potentially misspecified.

Given the permanent-transitory models role in supplying key inputs to structural
incomplete-market models, its potential misspecification is a serious concern for this framework. If used as inputs, the inconsistent estimates of permanent and transitory variances would generate different equilibrium outcomes. This undermines the reliability of incomplete-markets model predictions for policy analysis.

Reconciling these two approaches and identifying the nature of this misspecification is therefore an important research agenda. While this task goes beyond the scope of this thesis, a key interest of my study based on these findings is to assess whether the same inconsistency holds in the Australian data.

2.4 Australian Inequality Literature

In contrast to the substantive body of international research discussed above, the Australian literature is much less well established in these areas. Given that the related incomplete-markets literature has developed mainly throughout the past decade, this framework has not yet gained momentum within the Australian literature. Research relating to the key inputs of these models has also been limited. Although there has been extensive research documenting income inequality in Australia over time, there has much less empirical research relating to life-cycle inequality and the estimation of income dynamics.

For cross-sectional inequality patterns, recent income inequality growth has generated a renewed interest in this area of research for Australia. Recent studies including Greenville, Pobke, and Rogers (2013), Wilkins (2013) and Whiteford (2013) analyse inequality trends in Australia using ABS cross-sectional data. These papers are in consensus that income inequality in Australia increased from approximately 2003 up to 2008, stabilised or slightly declined between 2008-2009 and rose again in 2010 to 2011. The findings of these papers will be primarily used to inform the sections of this current research focused on documenting inequality patterns over time.

In relation to the other key interests of this current paper, life-cycle inequality patterns and wage dynamics, Borland (1999) identified these topics as areas of research in Australia for which in 1999 almost nothing was known. Specifically, he recognised that the absence of an appropriate longitudinal data source had prevented the identification of permanent and transitory income dynamics for Australia. He recommended that this be an avenue of research in the future.

8This would largely be due to differences in the size of the permanent variance, as this component that is uninsurable
9Other papers including Borland (1999) and Coelli and Wilkins (2009) look at the education earnings premium in Australia over time. These papers will be discussed further in Section 8.
The continued lack of an panel data source of adequate length over the past decade has meant that this research agenda highlighted by Borland (1999) has remained relatively untouched.\textsuperscript{10}

An exception to this observation is a study by Barrett, Crossley, and Worswick (2000). This paper observed that consumption inequality increased to a lesser degree than income inequality between 1975-1993. They used this as evidence to suggest that a substantial portion of overall income inequality was driven by the variance of transitory income shocks.\textsuperscript{11} However, this interpretation could not be verified as identifying the relative contribution of permanent and transitory income dynamics was not possible using cross-sectional data.

The HILDA survey now has a sufficient panel dimension to address this shortcoming and enable this type of research. In my following main analysis, I therefore utilise the panel dimension of HILDA and decompose residual inequality to obtain estimates of permanent and transitory inequality.\textsuperscript{12} Together, with my analysis of cross-sectional and life-cycle inequality trends, the estimation of these wage dynamics represents a main contribution of this paper.

\textsuperscript{10}The HILDA survey was introduced by The Melbourne Institute in 2001; however, throughout the decade its time dimension remained too short to enable this type of analysis.

\textsuperscript{11}This finding was based on Australian Bureau of Statistics (ABS) Household Expenditure Survey (HES) consumption and income data.

\textsuperscript{12}Another working paper by Higgins (2010) have also employed the HILDA Survey from 2001-2006 to model wage dynamics using an approach similar to Meghir and Pistaferri (2004) and an ARMA(1,1) specification.
3 Data and Measurement

3.1 Data description

This empirical analysis employs individual-level data sourced from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. HILDA is a longitudinal survey that draws a representative sample from private households dwelling in Australia. Administered by The Melbourne Institute, the HILDA survey began in 2001 and has eleven available waves to date. The survey interviews all sample household members aged 15 years and over on an annual basis. It tracks individuals over time even when they split-off from original dwellings and form new households. Interviewers collect individual and household level information regarding income, family background, education and employment. The first wave in 2001 interviewed 6,872 households and 13,969 individual respondents. The HILDA sample was replenished with a further 2,153 households and 5,477 individuals in 2011.

While most previous inequality studies in Australia employ either SIH or HES surveys from the Australian Bureau of Statistics (ABS), HILDA is a more appropriate choice in this instance due to its panel dimension. This key design feature implies the repeated sampling on the same set of individuals annually. As discussed previously, the panel structure of HILDA is crucial for analysing wage dynamics. The separate identification of permanent (long-run) and (short-term) residual inequality can only be achieved via longitudinal analysis and is not possible using cross-sectional data.

The stronger emphasis administrators have placed on consistency in survey design over time, is an additional benefit of the HILDA survey. Wilkins (2013) discusses adjustments that have been made to the collection of income information in both HILDA and ABS data sources. He identifies that the substantial changes made to both SIH and HES survey designs are a major limitation of these ABS datasets. For HILDA, on the other hand, he finds that when changes have been made, they have had less of an impact on the inter-temporal comparability of this survey.

---

1 For more information see The Hilda Manual: Summerfield, Freidin, Hahn, Ittak, Li, Macalalad, and Wooden (2012).
2 ABS Household Expenditure Survey (HES) and Survey of Income and Housing (SIH).
3 See Wilkins (2013) for further comparison between HILDA, ABS surveys, tax records and National Accounts as data sources for inequality studies.
4 Examples of such changes are the inclusion of salary sacrifice information in 2010 and a switch to computer generated interviewing in 2009.
In selecting the HILDA survey to study changes in the wage distribution across both time and life-cycle dimensions, there are some important issues that must be considered before estimating measures of inequality. Most of these issues are commonly encountered whenever using a micro-level household data source.

**Imputation.** Income variables taken from the HILDA survey have been treated for non-response using mainly a combination of Nearest Neighbor regression and Little and Su imputation methods. Summerfield et al. (2012) show the rates of income non-response for each wave. The rate of income non-response has decreased as the survey has become more established. In 2011, the percent of mean imputed income values was 3.4% and 4.8% for wages & salaries and total pre-tax income, respectively. In my analysis, I do not exclude imputed individuals as these figures represent relatively small portion of the sample.

**Top-coding.** Another data quality issue that can have implications for assessing inequality, particularly at the top of the wage distribution, is top-coding. This common censoring practice is used by surveys to protect the identification of top-income earners. The HILDA survey uses a weighted-mean top-coding procedure. They substitute any income values above a certain threshold with the weighted mean value of income for all individuals and households in the top-coded category. Wilkins (2013) states that top-coding in the HILDA survey should have negligible implications for inequality measures as this procedure is carried on only 0.2%-0.4% of the sample.

**Attrition.** For HILDA, similar to other longitudinal studies, attrition over time has a cumulative effect in eroding the representativeness of the sample. Attrition will introduce bias to the survey if it is non-random; that is, if particular subgroups are more likely to attrit than others. Summerfield et al. (2012) identify that individuals in the HILDA survey who are young, single, non-English speaking, unemployed or working in a low-skilled occupation have a higher likelihood of dropping out or refusing to participate in the study. These attributes are also associated with higher wage variability and individuals at the lower end of the wage distribution. Therefore, the higher propensity for attrition amongst people with these characteristics introduces a potential for bias within the sample towards excess stability in the wage distribution.

**Sample Representativeness.** A limitation of HILDA is that migrants arriving in Australia after 2001 have little chance of entering the sample. This represents

---

5See the Hayes and Watson (2009) for details of this approach.
6Heathcote et al. (2010) does not exclude imputed values for the Current Population Survey (CPS) when 26.8% total wages and salary earnings have been imputed
another source of declining representativeness relevant to this study. Australia’s immigration policy is targeted at stable skilled migration. By not accounting for the inflow of skilled migrant into Australia, the HILDA survey may therefore underrepresent young, more highly educated human capital in the panel. This issue has the potential to understate inequality across high and low education groups and introduce an age bias that may impact inequality measures taken over the life-cycle dimension.

Sample Top-Up. The HILDA survey aimed to minimize the effects of potential attrition bias and erosion of representativeness by replenishing the sample top-up in 2011. The main objective of administrators was to ensure the long-term representativeness of the panel. However, in the short-term the injection of new individuals and households had the potential to alter the structure of the wage distribution. Reassuringly, Wilkins (2013) performed robustness checks on this sample top-up and found that the inclusion or exclusion of this additional sample has minimal impact on the 2010-2011 wage distribution.

3.2 Sampling Strategy

The initial unbalanced sample obtained from HILDA consists of 26,028 individuals and 147,823 observations across years 2001 to 2011. An unbalanced panel is used to minimise the impact of attrition bias and to capture a richer, more informative sample. While this means individuals do not have to be present for every wave of the survey, the dynamic component of the study requires respondents be present for at least two consecutive years. The following selection criteria is applied to the initial sample.

Male full-time workers are only included if they are aged between 23 and 60 years. This excludes individuals at the tail ends of the working life-cycle where there are less observations due to lower labour force participation. Full-time self-employed workers are excluded due to their more complex income structure. Respondents enrolled in full time study are also dropped from the sample. Individuals are excluded if they earn less than half the minimum wage. This eliminates unrealistic outliers at the bottom of the wage distribution. The top of the wage distribution is not trimmed, as observations in this region are not unrealistic and it is unclear whether

\footnote{This criteria draws upon related literature and is based mainly upon the selection criteria detailed by Huggett et al. (2011) who use United States PSID data to study the sources of lifetime inequality.}

\footnote{A male sample is chosen in accordance with existing literature. Previous focus on a male sample is due to the fact that incomplete market models have traditionally defined a male ‘bachelor’ household.}

\footnote{Minimum wage information was obtained from: https://www.fairwork.gov.au/pay/national-minimum-wage/pages/default.aspx.}
trimming would remove actual outliers or genuine variation possibly important in the assessment of inequality. Those with missing key income or demographic information are also excluded. The final sample of interest constructed in accordance with these guidelines includes 3,906 individuals and 23,261 observations. A more detailed outline of this procedure is reported in Appendix A.1.

3.3 Key Variables of Interest

3.3.1 Income Measures

All income measures used in this paper have been converted into log and real 2011 terms. Income measures are adjusted using RBA CPI Year Ended June index to align with the fact that most income variables recorded in HILDA are reported in retrospective terms, as the amount earned in the previous financial year.

**Hourly Wage.** The key income variable of interest to this study is the real hourly wage rate as it is closest to an exogenous measure of income and represents the common input to the incomplete-markets framework. The hourly wage variable is constructed from two separate variables in HILDA. It is defined as:

\[
\text{Hourly Wage} = \frac{\text{Earnings (wscei)}}{\text{Usual Hours Worked Per Week (jbhruc)}}
\]

**Hours Worked.** The usual hours worked per week variable is a combination of usual reported hours per week, or if individuals indicate that their hours vary; hours per week they work on average.

**Earnings** is defined as weekly gross wages or salaries collected from all jobs and received in the previous financial year. This labour income variable also includes incorporated business wages and salary, but this portion is eliminated by the exclusion of self-employed workers from the sample.

**Pre-Tax Income** is defined as net financial year gross income including gross financial year labour income, business and investment income as well as both private and public transfers and pensions. This net income variable is constructed from two separate variables; this is achieved by first accounting for missing/negative values then subtracting a negative from positive income component.

---

10 For simulating life-cycle profiles in Section 7. I impose additional restrictions on my main sample which are described in Appendix C.1.

11 Income measures are deflated using Consumer Price Index information obtained from Reserve Bank of Australia (RBA) website under Statistical Table G1.

12 As the sample includes only full-time workers, by definition individuals work at least 35 hours per week.

13 This income measure is only used in the preliminary stages of analysis, see Figure 3.1.
3.3.2 Education

A binary education variable is constructed to compare inequality between and within education groups. Individuals are stratified into high and low education groups. This variable is derived from a question in HILDA that records the highest level of education attained by respondents. This classification is based on grouping education levels with similar years spent in education and also similar returns to education as represented by the first moment of earnings (See Appendix A.2). For an individual to be allocated to the high education group they are required to have accumulated at least fifteen years of education. ‘High’ education is therefore defined as those who have completed either an Undergraduate, Postgraduate or Graduate Diploma qualifications. Those in the ‘Low’ education group hold either Year 12 or below, Higher School Certificate or a TAFE Certificate/Diploma qualifications. Using this variable, I construct my high and low education subsamples which comprise of 1 054 (6 222) and 2 852 (17 039) individuals(obs.), respectively.

3.3.3 Control Variables

Key demographic variables taken from the HILDA survey are used in reporting measures of inequality over different subgroups and as a set of observable controls used when isolating residual inequality in Section 4.3 and 5.1. These personal characteristics include marital status, age, gender, union membership, migrant status, being an indigenous Australian, employment status, health status and state information. The choice of an appropriate set of covariates is guided primarily by existing empirical studies. Storesletten et al. (2004b) explains that the deterministic component of income serves mainly as an intermediary to identifying residual variation - the most important driver of inequality. These observable characteristics are therefore used solely to obtain a proper estimate of residual inequality. As they are not a central focus in this paper, their discussion here will be brief although a more detailed description is provided by Appendix 3.3.

3.4 Descriptive Statistics

Here, I present income and demographic descriptive statistics for both the main full-time male sample and education group sub-samples. While the majority of the main analysis centres around wage measures, it is still important to have an understanding of the underlying demographic composition of the samples.

Beginning with my full sample, Figure 3.1 shows the evolution of both average labour income and average pre-tax income between 2000-01 and 2010-11. Over the sample period, both income measures increased between 2001-02 to 2006-07, before
stabilising briefly around the onset of the GFC and then resuming growth. Pre-tax income is defined in the previous section as an aggregate measure of income and labour income is the annual value of gross weekly wages and salaries from the previous year. Both pre-tax income and labour income follow similar trends and fluctuations. Few deviations occur between the two series over the decade, one such disparity does occur in the brief spike that can be seen in pre-tax income around 2008-09\textsuperscript{14}.

Most importantly, Figure 3.1 illustrates that labour income makes up a large portion of overall pre-tax income, accounting for approximately 75% of this aggregate income measure. This indicates that the inequality patterns identified by this paper for the wage component of income are likely to also be a main determinant of aggregate income inequality trends for Australia in recent times.

Figure 3.2.a and 3.2.b explore trends in real average hourly wages and weekly hours for full-time males. This figure shows that average hourly wages have increased substantially in real terms by approximately 20.14%, over the sample period. In contrast, average weekly hours have remained quite stable, decreasing only slightly by 1.4%. Together, these two plots support the Greenville et al.\textsuperscript{[2013]} finding that average full-time hourly wages have driven the growth in labour income over the past decade.

\textsuperscript{14}This can be attributed to the increase in public transfers in the form of Government Stimulus Bonus Payments during this year.
Table 3.1 reports other demographic information for the main full-time male sample at both the beginning and end of the sample period. The changes in income represented by cross-sectional statistics in Table 3.1 align with the trends observed in Figures 3.1 and 3.2; with the statistics indicating only a slight decline in annual hours and a substantial increase in average annual labour earnings (as defined by gross wages & salaries) across the period.

Table 3.1: Descriptive Statistics for Full-time Male Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>2000-2001</th>
<th>2010-2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age(years)</td>
<td>39.4</td>
<td>41.4</td>
</tr>
<tr>
<td>% Graduates or Higher</td>
<td>27.7</td>
<td>29.6</td>
</tr>
<tr>
<td>% Married</td>
<td>65.0</td>
<td>58.8</td>
</tr>
<tr>
<td>% Union</td>
<td>37.9</td>
<td>31.4</td>
</tr>
<tr>
<td>% English Speaking Migrant</td>
<td>14.8</td>
<td>10.9</td>
</tr>
<tr>
<td>% Non-English Migrant</td>
<td>9.1</td>
<td>7.9</td>
</tr>
<tr>
<td>% Australian Born</td>
<td>76.1</td>
<td>81.2</td>
</tr>
<tr>
<td>% Indigenous</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Avg. Hours Worked</td>
<td>2 323</td>
<td>2 290</td>
</tr>
<tr>
<td>Avg. Annual Earnings ($)</td>
<td>70 996</td>
<td>84 714</td>
</tr>
<tr>
<td>No. Obs.</td>
<td>1 856</td>
<td>2 268</td>
</tr>
</tbody>
</table>

Looking at other demographic information, union representation declined by 6.5% while between 2000-01 and 2010-11 the proportion of male full-time respondents holding an undergraduate degree or higher increased modestly over time from 27.7% to 29.6% (1.9%). The representation in the sample of indigenous Australians also increased marginally by 0.5%.

The average age in years of individuals in the main sample increased by 2 years from 39.4 and 41.4 and the proportion of male full time workers who were married decreased by 6.2%. These statistics could be possibly indicative of recently observed
demographic changes of an ageing population and a shift towards single person households. This latter change has been driven by an increase in divorce rates and by individuals choosing to marry and have children later in life.\footnote{As marriage represents a sharing of resources and risk this demographic change would be expected to increase inequality at the household level.}

The percentage of English and non-English speaking migrants decreased by 3.83% and 1.26% respectively. The decline in the proportion of migrants in the sample, while only small, may be an indicator of the possible bias discussed in the previous section introduced by the underrepresentation of new migrants in the sample. As migrants tend to be young and highly educated, their inclusion may mitigate the ageing of the panel noted above and explain the relatively stable percentage of highly education male workers in the sample.

### 3.4.1 Descriptive Statistics by Education Group

Figure 3.3 plots the average hourly wage and hours respectively for both high and low education subgroups over the sample period. In Figure 3.3.a. both high and low education groups experienced an increase in average income in real terms and a decrease in average weekly hours, although the increase of the two variables for high education is more pronounced. Between 2000-01 and 2010-11, average hourly wage rose by 18.54% for the Low Education group, and grew slightly more for High Education group with a 20.69% increase in average hourly wage across the period.

In Figure 3.3.b. average weekly hours worked for the high education group had a greater decline in hours across the entire period, whereas for the low education group, average weekly hours remained fairly stable up until 2007-08, after which there was a slight temporary drop in hours up until 2009-10. A comparison of observable demographic characteristics across education groups is also provided in Appendix A.4.
4 Inequality over Time

This section begins my main analysis by first documenting cross-sectional trends in full-time male wage inequality in Australia for the period 2000-01 to 2010-11. Here, I draw on existing Australian studies to support my analysis of changes in the wage distribution over time.

4.1 Cross-Sectional Wage Inequality Trends

Figure 4.1. depicts four commonly used measures of wage inequality. Figure 4.1.a and Figure 4.1.b present the wage inequality trend for the entire distribution using two measures of dispersion; the variance of log hourly wages and the Gini coefficient of hourly wages. These two standard inequality indices show similar patterns across the period. Wage inequality fell from 2000-01 to 2001-02, then increased between 2002-03 to 2007-08, dampened in 2008-09 and resumed growth in 2009-10. Overall, inequality rose across the whole period up to 2010-11. The variance of log hourly wages increased by 2.7 log points from 0.172 to 0.199 and the Gini coefficient rose

---

1 The Gini Coefficient represents the ratio of the area between the line of equality and the Lorenz Curve and the area under the line of equality. A Gini coefficient of 1 indicates perfect inequality, 0 means perfect equality.

2 A dampening of inequality is observed for 2008-2009 in all figures except for Figure 4.1.c, in which the p50-p10 ratio remains stable.

---
by 2 points from 0.238 to 0.258 over the sample period. These statistics represent a substantial growth in the dispersion of full-time hourly wages in the 2000s.

The aggregate trend of wage inequality shown in the top half of Figure 4.1 is both qualitatively and quantitatively similar to findings of other Australian empirical studies. For instance, the observed change in the Gini coefficient represents a growth of 8.36%. In comparison, [Greenville et al. (2013)] find a 10.2% rise in the Gini coefficient of male hourly wages from 1998-99 to 2009-10. The variation between these two numbers can be reconciled by the use of difference data sources and slight differences in samples and time periods. [Greenville et al. (2013)] use ABS HES Survey and include all individuals as opposed to only those who are of working age. They also point out that HILDA generally has lower variation than ABS surveys as HILDA has not been exposed to the same inconsistencies in survey design that have been a major issue for ABS data.

Relative to other countries, Australia’s full-time wage inequality is above the OECD average. While Australia has more inequity than countries such as those in Northern and Eastern Europe, it is still ranked below the United States and Canada, who in 2011 were reported to have respective Gini coefficients of 0.4 and 0.36 for wages of all full-time workers.

Now looking at inequality trends across different parts of the wage distribution, Figure 4.1.c and Figure 4.1.d. use inter-decile ratios to compare inequality trends at the top and bottom of the wage distribution, relative to the median. The 50-10 ratio represents the differential in log hourly wages of full-time male workers at the 50th and 10th percentiles of the distribution; the 90-50 ratio is similarly defined. Figure 4.1.c. indicates that the 50-10 ratio remains relatively stable over time, increasing by only 4 points from 1.57 to 1.61 across the sample period. Figure 4.1.d shows that the 90-50 ratio increased by 10 points from 1.76 to 1.86 points during the same time frame. Comparison of these two figures therefore demonstrates that wages of the top 10% of the sample grew more relative to rest of the distribution in the 2000s.

Income growth at the top of the wage distribution is being increasingly recognised as a key trend driving inequality growth. [Atkinson and Leigh (2007)] investigate top income earners for Australia and find that the ratio of average income of the top 1% of earners to the rest of the population was 9:1 in 2002, which is larger

---

3 See [OECD (2011)]. These comparative OECD statistics are based on 2005 data. At this time Australia was reported to have a Gini coefficient of 0.26 based on ABS survey Data, for all full-time working age individuals (including females and self-employed)

4 It should be noted that the use of micro survey data can be affected by top-coding and imputation. While [Summerfield et al. (2012) and Wilkins (2013)] provide evidence that these issues should not be a serious concern for the HILDA survey, administrative tax-recod data is a good alternative for further investigation of the top of the wage distribution.
than it has been at any point in time over the past 50 years. This rapid rise in the income of top earners has been observed in other developed countries including the United States and Canada. Atkinson and Leigh (2007) attribute this growth to international demand for talent, a decline in the progressivity of tax and transfers and more favourable incentive schemes, especially for executives, managers and financial professionals.\footnote{In a related study, Atkinson and Leigh (2013) use administrative tax data to show that the trends in the share of overall income accruing to Australia’s top income earners bear close similarities with their Canadian, UK and US counterparts.}

### 4.2 Dispersion of Hours, Wages and Earnings

![Graphs showing dispersion of hours, wages, and earnings](image)

**Figure 4.2:** Average Earnings, Wage and Hours Growth: 2000-01 to 2010-11

Figure 4.2 looks more closely at how changes in the distribution of hours and wages impact upon the distribution of earnings. Firstly, individuals are ranked by earnings and then average hours, hourly wages and earnings for the 0-10, 45-55 and 90-100 deciles are calculated relative to the base year, 2000-01. This figure shows the percentage changes of these components across different deciles of the distribution.

Earnings is defined as hourly wage multiplied by hours worked. Figure 4.2.a and Figure 4.2.b describe the growth in hours and hourly wages, respectively. Comparison of these two figures to Figure 4.2.c which presents the growth across
deciles in earnings, shows that the rise in labour earnings is being driven almost completely by the strong growth in hourly wages.

By looking at the different deciles of the distribution it is evident that earnings and hourly wages have grown faster among high earners at the top of the distribution than for median and low earners, especially in the later half of the 2000s.

In Figure 4.2.b, hourly wages for high income earners increased by approximately 25% in the past decade. Individuals at the median or at the bottom of the earnings distribution, also experienced growth up until 2008 and the growth began to stabilise or decline for the rest of the sample period. The bottom and middle decile experienced a total growth in hourly wages of only 9.44% and 5.57%, respectively. This implies that total growth at the end of the period was much lower in these deciles relative to individuals at the top of the distribution.

These results indicate that recent cross-sectional trends in earnings and wage inequality appear to be driven predominantly by the rapid growth in hourly wages, particularly for the top of the wage distribution.
4.3 Between and Within-Group Inequality

To further identify sources of wage inequality, I next analyze both between-group and within-group inequality. Figure 4.3 demonstrates the role of the key observable characteristics, education and experience in explaining inequality trends as well as the contribution of unobservable factors to the rising dispersion of wages.

Figure 4.3.a illustrates the education wage differential as defined by the ratio of the average wages of high to low education groups. The education wage premium remains relatively unchanged up until 2006-07, rises steadily between in 2007-08 to 2008-09 before flattening out again up until 2010-11. The finding that the education premium is relatively stable across the sample is consistent with other Australian research including Borland (1999) and Coelli and Wilkins (2009). This observed trend will be discussed in greater detail in Section 8.1. Accompanying the evolution of the education wage premium is the plot of the post-secondary completion rate or ‘supply’ of highly educated workers in the sample population presented in Figure 4.3.b. This figure shows that the completion rate increased by approximately 1.8% over the ten year period.

Figure 4.3.c displays the experience wage premium. This is calculated as the ratio of average wages of 44 to 55 year olds to 25 to 35 year olds. This premium has remained relatively stable over time. Although the ratio does decline between 2005-04 to 2007-06 before increasing slightly back towards to its initial level.
Based on these observable characteristics, both between-group wage differentials appear fairly stable and do not provide a great deal of insight into the movements in wage inequality patterns over the sample period observed in Figure 4.1. Therefore, I turn to consider within-group or residual inequality in Figure 4.3.d. First, the residual component \( \hat{u}_{i,t} \) is estimated by year using the following regression:

\[
\ln(y_{i,t}) = X_{i,t}\psi_{i,t} + u_{i,t}
\]  

(4.1)

where \( \ln(y_{i,t}) \) is the log of hourly wages and \( X_{i,t} \) is a vector of controls including a cubic polynomial of age, education, migrant status, marital status, indigenous status, union membership, health status as well as state fixed effects.\(^6\)

I then plot the time-varying variance of \( \hat{u}_{i,t} \) with the variance of raw log hourly wages. In Figure 4.3.d it can be seen that after controlling for these observable characteristics, residual wage inequality bears striking resemblance to the ‘raw’ inequality of log wages. The two plots follow a very similar evolution in both non-stationary trend and cyclical movements and on average across the period, residual wage inequality accounts for 81.14% of overall cross-sectional inequality.\(^7\)

This provides strong evidence that unobservable characteristics or residual inequality is an important driver of wage inequality in Australia.

To recap, this section identifies key features of recent cross-sectional inequality trends. These facts are especially useful for quantitative research conducting cross-sectional analysis and interested in macroeconomic implications of rising inequality over time (see Heathcote et al. (2010), for example). I find a substantial rise in wage inequality from 2000-01 to 2010-11 across the entire distribution and observe particular growth top wage earners. I also observe that rising hourly wages is a key driver to overall earnings growth and dispersion. Finally, I document evidence that indicates between-group inequality in Australia has remained relatively constant over the sample period. Most importantly, I find that residual or within group inequality is a main determinant of overall wage inequality trends. The next section will provide further insight into this significant component of wage inequality by deconstructing it into its permanent and transitory components.

\(^6\)This specification is based on the set of controls used by previous studies, namely Heathcote et al. (2010) and Huggett et al. (2011).

\(^7\)This percentage value aligns with the share of residual inequality recorded by other studies including Heathcote et al. (2010) who finds that residual inequality accounts for approximately 79% of overall wage inequality.
5 Dynamics of Inequality

This section will develop a non-stationary unobservable components model to decompose residual (within-group) wage inequality. The aim here is to obtain the most important inputs required for implementing an incomplete markets model on Australia data. I use the panel dimension of HILDA and two separate sets of covariance conditions to estimate the variance of permanent and transitory wage innovations. I then compare my results to a key finding put forward by Heathcote et al. [2010].

5.1 Statistical Model

The following specification used to model the dynamic component of wages has been a typical approach within empirical inequality literature since Gottschalk and Moffit [1994]. Firstly, residuals from the following regression are captured:

\[
\ln(y_{i,t}) = X_{i,t}\psi_{i,t} + u_{i,t} \tag{5.1}
\]

where \(\ln(y_{i,t})\) is the log of hourly wages or earnings, \(X_{i,t}\) is a vector of controls including a cubic polynomial of age, education, migrant status, marital status, indigenous status, union membership, health status as well as state fixed effects which represents the deterministic component of income, \(\psi_{i,t}\) is a vector of time varying coefficients and \(u_{i,t}\) is the unobservable component of wages. The dynamics of the residual are modelled by the following process:

\[
u_{i,t} = z_{i,t} + \epsilon_{i,t} \tag{5.2}
\]
\[
z_{i,t} = z_{i,t-1} + \eta_{i,t} \tag{5.3}
\]

The above equations represent a permanent-transitory model; where \(z_{i,t}\) represents the permanent (unit-root) component of AR(1) form and \(\epsilon_{i,t}\) represents the uncorrelated transitory component of residual wages. This specification of the stochastic wage process is favoured by quantitative macroeconomists for its parsimony. The \(\eta_{i,t}\) and \(\epsilon_{i,t}\) terms represent the respective permanent and transitory shocks and are assumed to be orthogonal to one another and serially uncorrelated. Both follow an

\footnote{The same specification is used as equation (4.1). Time fixed effects are not required as panel analysis means that regressions are carried out by year.}
i.i.d distribution and have time-varying conditional variances:

\[
\eta_{i,t} \sim (0, \sigma_{\eta,t}) \quad \epsilon_{i,t} \sim (0, \sigma_{\epsilon,t})
\] (5.4)

The variance of residual wages is equal to the sum of the variances of the permanent and transitory components. It is important to identify the relative contributions of both for two reasons. Firstly, residual inequality is identified in Section 4.3 as the main driver of cross-sectional inequality. Therefore decomposing within-group inequality can help to better interpret the impact of rising inequality over time. Secondly and most importantly, permanent and transitory shocks to wages have very different implications when fed through the incomplete-markets framework. These shocks represent the exogenous forces driving the mechanisms of these models. Transitory shocks are insurable whereas permanent shocks are generally uninsurable and transmit through to consumption and welfare. Having accurate estimates of these empirical inputs is therefore crucial.

The separate identification of permanent and transitory inequality is not possible using the cross-sectional variance of log-wage residuals alone. To obtain these estimates, longitudinal analysis of HILDA and the appropriate variance-covariance conditions are required. I use these techniques to get estimates of \(\sigma_{\epsilon,t}\) and \(\sigma_{\eta,t}\).

Following Heathcote et al. (2010), I employ two conventional approaches from within the literature to estimate this model. I first use the ‘levels’ approach, a method used by macroeconomic researchers (e.g. Storesletten et al. (2004a) and Heathcote et al. (2010)) and obtain estimates using covariance conditions on the level of log wage residuals. Second, I adopt the ‘difference’ approach, a method well-established in labour economics literature (e.g. MaCurdy (1982); Abowd and Card (1989) and Meghir and Pistaferri (2004)). This approach uses covariance conditions based on log wage growth-rates. The two methods are both standard approaches and theoretically should provide the similar estimates of \(\sigma_{\eta}\) and \(\sigma_{\epsilon}\).²

²In contrast to Heathcote et al. (2010), I use annual rather than bi-annual data, this slightly alters the required covariance conditions.
Levels Approach. This first approach estimates empirical moments using the following variance and covariance conditions:

\[ u_{i,t} = z_{i,t-1} + \eta_{i,t} + \epsilon_{i,t} \]

\[
\text{Var}(u_{i,t}) = \text{Var}(z_{i,t-1}) + \sigma_{\eta,t} + \sigma_{\epsilon,t} \\
= \text{Var}(z_{i,t-2}) + \sigma_{\eta,t-1} + \sigma_{\eta,t} + \sigma_{\epsilon,t} \\
(5.5)
\]

\[
\text{Cov}(u_{i,t+1}, u_{i,t}) = \text{Cov}(z_{i,t-1} + n_{i,t} + n_{i,t+1} + \epsilon_{i,t+1}, z_{i,t-1} + \eta_{i,t} + \epsilon_{i,t}) \\
= \text{Var}(z_{i,t-1}) + \sigma_{\eta,t} \\
(5.6)
\]

\[
\text{Cov}(u_{i,t-1}, u_{i,t}) = (z_{i,t-2} + \eta_{i,t-1} + \epsilon_{i,t-1}, z_{i,t-2} + \eta_{i,t-1} + \eta_{i,t} + \epsilon_{i,t}) \\
= \text{Var}(z_{i,t-2}) + \sigma_{\eta,t-1} \\
(5.7)
\]

\[
\text{Cov}(u_{i,t+1}, u_{i,t}) - \text{Cov}(u_{i,t-1}, u_{i,t}) = [\text{Var}(z_{i,t-1}) + \sigma_{\eta,t} + \sigma_{\epsilon,t}] - [\text{Var}(z_{i,t-2}) + \sigma_{\eta,t-1}] \\
= \sigma_{\epsilon,t} \\
(5.9)
\]

\[
\text{Var}(u_{i,t}) - \text{Cov}(u_{i,t+1}, u_{i,t}) = [\text{Var}(z_{i,t-2}) + \sigma_{\eta,t-1} + \sigma_{\eta,t} + \sigma_{\epsilon,t}] - [\text{Var}(z_{i,t-2}) + \sigma_{\eta,t-1}] \\
= \sigma_{\eta,t} + \sigma_{\epsilon,t} \\
(5.10)
\]

Subtract (5.9) from (5.10) to obtain \( \sigma_{\eta,t} \)

Differences Approach. The model is then estimated in a similar fashion using an alternative set of variance-covariance restrictions and taking the first differences of residual wages.

\[
\Delta u_{i,t+1} = u_{i,t+1} - u_{i,t} \\
= (z_{i,t} + \eta_{i,t+1} + \epsilon_{i,t+1}) - (z_{i,t} + \epsilon_{i,t}) \\
= \eta_{i,t+1} + \epsilon_{i,t+1} + \epsilon_{i,t} \\
(5.11)
\]
The theoretical variance-covariance structure implies:

\[
\begin{align*}
Var(\Delta u_{i,t+1}) &= \sigma_{\eta,t+1} + \sigma_{\epsilon,t+1} + \sigma_{\epsilon,t} \\
Cov(\Delta u_{i,t+1}, \Delta u_{i,t}) &= (\eta_{i,t+1} + \epsilon_{i,t+1} - \epsilon_{i,t}, \eta_{i,t} + \epsilon_{i,t} - \epsilon_{i,t-1}) \\
&= -\sigma_{\epsilon,t}
\end{align*}
\]

By imposing these moment restrictions and estimating the above moments empirically, Equation (5.13) identifies \(\sigma_{\epsilon,t}\) for \(t = 2003, 2004, \ldots, 2010\). The parameter \(\sigma_{\eta,t}\) is then identified by:

\[
\sigma_{\eta,t+1} = Var(\Delta u_{i,t+1}) + Cov(\Delta u_{i,t+2}, \Delta u_{i,t+1}) + Cov(\Delta u_{i,t+1}, \Delta u_{i,t})
\]

(5.14)
5.2 Results

Figures 5.1.a and 5.1.b plot the variances of both permanent and transitory shocks to wages over time. The level and difference moment estimates obtained are presented together in each graph. Inspection of both graphs in Figure 5.1. indicates that the transitory variance is larger than permanent variance when estimated from both level and difference approaches.

In Figure 5.1.a, the estimated variance of the permanent shock from both approaches appear to follow similar pattern over the period; the main difference between the two estimates is their magnitude. The difference approach reports a higher permanent variance at all points of the sample period. It is also noted that the level approach generates negative variances in some years. For the variance of the transitory shock in Figure 5.1.b the opposite pattern is observed. In this case, it is the level approach that estimates a larger transitory variance over the period.

In theory, as long as the model has been correctly specified, the two approaches should generate similar estimates. However, these results indicate that the two methods deliver substantially different estimates of permanent and transitory variances. This implies that using difference or level conditions does impact upon the estimates of permanent and transitory parameters.

While this evidence appears surprising, this result represents the same empirical regularity that has been observed in previously discussed international literature (See Section 2.3, page 10.). Specifically, the diverging parameter estimates of the level and difference approach identified in Figure 5.1 mirror the findings of Heathcote.

This result supports Barrett et al. (2000)’s finding based on joint income and consumption inequality patterns in ABS data that changes in the variance of transitory income fluctuations appear to account for a substantial portion of overall income inequality in Australia.

Table 5.1: Estimates of Wage Dynamics: Cross-Country Comparison

<table>
<thead>
<tr>
<th>Country</th>
<th>Data</th>
<th>Sample Period</th>
<th>Method</th>
<th>$\sigma^2_\eta$</th>
<th>$\sigma^2_\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Australia</strong></td>
<td>HILDA</td>
<td>2001-2011</td>
<td>Difference</td>
<td>.013</td>
<td>.023</td>
</tr>
<tr>
<td>(Current Paper)</td>
<td></td>
<td></td>
<td>Level</td>
<td>.001</td>
<td>.030</td>
</tr>
<tr>
<td><strong>United States</strong></td>
<td>PSID</td>
<td>1967-2002</td>
<td>Difference</td>
<td>.028</td>
<td>.060</td>
</tr>
<tr>
<td>Heathcote et al. (2010)</td>
<td></td>
<td></td>
<td>Level</td>
<td>.019</td>
<td>.085</td>
</tr>
<tr>
<td><strong>Canada</strong></td>
<td>SLID</td>
<td>1993-2005</td>
<td>Difference</td>
<td>.055</td>
<td>.061</td>
</tr>
<tr>
<td>Brzozowski et al. (2010)</td>
<td></td>
<td></td>
<td>Level</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Germany</strong></td>
<td>GSOEP</td>
<td>1984-2004</td>
<td>Difference</td>
<td>.030</td>
<td>.045</td>
</tr>
<tr>
<td>Fuchs-Schündeln et al. (2010)</td>
<td></td>
<td></td>
<td>Level</td>
<td>.010</td>
<td>.075</td>
</tr>
<tr>
<td><strong>Sweden</strong></td>
<td>LINDA</td>
<td>1978-2004</td>
<td>Difference</td>
<td>.040</td>
<td>.010</td>
</tr>
<tr>
<td>Flodén and Domeij (2010)</td>
<td></td>
<td></td>
<td>Level</td>
<td>.006</td>
<td>.061</td>
</tr>
</tbody>
</table>

Note: $\sigma^2_\eta$ and $\sigma^2_\epsilon$ denote permanent & transitory variances respectively

Table 5.1 provides a direct comparison of the average permanent and transitory variances estimated in this paper to those identified by Heathcote et al. (2010) for the U.S.; Brzozowski et al. (2010) for Canada; Fuchs-Schündeln et al. (2010) for Germany and Flodén and Domeij (2010) for Sweden.

All countries in Table 5.1 employ national longitudinal surveys that like the HILDA survey belong to the Cross National Equivalent File (CNEF). Each national study uses the panel dimension of these data sources and the level and difference approach described in detail by Heathcote et al. (2010) to apply a similar version of the permanent-transitory model as defined by equations (5.1)-(5.2).

This exercise for each country delivered the results shown in Table 5.1. These countries estimated moments from across different sample periods, largely determined by data availability. For Canada, Brzozowski et al. (2010) only report the results for their level estimates of wage dynamics graphically. 4

It is evident in Table 5.1, that the ‘inconsistency’ of the difference and level approach holds across countries. Firstly, the variance of the transitory shock is larger across both approaches, for all nations. Secondly, all countries report a larger average variance of the permanent wage innovation using the difference method. Finally, the transitory variance is larger when estimated under level conditions.

4Based on graphical evidence, Brzozowski et al. (2010)’s level estimates of permanent and transitory moments are approximately 0.01 and 0.14, respectively.
Therefore, it appears that the salient features identified in Figure 5.1 are commonalities across all of these countries. This even includes the negative variances observed for permanent wage shocks estimated using the levels approach – [Heathcote et al. (2010)] similarly record negative variances for this component using the same method.

[Heathcote et al. (2010)] use their parameter estimates and the results of the other nations shown in Table 5.1 as evidence in support of the misspecification of this permanent-transitory model. They present this inconsistency as an ‘empirical puzzle’ in the literature. Based on the results shown in Figure 5.1 and Table 5.1, it is clear that this puzzle also holds true for Australian data.

This Australian finding provides important additional evidence from a new data sources and a new country to further validate [Heathcote et al. (2010)]’s argument that the conventional permanent-transitory model is misspecified.

As these parameters represent the key inputs to incomplete-markets models, the potential misspecification of the permanent-transitory model is a serious concern that undermines this framework and its role in reliable policy analysis. For this reason, there is an active area of research dedicated to reconciling the inconsistency between level and growth-rate estimates of the wage process.
6 Life-Cycle Inequality

This section documents how inequality behaves over the life-cycle. This analysis specifically aims to provide the empirical patterns necessary for informing future incomplete-markets models that incorporate an explicit life-cycle structure\footnote{See for example \cite{storesletten2004}; \cite{guvenen2007}; \cite{heathcote2010} and \cite{kaplan2012}.}. The approach used here was first developed by Deaton and Paxson (1994) and the model specification and identifying restrictions employed are based on the empirical strategy presented by Huggett et al. (2011).

The procedure used to examine lifetime changes in the wage distribution begins with the construction of a rotating age panel. This ensures sufficient observations in each age-year category. From there, I calculate the moments of interest for individuals in each age-year bin of the sample. I then use these estimates to construct their corresponding life-cycle profiles. Finally, I repeat this exercise to obtain life-cycle patterns of inequality for my two education subsamples.

The challenge in obtaining a clear view of life-cycle inequality lies in distinguishing actual age effects from cohort or time effects. This is important especially given the rise in cross-sectional inequality in Australia identified in Section 4.1. The analysis will address this issue by presenting age-profiles estimated from both a cohort and time perspective.

6.1 Model Specification and Identification

To characterise the distribution across age, statistics measuring the mean, dispersion and skewness of wages are obtained. Similarly to my cross-sectional analysis in Section 4.1, two measures of dispersion are employed; both the residual variance of log hourly wages and the Gini coefficient. Skewness, defined as the ratio of mean to median wages is also included as an indicator of the portion of wages accruing to top earners.

The following general model is used:

\[
    h_{a,t} = \beta_0 + \beta_a D_a + \beta_t D_t + \beta_c D_c + \epsilon_{a,t} \tag{6.1}
\]

In the equation above, where $\beta_a$, $\beta_c$ and $\beta_t$ represent the respective coefficients on
three full sets of age $D_a$, cohort $D_c$ and year dummies $D_t$. The age-profile for each moment is formed from the estimated vector of age coefficients, $\hat{\beta}_a$.

**Cohort Vs. Time.** The identification strategy of this method relies on restrictions which address the collinearity of the following relationship:

$$Age(a) = Year(t) - Cohort(c)$$

From the equation above it is clear that directly capturing how age impacts upon inequality over the lifetime of an individual is complicated by the relative contributions of time and cohort effects. For instance, what might be identified as an age-driven change in inequality could just as easily be interpreted as being a time or cohort effect.

Cohort effects can be described as conditions specific to an individual’s birth year. They represent historical factors such as the size of a cohort in a given birth year as well as differences in quality and access to education across cohorts. Time effects on the other hand captures changes in the aggregate economic environment across calendar year. This includes business cycle fluctuations, economy-wide productivity shocks, trade liberalisations, demographic shifts and economic growth.

Separating out pure age effects in the presence of these time and cohort effects is not generally possible due to their inter-dependency. To isolate age-inequality effects, I have adopted an identification strategy commonly used by existing empirical studies. This involves taking either a cohort or time approach and imposing normalising restrictions accordingly. As there is a lack of consensus within the literature regarding the most appropriate stance to take, the results here present estimates of age effects using both methods.

The two regressions below depict how this is been achieved:

$$h^c_{a,c,t} = \beta_a D_a + \beta_c D_c + \epsilon_{a,t}$$

$$h^t_{a,c,t} = \beta_a D_a + \beta_t D_t + \eta_{a,t}$$

The first equation represents the cohort view; year effects are restricted to zero and a regression is run on a full-set of age and cohort dummies to control for cohort effects. The second equation takes a time view, normalizes cohort effects to zero and a regression is instead run on a full-set of age and year dummies.

The estimated vector of age coefficients from both approaches are extracted to construct life-cycle profiles that are then reported jointly in Figure 6.1.

---

2See Heathcote et al. (2005) for discussion of the two views.
Rotating Age Panels. An additional consideration here is that HILDA is a relatively small survey, this implies that there could be insufficient observations for every single age in each year to obtain good estimates of the statistics of interest. This issue is addressed in two ways. First, the selection criteria includes only those between the ages of 23 and 60 years and excludes the ages at the extremes of the working life-cycle where there are fewer observations. Second, overlapping rotating age panels are constructed to ensure sufficient cell sizes. These centered 5 year age cells are built by defining an individual’s age as \( 'a' \) if their actual age lies within the range of \( 'a-2' \) and \( 'a+2' \). For instance an individual whose age lies between 23 to 27 years is assigned the starting age of 25. This rotating age panel structure ensures at least 100 observations in each cell in the range 23-27 to 56-60.

6.2 Results

Figure 6.1: Life-Cycle Wage Inequality: Time & Cohort Views

Figure 6.1 presents the life-cycle profiles for the distribution of wages for the full sample. Two sets of estimates are reported in each graph, representing age effects that have been separately conditioned on cohort and year effects. These profiles have been constructed from the series of 35 coefficients on age dummies taken from the two regressions described by Equation (6.3) and (6.4). This exercise is also repeated for both high and low education samples. Life-cycle profiles for these groups shown in Appendix B.1 and results are discussed below.
In all cases in Figure 6.1, it is clear that the cohort and time views produce qualitatively similar results; however, precise magnitudes are sensitive to the stance taken. Controlling for cohort effects results in steeper profiles than controlling for time effects across all statistics. Similar findings have been reported in relevant empirical studies including Heathcote et al. (2005), Brzozowski et al. (2010) and Heathcote et al. (2010).

Heathcote et al. (2005) argue that this can be attributed to the fact that the cohort approach does not control for rising time trends. This might also be the case here for Australian data. The rising time trend observed for the first and second moment of real hourly wages in Figure 3.2.a and Figures 4.1.b-c respectively, may be confounding with age effects and attributing to the steeper cohort profiles observed in Figure 6.1.

Figure 6.1.a reports the life-cycle pattern for the mean of real log hourly wages. The time age-profile of mean hourly wages is much more stable than the cohort profile. Both trajectories are consistent with human capital theory predictions which state that wage is low early in life and increases at a decreasing rate over the working lifetime. Indeed, my results show that wages somewhat stabilize in growth at middle age around 35-40 years. However, the profiles do not exhibit the downward sloping trend at older ages predicted by human capital theory and the traditional hump-shaped pattern is not observed. This can mainly be attributed to the chosen income measure of hourly wages which has already been adjusted for endogenous labour supply. Individuals approaching retirement who spend less time working by reducing their hours should not be expected to also experience any change in their hourly wage rate.

The key finding of this exercise is shown in Figure 6.1.b and Figure 6.1.c. which both confirm that wage inequality increases over the life-cycle. Together, these two figures illustrate that the observed increase in life-cycle inequality is robust to both the variance of log hourly wages and the Gini Coefficient. Figure 6.1.d indicates the skewness of hourly wages is also increasing with age. This result implies that mean hourly wages increases by a greater magnitude than the median over the lifetime, suggesting significantly higher expansion in the upper wage region. Figure 6.1.d therefore provides accompanying evidence for an increasingly uneven distribution of wages with age.

4They base this finding on evidence of life-cycle and cross-sectional patterns of inequality observed in the United States in the 1980s.
Both inequality profiles of real hourly wages depicted by Figure 6.1.b and 6.1.c display some curvature when taken from both a cohort and time view. Inequality rises until around age 43-47 after which wage dispersion does not grow as rapidly. This is in contrast to studies including Brzozowski et al. (2010) and Blundell and Etheridge (2010) that find evidence for a monotonic linear profile of the log variance of wages. However, my results do align with Heathcote et al. (2010) who also report similar curvature in the life-cycle inequality profile of log wages. The finding in Figure 6.1.b and 6.1.c that wage inequality fans out over the life-cycle therefore provides evidence to support the view that highly persistent wage shocks are incurred over the working lifetime.

Appendix A.5 provides a supplementary look at the life-cycle profiles for the two education subsamples. I estimate age profiles for both the mean and variance of log hourly wages. For average log hourly wages, the higher education group is shown to have a steeper profile from ages 23-27 to 33-37 relative to the low education group. This is indicative of highly educated workers realizing their return on investment to education, which is traditionally completed in Australia around the age of 22. The steep profile early in the working life results in higher average hourly wages for the high education group across the remainder of the working lifetime.

Life-cycle inequality for the two education groups is depicted in Appendix A.5. by the variance of log hourly wages. The inequality patterns of the two education groups are somewhat more difficult to interpret. Taken from the cohort view, both education groups report an inequality profile of similar magnitude; however, when taken from the time view, the high education group shows a higher level of inequality over the life-cycle. The shape of the inequality age profiles for the high and low education groups are also somewhat dissimilar. The profile for the high education group appears to be more concave whereas the low education profile assumes a more linear trajectory. For both education groups however is it clear that inequality increase over the life-cycle irrespective of whether one conditions on either cohort or time effects.

In summary, this section characterizes life-cycle patterns of the wage distribution. I find strong and conclusive evidence for increasing wage inequality over the life-cycle. This result is robust to both measure of inequality, identification strategy and education subsample. Such a finding has significant implications for quantitative research interested in consumption, demographic changes, health, savings and portfolio decisions, human capital accumulation and a wealth of other related areas.
7 Simulated Life-Cycle Profiles

This section returns to the ‘empirical puzzle’ discussed previously in Section 5. Here, I evaluate my two sets of diverging parameters that were estimated using the difference and levels approaches described in Section 5.1.

I first simulate two stochastic wage processes using each set of permanent-transitory moments. I estimate life-cycle inequality profiles using this simulated data. These profiles are then used to assess which set of parameters yields the closest replica to the actual age-inequality profile estimated by the procedure outlined in Section 6.1. For robustness, I simulate these age-profiles using two measures of inequality; the variance of log hourly wages and the Gini Coefficient. I also simulate inequality profiles using the variance measure for both high and low education groups.

The unbalanced, dynamic nature of the panel data complicates this exercise. In particular, it means that each individual in the sample may have a different initial or terminal year. It also implies that there can be respondents exiting and re-entering the sample at different times. In order to address these issues and simulate a stochastic wage process for each individual, additional restrictions must be imposed on my original sample.

First, I include only individuals in years for which I have obtained permanent and transitory parameter estimates (2003-2010). I then also exclude individuals who exit and re-enter the survey. The resulting sample represents an alternative to using a balanced panel for this analysis. A balanced panel would be easier in many regards; however, it would reduce the sample size considerably and introduce potential attrition bias. The sampling strategy I employ still allows for the simulation of wage dynamics whilst also retaining as much information from the sample as possible.

The sample selection criteria and the estimation procedure carried out for this simulation are presented in Appendix C.1 and C.2, respectively. In Figure 7.1 and 7.2, I carry out a goodness-of-fit exercise.

---

1 2002 is also included to obtain a value of initial value of permanent income for individuals in 2003. This year is not used for simulation because the difference approach did not estimate moments for 2002.
7.1 Implications of Permanent and Transitory Parameters over the Life-Cycle.

In Figure 7.1, I plot the level and difference simulated profiles together with the actual profile estimated from the data. To construct these profiles I employ the model and identification strategy outlined in Section 6.1 This implies that I must estimate life-cycle inequality profiles that control separately for time and cohort effects. Therefore, Figure 7.1.a and 7.1.b present estimated age effects conditioned on cohort effects and time effects, respectively.

From both Figure 7.1.a and Figure 7.1.b, it is visually evident that the difference parameters generate a profile that is the closest fit to the actual inequality profile estimated from the data for both cohort and time profiles.

Taking the cohort view, Figure 7.1.a reports that both difference and level simulated profiles have underestimated the actual age profile; although the difference profile does so to a less extent. The time approach in Figure 7.1.b produces simulated profiles that are much closer to their empirical counterpart. In this case also, the difference approach provides a more accurate estimate of the observed inequality pattern.

Table 7.1 shows the squared distance of the level and difference profiles from the actual profile at different stages of the life cycle. The Mean Squared Error (MSE) is reported for the two simulated profiles to assess their relative fit to the actual inequality profile.
The MSE estimates in Table 7.1 align with my previous visual observations in Figure 7.1 and confirms that the simulated difference profile is a closer fit to the actual life-cycle inequality profile across both cohort and time approaches. For the cohort (time) view, the difference profile has an MSE estimate of 0.0069 (0.0005) as opposed to the level profile which is further away with an MSE estimate of 0.0209 (0.0042). This suggests that the permanent and transitory parameters estimated using the first-difference approach have greater predictive accuracy.

For robustness, I evaluate life-cycle inequality profiles simulated from the two sets of permanent-transitory parameters using the Gini coefficient as an alternative measure of inequality. Figure 7.2 shows that the simulated and actual life-cycle inequality profiles depict analogous results to those presented in Figure 7.1.

In Figure 7.2.a, the simulated difference profile again appear to provide a closer fit to the actual profile. In Table 7.2 this observation is confirmed, the MSE estimate for the difference profile is found to be 0.0016 whereas the simulated levels profile reports a MSE estimate of 0.0031. In Figure 7.2.b the simulated profiles are much closer in distance to their actual counterpart. This makes it difficult to compare the two simulated profiles graphically; however, Table 7.2 indicates that again the difference approach is a marginally closer fit.

<table>
<thead>
<tr>
<th>Age</th>
<th>Cohort Difference</th>
<th>Cohort Level</th>
<th>Year Difference</th>
<th>Year Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-27</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>33-37</td>
<td>.0030</td>
<td>.0088</td>
<td>.0001</td>
<td>.0055</td>
</tr>
<tr>
<td>43-47</td>
<td>.0076</td>
<td>.0258</td>
<td>.0001</td>
<td>.0085</td>
</tr>
<tr>
<td>58-60</td>
<td>.0068</td>
<td>.0239</td>
<td>.0001</td>
<td>.0065</td>
</tr>
<tr>
<td>MSE</td>
<td>.0069</td>
<td>.0209</td>
<td>.0005</td>
<td>.0042</td>
</tr>
</tbody>
</table>

Table 7.1: MSE estimates: Simulated vs Actual Variance Profiles

Figure 7.2: Actual vs. Simulated Age Profile: Gini Coefficient
Table 7.2: MSE estimates: Simulated vs. Actual Gini Profiles

<table>
<thead>
<tr>
<th>Age</th>
<th>Cohort Difference</th>
<th>Cohort Level</th>
<th>Year Difference</th>
<th>Year Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-27</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>33-37</td>
<td>.0009</td>
<td>.0013</td>
<td>.0001</td>
<td>.0002</td>
</tr>
<tr>
<td>43-47</td>
<td>.0019</td>
<td>.0035</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>58-60</td>
<td>.0031</td>
<td>.0069</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>MSE</td>
<td>.0016</td>
<td>.0031</td>
<td>.0000</td>
<td>.0001</td>
</tr>
</tbody>
</table>

I also carry out this procedure on my high and low education subsamples. Results are shown in Appendix C.3. For the low education group, neither the level or difference approach show a good fit to the actual data. However, for the high education group; the permanent and transitory parameters of the difference approach again produce a closer replica.

It should be noted that the simulation for my two education samples could be limited by a small sample size. The additional restrictions required for simulation (as shown in Appendix C.1) significantly reduces the original sample. Further disaggregation of this smaller sample into high and low education groups may have an impact on results.

In summary, these findings support the view that the difference set of permanent and transitory parameters appear to simulate life-cycle inequality profiles that more closely represents the lifetime patterns observed in actual Australian data. Further research here is required into reconciling the inconsistent estimates of the wage process in growth rates and levels.
8 The Role of Education

Supplementary to my main analysis, this section disaggregates my sample and allows for differences in wage inequality by level of education. I assess between education group inequality over the 2000s and compare results to existing Australian and international evidence. I then decompose within-group inequality by applying the unobservable components model specified in Section 5.1 to obtain education-specific permanent and transitory variance parameters.

8.1 Wage Inequality Trends by Education Group

Figure 8.1: Wage Inequality Trends by Education: 2000-01 to 2010-11

Figure 8.1 reports differences in inequality measures across education subgroups; these results show that inequality is greater for highly educated workers at all points of the wage distribution.

Inspection of Figure 8.1.a and 8.1.b highlights similar trends across both variance and Gini inequality measures. In Figure 8.1.a, the variance of log hourly wages for highly educated workers rose by 1.8 log points from 0.186 to 0.204 between 2000-01 and 2010-11. Over the same time period, the variance measure of the low education group increased from 0.135 to 0.159 log points, representing an overall gain of 2.4

1Education groups are defined by Section 3.3.2.
log points; however, it should be noted that most of this increase was experienced for the low education group in the last two years of the sample.

Figure 8.1.b shows that the Gini coefficient follows a similar evolution for both education groups. The Gini increased by 1.6 points from 0.239 to 0.255 for highly educated workers over the sample period. In contrast, a net increase of 2.1 points, from 0.209 to 0.230 is reported for the low education group, again with the largest growth occurring between 2008-09 to 2010-11.

Figure 8.1.c and Figure 8.1.d compares the inter-decile ratios of the two education groups. Together, these figures communicate that there are larger observed differences between the 50-10 ratio of the two groups than there are for the 90-50 ratio. This result indicates that there is a greater distance between the bottom decile and the median for the high education group compared to the low education group; whereas at the top of the wage distribution, relative to the median, the two education groups have a similar spread.

Overall, the gap between high and low education groups remains relatively constant on average for all inequality measures over the sample period. This finding aligns with the stable behaviour of the education-wage premium in Figure 4.2. These results represent quite a different pattern to what has been observed in other countries, particularly in the United States.

In the U.S., a rising college wage premium driven by the increasing demand for high-skilled, high ability workers has been the leading explanation for observed inequality trends since the 1980s. An extensive body of research has traditionally attributed this increase in demand to skill-biased technological change (SBTC). Earlier models of this hypothesis argued that SBTC was created by the automation/computerisation of the workplace. More recent models have incorporated the routine nature of jobs and international trade into their framework in an attempt to explain differences in observed inequality patterns across countries.

Interestingly, to date there has been limited empirical evidence in support of the SBTC hypothesis in Australia. In stark contrast to the United States, Australian studies including Borland (1999) and Coelli and Wilkins (2009) have shown that the education earnings premium in Australia has remained stable from the early 1980s up until 2004. The results in Figure 8.1 are therefore consistent with earlier Australian studies and provide empirical evidence to suggest that rising wage

\[\text{See Katz and Murphy (1992) and Acemoglu (2002).}\]

\[\text{Both papers present varying arguments for this disparity. Borland (1999) argues that a rapid increase in supply of highly educated workers explains the stability of the college wage premium, whereas Coelli and Wilkins (2009) put forward that changes in credentials of certain occupations could also account for there being no change in the education differential up until 2004.}\]
inequality in Australia has predominantly occurred within-education groups.

8.2 Within-Group Inequality by Education Group

This section of the analysis explores within-group wage inequality across education groups by assessing whether the stochastic wage process differs by education. The underlying motivation of this exercise is to obtain education-specific permanent and transitory parameters and to assess whether the ‘empirical puzzle’ identified in Section 5 holds across both high and low education levels. The permanent-transitory model from Section 5.1 is applied to both subsamples, again using both the ‘level’ and ‘difference’ estimation procedures.

Figure 8.2 illustrates the variance of the permanent and transitory shocks to log real wages against time for both high and low education groups. Figure 8.2.a and Figure 8.2.b depict similar magnitudes over time for the variance of the permanent shock for both high and low education groups; although the low education group fluctuates more across the sample period. For both education groups, the level approach reports a negative permanent variance in some years, which again, similar to the full sample estimates, is a signal of misspecification.

Figure 8.2.c and Figure 8.2.d represent the variance of the transitory shock for both high and low education groups, respectively. In this instance, the variance of the high education transitory shock fluctuates more than that of the low education group.
To further investigate these results, Table 8.1 reports the average variances of both the permanent and transitory shocks taken for both education groups across the sample period (2001-02 to 2009-10).

Table 8.1: Estimates of Wage Dynamics by Education Group

<table>
<thead>
<tr>
<th>Approach</th>
<th>Sample</th>
<th>$\sigma_n^2$</th>
<th>$\sigma_\epsilon^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels</td>
<td>High</td>
<td>.0002</td>
<td>.0329</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>.0024</td>
<td>.0308</td>
</tr>
<tr>
<td>Difference</td>
<td>High</td>
<td>.0230</td>
<td>.0164</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>.0242</td>
<td>.0119</td>
</tr>
</tbody>
</table>

Note: $\sigma_n^2$ and $\sigma_\epsilon^2$ denote permanent & transitory variances respectively

Comparison across education groups in Table 8.1 illustrates that the average variance of the transitory shock is larger for both high and low education groups when estimated using the levels approach. In contrast, the average variance of the permanent shock is greater when estimated using first-difference conditions. As before, it is clear that the two approaches cannot simultaneously produce estimates that are consistent with one another.

From Figure 8.2 and Table 8.1, it is apparent that the ‘empirical puzzle’ discussed in Section 5.2 is robust across education groups; the level and difference approach produce a similar pattern of diverging estimates for the two education groups as to what was observed for the full sample.

---

4Similar to results in Section 5.2, the difference approach does not estimate moments for 2001-02.
9 Summary and Conclusion

In this thesis, I document key wage inequality patterns and wage dynamics useful for informing future quantitative macroeconomic research in Australia.

I analyse cross-sectional and life-cycle trends in wage inequality and find robust and conclusive evidence that wage inequality is increasing across both dimensions. I also employ the panel structure of the HILDA survey using ‘level’ and ‘difference’ estimation methods to estimate a conventional permanent-transitory model and decompose residual wage inequality. I find that these two approaches deliver substantially different estimates of the variance of permanent and transitory shocks to wages. The disparity in my findings aligns with existing studies including Heathcote et al. (2010) who have identified this inconsistency to be an empirical puzzle within the literature.

I further investigate this inconsistency by simulating life-cycle inequality profiles using my diverging sets of permanent and transitory parameters. I find that the life-cycle inequality profiles generated using the difference approach provide a closer comparison to lifetime patterns observed in the actual data. Additionally, I examine both between and within education group inequality. I present evidence for a stable education wage premium and provide education-specific permanent and transitory parameters.

This research represents an important precursor to investigating the welfare implications of rising wage inequality in Australia. Crucial steps have been taken to enable the analysis of a wide range of policy issues within an incomplete-markets framework in the future. However, for this to be possible, some important areas of inquiry still remain.

First, the results of my research have focused solely on providing wage inputs to an incomplete-market model. Further empirical insight into how these wage inputs transmit through insurance and redistributive mechanisms is required. This can be achieved by compiling a more extensive collection of empirical facts documenting trends in the distribution of hours, earnings, disposable income, consumption and wealth within Australia at both individual and household levels.

Limitations imposed by the data restricted my analysis of these areas. The periodical collection of expenditure and wealth data and the fact that the survey is an individual level rather than a household level panel were constraining factors of this
research. The former issue prevented the estimation of consumption inequality which would have been particularly useful to examine in conjunction with my estimates of permanent and transitory wage inequality. The second issue did not allow for the analysis of household-level earning dynamics and intra-household insurance mechanisms.

Future research may therefore benefit from integrating consumption information from other data sources (e.g. ABS Household Expenditure Survey) and also from investigating potential ways to construct a household panel using cross-wave identifiers provided by HILDA.

A more pressing research agenda lies in addressing the empirical puzzle confirmed in this paper and previously identified by other researchers. The function of the permanent-transitory model in supplying inputs for the incomplete markets framework means that this model's potential misspecification is a serious concern for quantitative macroeconomists. For this reason, researchers are currently trying to reconcile the level and difference approaches by allowing for more complex error structures in the wage process. In Australia, this area of research is restricted by the relatively short time dimension of HILDA.

That said, this Australian finding provides important insights and new evidence that can be used to inform ongoing research in this area. The importance of providing Australian evidence to this field of the literature is further emphasized by the future potential of heterogeneous-agent incomplete-market models to enable quantitative research into the welfare implications of rising wage inequality in Australia.
A Data

A.1 Sampling Strategy: Male Full-Time Sample

Table A.1: Sampling Strategy: Full-time Male Sample

<table>
<thead>
<tr>
<th>Description</th>
<th>Dropped</th>
<th>Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial unbalanced sample</td>
<td>216 368</td>
<td></td>
</tr>
<tr>
<td>Missing wave information</td>
<td>15 026</td>
<td>201 342</td>
</tr>
<tr>
<td>Respondents only</td>
<td>53 519</td>
<td>147 823</td>
</tr>
<tr>
<td>Age restrictions</td>
<td>46 773</td>
<td>101 050</td>
</tr>
<tr>
<td>Employed-full time</td>
<td>47 444</td>
<td>53 606</td>
</tr>
<tr>
<td>Full-time student</td>
<td>339</td>
<td>53 267</td>
</tr>
<tr>
<td>Not self employed</td>
<td>9 300</td>
<td>43 697</td>
</tr>
<tr>
<td>Male only</td>
<td>16 318</td>
<td>27 649</td>
</tr>
<tr>
<td>Miscoded or zero weekly income</td>
<td>131</td>
<td>27 518</td>
</tr>
<tr>
<td>Miscoded or zero annual income</td>
<td>323</td>
<td>27 195</td>
</tr>
<tr>
<td>Drop if half the minimum wage</td>
<td>151</td>
<td>27 044</td>
</tr>
<tr>
<td>Missing demographic information</td>
<td>66</td>
<td>26 978</td>
</tr>
<tr>
<td>At least two consecutive waves</td>
<td>3 717</td>
<td>23 261</td>
</tr>
</tbody>
</table>

A.2 Education Groups: Trends in Average Weekly Earnings

![Real Average Earnings by Education Group]

Real Average Earnings
By Education Group

- Post-Graduate
- Graduate Diploma
- Undergraduate
- Tafe Certificate/Diploma
- Higher-School Certificate
- Year 11 or Below
A.3 Description of Key Demographic Variables

Table A.2: Description of Key Demographic Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Education</td>
<td>Dummy variable equal to 1 if Postgraduate- Masters or Doctorate, Undergraduate degree, Graduate diploma, 0 if Certificate or Diploma, Higher School Certificate or Year 11 or below.</td>
</tr>
<tr>
<td>Age</td>
<td>Age as at last birthday, on 30th June preceding interview.</td>
</tr>
<tr>
<td>Male</td>
<td>Dummy variable equal to 1 if male, 0 female.</td>
</tr>
<tr>
<td>Married</td>
<td>Dummy variable equal to 1 if legally married or defacto; 0 otherwise.</td>
</tr>
<tr>
<td>Full-Time Employee</td>
<td>Dummy variable equal to 1 if full-time employee, 0 otherwise.</td>
</tr>
<tr>
<td>Employment Status</td>
<td>Categorical variable equal to: 1 if full-time, 2 if part-time, 3 if NILF.</td>
</tr>
<tr>
<td>Self-Employed</td>
<td>Dummy variable equal to 1 if self employed, 0 otherwise.</td>
</tr>
<tr>
<td>Full-time Student</td>
<td>Dummy variable equal to one if enrolled in full-time education.</td>
</tr>
<tr>
<td>Migrant</td>
<td>Categorical variable equal to: 1 if English speaking migrant 2 if Non-English speaking migrant, 3 if Australian.</td>
</tr>
<tr>
<td>Union</td>
<td>Dummy variable equal to 1 if belongs to union or employee association.</td>
</tr>
<tr>
<td>Health</td>
<td>Dummy variable equal to 1 if long term health condition.</td>
</tr>
<tr>
<td>State</td>
<td>Categorical variable indicating Australian state or territory of residence.</td>
</tr>
<tr>
<td>Indigenous</td>
<td>Dummy variable equal to 1 if Aboriginal or Torres Strait Islander.</td>
</tr>
</tbody>
</table>

A.4 Descriptive Statistics for Education Subsamples

Table A.3: Descriptive Statistics by Education Group

<table>
<thead>
<tr>
<th></th>
<th>High Education</th>
<th>Low Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age(years)</td>
<td>39.2</td>
<td>40.9</td>
</tr>
<tr>
<td>% Married</td>
<td>69.2</td>
<td>67.8</td>
</tr>
<tr>
<td>% Union</td>
<td>33.1</td>
<td>32.5</td>
</tr>
<tr>
<td>% English Speaking Migrant</td>
<td>18.2</td>
<td>11.9</td>
</tr>
<tr>
<td>% Non-English Migrant</td>
<td>13.0</td>
<td>13.1</td>
</tr>
<tr>
<td>% Australian Born</td>
<td>68.9</td>
<td>74.9</td>
</tr>
<tr>
<td>% Indigenous</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Avg. Hours Worked</td>
<td>2 378</td>
<td>2 285</td>
</tr>
<tr>
<td>Avg. Annual Earnings $</td>
<td>93 982</td>
<td>109 826</td>
</tr>
<tr>
<td>No. Obs.</td>
<td>517</td>
<td>671</td>
</tr>
</tbody>
</table>
B  Life-Cycle Inequality

B.1 Life-Cycle Profiles by Education Group

Figure B.1: Life-Cycle Wage Inequality by Education: Time & Cohort Views
C Simulated Life-Cycle Inequality Profiles

C.1 Sampling Strategy: Simulated Life-Cycle Inequality Profiles.

Table C.1: Sampling Strategy: Simulated Life-Cycle Profiles

<table>
<thead>
<tr>
<th>Description</th>
<th>Dropped</th>
<th>Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial unbalanced sample</td>
<td>216 368</td>
<td>1 342 026</td>
</tr>
<tr>
<td>Missing wave information</td>
<td>15 026</td>
<td>201 342</td>
</tr>
<tr>
<td>Respondents only</td>
<td>53 519</td>
<td>147 823</td>
</tr>
<tr>
<td>Age restrictions</td>
<td>46 773</td>
<td>101 050</td>
</tr>
<tr>
<td>Employed-full time</td>
<td>47 444</td>
<td>53 606</td>
</tr>
<tr>
<td>Full-time student</td>
<td>339</td>
<td>53 267</td>
</tr>
<tr>
<td>Not self employed</td>
<td>9 300</td>
<td>43 697</td>
</tr>
<tr>
<td>Male only</td>
<td>16 318</td>
<td>27 649</td>
</tr>
<tr>
<td>Miscoded or zero weekly income</td>
<td>131</td>
<td>27 518</td>
</tr>
<tr>
<td>Miscoded or zero annual income</td>
<td>323</td>
<td>27 195</td>
</tr>
<tr>
<td>Drop if half the minimum wage</td>
<td>151</td>
<td>27 044</td>
</tr>
<tr>
<td>Missing demographic information</td>
<td>66</td>
<td>26 978</td>
</tr>
<tr>
<td>Keep Waves 2-10 only (a)</td>
<td>7 993</td>
<td>18 985</td>
</tr>
<tr>
<td>Exclude those who exit and re-enter (b)</td>
<td>2 507</td>
<td>16 478</td>
</tr>
<tr>
<td>At least three consecutive waves (c)</td>
<td>1 065</td>
<td>15 413</td>
</tr>
</tbody>
</table>

C.2 Summary: Steps for Simulating Life-Cycle Profiles

1. Construct sample as outlined in Appendix A.2.
   (a) Exclude 2001 and 2011 from the sample as the P-T model does not estimate moments for these years. Although only moments from 2003 to 2010 are used for simulation.
   (b) Include only individuals who are present for at least three consecutive waves.
   (c) Exclude individuals who exit and re-enter the sample.

---

1 2002 is still included to obtain initial values for 2003, see Step 3 and 4. (the Level estimates for 2002 are not used for simulation as there are no corresponding Difference Estimates)
2. Create an identifier for each individual’s initial \((t=a)\) and terminal year \((t=b)\).

3. Run regression (5.1) by year and capture the residual, \(u_{i,a}\), for every individual.

4. Assign the residual, \(u_{i,a}\), as the initial value of the permanent component, \(z_{i,a+1}\) to every individual who entered the sample in year \(t=a\).

5. For the Difference Approach:
   
   (a) Generate the random normal variables: \(\epsilon_{i,t} \sim (0, \hat{\sigma}_{\epsilon,t})\) and \(\eta_{i,t} \sim (0, \hat{\sigma}_{\eta,t})\), where \(\hat{\sigma}_{\epsilon,t}\) and \(\hat{\sigma}_{\eta,t}\) are the estimated permanent and transitory variances obtained by the difference approach described by equations (5.5)-(5.10).

   Take a random draw from each time-varying distribution for every individual across every year they appear in the sample. This will generate \(\hat{\eta}_{i,t}\) and \(\hat{\epsilon}_{i,t}\).

   (b) By individual, construct the error term, \(\hat{u}_{i,t}\), defined by equations (5.2) and (5.3), using \(\hat{\eta}_{i,t}\) and \(\hat{\epsilon}_{i,t}\) from year \(a+1\) to year \(b\).

   (c) Estimate the coefficient parameters of regression (5.1) for each year. Construct the wage, \(\hat{y}_{i,t}\), process using the estimated coefficients, the observed values of the control variables and \(\hat{u}_{i,t}\) for each year.

6. Repeat Step 5 for the Levels Approach.

   (a) Interchange equations (5.5)-(5.10) with equations (5.11)-(5.14).

7. Simulate Step 5 and 6 with 100 repetitions.

8. Construct life-cycle profiles of the variance of log hourly wages using simulated data from both approaches and applying the model specified by Section 6.1.

9. Compute the average difference and level simulated inequality profile and compare to the observed inequality age profile estimated using the actual data.

10. Repeat Step 8 and 9 using Gini coefficient as measure of inequality.

11. Repeat Step 1 to Step 9 for high and low education subsamples.

12. Assess the goodness of fit.

---

2 Values of \(\hat{\sigma}_{\eta,t}\) in the level approach that were reported to be negative are set to zero.

3 Time \(t=a\) is not included as the starting year is used to obtain \(z_{i,a+1}\).

4 100 profiles will be generated.
C.3 Simulated vs. Actual Life-Cycle Inequality Profiles by Education Group

Figure C.1: Actual vs. Simulated Life-Cycle Inequality Profiles by Education
Bibliography


55


