Essays in Applied Microeconometrics

Aganitpol Sivakul

A thesis submitted for the degree of Doctor of Philosophy in Economics

St Peter's College

University of Oxford

Trinity Term 2015

Essays in Applied Microeconometrics

Aganitpol Sivakul, St Peter's College

A thesis submitted for the degree of Doctor of Philosophy in Economics,

University of Oxford, Trinity Term 2015

Abstract

This thesis is a collection of three independent essays that applies microeconometrics techniques to empirically study topics in development and labour economics.

The first chapter uses evidence from a natural experiment in Bangladesh, where households were treated to different types of transfer, food grains and cash, at different periods in time, to test the effect of these transfers on household consumption behaviour. Using the fixed effect instrumental variable model, the estimation results show that though in-kind transfers did cause households to consume more grain than they would have chosen under equal-value cash transfers, the impact on calorie consumption and children health status is minimal. Households that received cash were able to reallocate their funds more effectively, and chose to spend their extra income on clothing and children's non-food consumption, while at the same time spending no more on vices.

The second chapter investigates the dynamics of living standards in Thailand. Income and earnings processes are first modelled after the statistical Galton-Markov process before being extended to follow a more structural permanent earnings model. Empirical estimations of income and earnings persistence in Thailand employ both constructed pseudo-panel data from Thailand's Labour Force Surveys and the Townsend Thai panel data. Galton-Markov estimates found conditional persistence to be low in Thailand. However, quantile regression estimates find that persistence is low at the bottom of the distribution but high at the top, indicating a divergence in earnings as time passes. A study of the covariance structure of earnings finds that total variation in the earnings process is predominantly driven by moderately persistent transitory components following the AR(1) process.

The third chapter attempts to empirically fit the power-law distribution and study the dynamics of inequality, especially at the upper end, of the income and consumption distribution in Thailand. We find that using the popular but incorrect method based on the linear regression approach will lead to researchers drawing a wrong conclusion. Regression estimates of the power-law exponent, α, provide strong evidence of power-law fit in Thailand. However, from the implementation of the superior Clauset et al. method, the evidence in support of the power-law fit is much weaker. Estimates of α for both income and consumption suggest that there is low inequality at the top in Thailand but further inspection finds that there is a high level of persistent between-group inequality between the top and bottom ends of the distribution. In addition, following Battistin et al. (2009), we find weak support for Gibrat's law of proportional random growth as the income-generating process in Thailand.

Word Count and Format

The computerised word count is 88,133. Taking page 6 as a representative page, the thesis counts 369 words per page. Multiplying this number by **270** pages (excluding bibliography) results in an overall word count of **99,630**.

The thesis consists of five chapters, including the introduction and conclusion chapters. A typical chapter contains a number of sections. Some chapters contain appendices.

Declaration

Chapter 2 of the thesis builds on my MPhil thesis submitted in Trinity Term 2011 as part of the requirements for the degree of MPhil in Economics, University of Oxford. The title of the MPhil thesis was "An Essay in Empirical Development Economics: Cash Versus In-Kind Transfers, The Case of Bangladesh".

Acknowledgements

First and foremost, I would like to thank my doctoral supervisor, Professor Ian Crawford for his continuous support, insight and valuable advice. I have learnt tremendously under his guidance and am infinitely grateful for all his assistance in the completion of this thesis.

I would also like to thank Professor Albert Park for his supervision of my master thesis, which has been extended and included as the second chapter in this doctoral thesis. My gratitude also goes to Dr. Bob Baulch at CPRC and Wahid Quabili at IFPRI for their help with the Bangladesh data set used in this chapter.

I am also grateful to friends in and outside of the Department of Economics for their help and support throughout the endeavour. La-Bhus Jirasavetakul, Jaree Pinthong and Nuwat Nookhwun deserve special mention. Additionally, I extend my thanks to participants of the Gorman Workshop, the Applied Microeconomics Workshop, the Rimini Conference in Economics and Finance, and the Hong Kong Economic Association Conference for their helpful comments on the thesis chapters.

Last but certainly not least, I wish to thank and dedicate this thesis to my parents. Without their constant combination of admonishment and encouragement, the completion of this thesis would not have been possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis is a collection of three independent essays that applies microeconometrics techniques to empirically study topics in development and labour economics. Chapter 2 addresses a question in the empirical development economics literature concerning the evaluation of transfer programmes. Specifically, the chapter seeks to determine whether the nature of the transfers, in-kind vs. cash, matters for the effectiveness of the programme in influencing households' behaviour, and which is superior. Chapter 3 and 4 address issues in the empirical labour economics literature using Thailand as a case study. Chapter 3 examines the dynamics of living standards in Thailand, focusing on income persistence and mobility, while Chapter 4 seeks to study the behaviour of the income and consumption distribution and the dynamics of inequality in Thailand.

Chapter 2, titled "Cash Versus In-Kind Transfers, The Case of Bangladesh", empirically analyses the effect of in-kind versus cash transfers on household consumption behaviour. A significant number of redistributive welfare programmes around the world are often implemented in-kind. Large redistribution programmes that have been heavily studied, such as the PROGESSA programme in Mexico or Bolsa Familia in Brazil, are in-kind transfer programmes. This common adoption is a puzzling fact given that providing benefits in-kind rather than in cash not only lowers recipients' utility but is often subject to higher administration costs. While there are many possible explanations of why transfers more often occur in-kind rather than in cash, one of the main justifications used to provide the motivation behind most in-kind programmes is that of "paternalism".

Paternalistic in-kind transfers intentionally cause a distortion of recipients' behaviour in

the hope of generating a positive social welfare gain as a result. The use of paternalistic in-kind transfers can only be justified when there exist negative externalities to social welfare in terms of the consumption habits of certain groups of the population that were eventually targeted or when the notion of consumer rationality is no longer respected. If this were truly the case, it is theoretically possible for a distorting in-kind transfer to improve social welfare. This chapter attempts to provide additional empirical evidence contributing to the debate between in-kind versus cash transfers. More specifically, it provides a test of the prediction of the theory of consumer demand under distorting in-kind transfers. The empirical exercises also seek to address whether such distorting paternalistic acts are really necessary given the cost involved.

Using evidence from a natural experiment in Bangladesh, where a set of households were treated to different types of transfer, food grains and cash, at different periods in time, we test whether this form of paternalism is necessary. A fixed effect instrumental variable model is used in the estimation and the comparison of households' behaviour under each type of transfer. This estimation strategy fixes the endogeneity in treatment variables as a result of failure to account for time-invariant household-specific effect and selection into treatment in previous studies of the programme (Ahmed, 2005).

Chapter 3, titled "Dynamics of Living Standards in Thailand", investigates income persistence and mobility in Thailand. Income persistence and mobility, which pertain to the dynamics of the income distribution and how it is changing over time, are crucial concepts. The degree of persistence in income is often seen as a measure of the equality of opportunity, and of the flexibility and freedom of movement, in the labour market (Atkinson et al., 1992). Individuals in a society where incomes are less persistent would enjoy a greater incentive to exert effort and climb up the income distribution than individuals in a society where income persistence is high.

Employing both constructed pseudo-panel data from repeated cross-sections of Thailand's Labour Force Surveys (LFS) (1985–2011) and the Townsend Thai panel data (1997–2010), income and earnings processes are first modelled after the statistical Galton-Markov AR(1) process. Galton-Markov estimates obtained from the Townsend panel data using the ordinary least squares technique are biased and likely to understate the degree of persistence as a result of measurement error in income, non-random attrition and non-representativeness

of the sample. Dynamic pseudo-panel methods are applied to pseudo-panel data constructed from nationally representative LFS to obtain consistent estimates of unconditional and conditional persistence in the presence of measurement error and sample attrition. Using the quantile regression technique, we extend the analysis to other parts of the income distribution to investigate whether households at different parts of the distribution experience different degrees of income persistence.

In addition, moving on from the statistical Galton-Markov framework to a more structural permanent earnings model set-up, we seek to model the covariance structure of the earnings process by decomposing earnings into permanent and transitory components and studying their evolution over the life cycle and over time.

Chapter 4, titled "Power Law in the Income Distribution: Evidence from Thailand", attempts to empirically fit the power-law distribution and study the dynamics of inequality, especially at the upper end, of the income and consumption distribution in Thailand. Discovery of the power-law distribution help provides evidence that the concentration of income and consumption at the top is at a higher level than predicted by the normal or lognormal framework.

Power-law distributions are fitted to income and consumption data from Thailand's Household Socio-Economic Survey (SES, 1986 - 2006). In this chapter, we illustrate the shortcomings of the popular but incorrect method of fitting the power law based on the linear regression approach, which will lead to researchers drawing a wrong conclusion. Instead, we advocate for the implementation of a superior method proposed by Clauset et al. (2009) based on the maximum likelihood estimation of the power-law exponent, α, and the non-parametric method to estimate x_{min}, the minimum cut-off value above which the power-law distribution applies. Clauset et al. (2009) also provide a more credible method for testing the goodness of the power law fit both on an absolute basis and relative to other competing distribution. Power-law estimates obtained are then studied to shed light on the dynamics of inequality in Thailand. In addition, we follow the work of Battistin et al. (2009) to further investigate whether Gibrat's law of proportional random growth is a valid underlying income-generating process in Thailand.

Chapter 2

Cash Versus In-Kind Transfers, The Case of Bangladesh

2.1 Introduction

In virtually all countries, developed and developing, a significant amount of redistribution takes place in-kind (Currie and Gahvari, 2008). In-kind transfers are often designed to change consumption behaviour and promote consumption of the good being transferred. Nonetheless, if there were zero transaction costs, there would be no difference between in-kind and cash transfers if the good could be marketed. Also, if the household already consumes a large amount of the good, the transfer may be infra-marginal and, hence, treated like cash. However, where transaction costs are non-negligible and the household originally consumed little of the in-kind good (making the transfer extra-marginal), we might expect greater consumption of the transferred good.

It is well established in economic theory that a cash transfer is weakly superior in terms of the recipient's utility than an equal-value in-kind transfer, as an in-kind transfer may act as a constraint on the consumption choice of the recipient. It is therefore, at first glance, puzzling that many redistribution programmes give out benefits in-kind as opposed to in cash. Trying to understand why governments conduct their welfare programmes in-kind has been a much-researched question in Economics.

Many theoretical explanations have been put forward to explain this phenomenon. These include one based on imperfect information and targeting of recipients of the welfare pro-

gramme. In the literature, it is argued that when the government has imperfect information on individuals' need of the transfers, provision of transfers in-kind, where the type of goods given can be manipulated, induces more "self-targeting" among recipients than cash. Recipients self-select themselves in to and out of welfare programmes and allows the benefits to go to those who need them the most, resulting in a more efficient transfer system (Nichols and Zeckhauser, 1982). Conversely, when transfers are given out in cash, all individuals will have an incentive to participate regardless of their needs.

New political economy considerations must also be borne in mind. Bearse et al. (2000) argue that in-kind transfers are of very poor quality in poorer countries. As a result, the rich will not participate in in-kind transfer schemes, making the programmes more redistributive, and encouraging the median voters to allocate a larger share of government expenditures to in-kind transfers. Political arguments are also often used to justify the appeal of universal in-kind transfer programmes.

One explanation that has received a lot of attention in the theoretical literature argues that in-kind programmes help improve the efficiency of the tax system. The idea behind this is that provision of in-kind goods complementary to labour can help reduce distortions in the labour supply caused by the tax system. While taxes such as an income tax may have a negative impact on labour supply, redistributing the tax budget through in-kind transfer programmes that are designed to have a positive impact on labour supply may help reduce this distortion. Infant and school-based feeding programmes that improve children's health or in-kind programmes that increase school participation will possibly have a positive impact on future productivity and labour market participation. A pure unconditional cash transfer cannot be designed to promote these labour market outcomes.

Other explanations include pecuniary effects (Coate et al., 1994), where provision of transfers in the form of in-kind goods has a direct effect on local supply. By increasing the local supply of these goods, in-kind transfers induce a general equilibrium effect, such as a fall in local prices of the in-kind goods and complementary products, which can increase recipients' utility more than a cash transfer.

The Samaritan's dilemma occurs when government faces a choice between a commitment not to bail out today's recipient in the future and letting the recipients' current decisions be distorted in anticipation of future transfers. One context where this applies is in human

capital accumulation, where the poor have less incentive to invest in activities that reduce the chances of being poor in the future as they bear high present costs and, when they are successful, run the risk of having the benefits taken away from them (Bruce and Waldman, 1991). Provision of transfers in-kind through a job-training programme may therefore be more effective in dealing with these dilemmas than cash programmes such as unemployment benefits[1].

By far the most common explanation provided for in-kind transfer is "paternalism". Provision of paternalistic in-kind transfers can be justified mainly by two reasons. The first justification is based on externalities, where differences exist between social and individual welfare functions within the society or, in an intra-household setting, when members of households make decisions that are not in the interests of other household members.

This externality justification for paternalism takes many forms in the theoretical literature. In one approach, the quantity of some goods directly enters the society's social welfare function (Pazner, 1972), making the social welfare function become non-individualistic. Another approach involves using interdependent preferences amongst donors and recipients. Under this interpretation, interdependent preferences give rise to a consumption externality that may justify public provision and the use of in-kind transfers (Daly and Giertz, 1972; Garfinkel, 1973).

The second justification for paternalism abandons the assumption of complete rationality among recipients. In this case, where the recipient is rationally bounded, it is argued that an act of "libertarian paternalism" by a central planner that alters individual behaviour could result in a better outcome (Thaler and Sunstein, 2003). In either case, social welfare could be improved by selecting paternalistic policies with the goal of influencing the choices of affected parties in a way that makes these parties better off. Many development projects, including the recently popularised conditional cash transfers, often have an element of paternalism in them. All of these justifications for in-kind transfer require that the in-kind transfers actually increase consumption of the transferred good more than cash transfers.

However, while the motivations for in-kind transfers are well established in the theoretical literature, there is very little empirical evidence on how redistribution programmes alter consumption choice. Studies conducted on the United States Food Stamp Programme showed

[1] A recent paper by Currie and Gahvari (2008) provides a detail account of all these motivations for in-kind transfer stated here.

that food vouchers are infra-marginal for most recipients and thus are treated like cash. In terms of those recipients whose consumption is distorted, Whitmore (2002) found that food stamp recipients have access to a well developed resale market in food stamps and that over-provided stamps that are not sold tend to induce consumption of some non-nutritious foods, such as soft drinks. This seems to suggest the failure of the programme to induce the intended constraint on the recipient's behaviour. The developed country context, however, is very different from the one studied here, where the focus is on transfer programmes in low-income countries.

There is a small, recent body of evidence that directly compares differences in outcomes under cash and in-kind transfers in low-income countries. Cunha (2010) compares the effect on household consumption and health outcomes between households in villages that were randomly assigned either equal-valued cash or in-kind transfers or no transfer at all in the Programa de Apoyo Aimentario (PAL) in rural Mexico. Using the difference-in-difference method to control for baseline differences of households, the author finds no evidence that the in-kind food transfer induced more food to be consumed by beneficiaries than did an equal-valued cash transfer. While the in-kind food transfers, which were made up of more than 10 food items, did indeed alter the types of food consumed in the beneficiary households, the households that received cash were found to have consumed equally nutritious foods and the extra cash was not spent on vices, such as alcohol and tobacco. There were also few differences in the health outcomes between children of households under the two types of transfer. As a result, Cunha found little evidence in support of the necessity of paternalism.

Jayachandran et al. (2010) provide evidence of the existence of a pecuniary effect. The authors found that the price decline in in-kind villages increases the programme's net transfer by 12 percent for the recipients of food transfers, while the price increase in cash villages dissipates 11 percent of the transfer. This provides some evidence of the impact of the price effect, which in isolated market could be large in magnitude and have a substantial effect on the beneficiaries of the transfer programme. Skoufias and Gonzalez-Cossio (2008) also found that, although both in-kind and cash transfers did increase consumption, there were no significant differences between the two and there was no evidence that the programme, irrespective of the type of the transfer, affected overall labour participation. That said, it did cause a significant switch from the agricultural to non-agricultural sectors, perhaps as a

result of the programme freeing households from their credit constraints. Both studies were on the same PAL programme in Mexico and both used the difference-in-difference method to identify the programme effects.

While this evidence from a randomly controlled trial provides strong evidence on the difference - or lack thereof - between in-kind and cash transfers from many perspectives, this chapter makes a new contribution to this literature using a different approach. In a random experiment, different households drawn randomly from a population are given different types of transfers and their outcomes are then compared to each other. The comparison of outcomes between different households is made possible by the random draw, which ensures that treatment is independent of potential outcome. While the data set used in this chapter is non-experimental, it utilises a natural experiment in which households in the survey were treated with different types of transfers, in-kind and cash, at two different periods in time. This makes it theoretically possible to identify the difference in outcomes within the same households who were given an in-kind transfer in one year and a cash transfer in the others.

In addition, the setting in Bangladesh, which is one of the poorest countries in the world, is different from richer developing countries in Central or Latin America such as Mexico. Households in Bangladesh are likely to have poorer access to markets for resale where they can offload the over-provided food grains. The Bangladesh programme also concentrated the transfer in one basic food item (wheat or rice) rather than offering a bundle of goods, perhaps increasing the likelihood that in-kind transfer was extra-marginal. Given the paucity of existing rigorous studies on this issue, this chapter provides valuable empirical evidence on whether the findings from Mexico hold true in other developing country contexts.

The data for this study is a longitudinal data set collected by the International Food Policy Research Institute (IFPRI) in Bangladesh. The panel data set covers the period from 2000 to 2006, during which Bangladesh implemented two conditional transfer programmes where one was a continuation of the other. The first, the Food for Education programme (FFE), started in 1993, provided an in-kind food grain subsidy to eligible households conditional on their children being enrolled in primary school, and was still active in 2000. By the second round of the survey in 2003, the FFE programme had been replaced by the Primary Education Stipends programme (PES), which was a continuation of the FFE programme with the only difference being that transfers were now given out in cash.

A few studies have compared the impact of food and cash transfer programmes in Bangladesh. A study by Del Ninno and Dorosh (2003), using data from various food grain distribution and cash transfer programmes in Bangladesh, finds that the marginal propensity to consume out of wheat transfer in-kind is significantly higher than out of cash transfers. A previous study of the FFE programme using the first year of the data examined in this chapter by Ahmed and del Ninno (2002)[2] finds the FFE programme had a significant effect on increasing calorie consumption, although the study found no effect on the health status of beneficiaries. All of these studies, however, used only non-experimental cross-sectional data, which likely leads to bias due to the endogeneity of treatment associated with unobserved heterogeneity across households.

Ahmed (2005) compared in-kind and cash transfers using the first two rounds of same IFPRI data set but suffer from severe shortcomings. By separately analysing each programme, the author found that beneficiaries of the FFE programme consumed significantly more food than non-beneficiaries, while the same is not true for the beneficiary of the PES programme, therefore concluding that there is a difference between the two programmes. However, Ahmed did not take full advantage of the panel structure of the data set by estimating the impact of both programmes separately. Using only cross-sectional data in the first wave of the FFE programme, the study failed to control for unobserved heterogeneity across households (especially among beneficiaries and non-beneficiaries of the programme), which again results in a biased estimation of programme impact. The study also use only one-third of the sample to estimate the impact of the PES programme by only looking at households who lived in a location with no FFE programme in 2000 but with the PES programme later in the 2003 round. While this might allow for cleaner estimates of the impact of the PES programme using the difference-in-difference estimation technique, Ahmed did not make use of the valuable remaining two-thirds of the samples that would have allowed the testing of the effect of those who switched status. In addition, given the geographical targeting of the earlier FFE programme, a location with no FFE programme in 2000 would have been richer in comparison to a location with the programme. These differences will be problematic when comparing impact of the FFE and PES programmes. By running two separate regressions on two different samples, a formal test of the difference in the impact is not possible.

This chapter contributes to the existing literature by providing new empirical evidence on

[2]This study only make use of the first wave of the panel data set (year 2000) used in this chapter

the difference in impact of in-kind versus cash transfers using a credible identification strategy that controls for household fixed effects and exploits an interesting natural experiment in Bangladesh. This chapter benefits from the latest public release of the data from the most recent 2006 survey of the same set of households. In this period, the PES programme is still active and the availability of more data enables more precise estimation of programme impact. Only a study by Baulch (2010) on the medium-term impact of the two education transfer programmes has been done using the full three waves of the data set. Baulch found the impact of these education transfer programmes to be very small relative to a programme of its size. However, unlike this chapter, Baulch's study did not focus on comparing in-kind and cash transfers. This chapter, to the best of my knowledge, is one of the first that used the new complete IFPRI data set on Bangladesh to compare the impact of in-kind and cash transfers. It is also the first study to directly compare the difference in outcome on the same set of households that lived through a period where the mode of transfer received was changed from in-kind to cash. The evidence presented in this chapter thus should contribute to current debates on in-kind versus cash transfers.

Using the fixed effect instrumental variable (FE-IV) approach in the analysis, which controls for unobserved time-invariant household-specific fixed effects and the endogeneity of programme participation, the difference in the impact of both programmes on calorie consumption was found not to be statistically significant. However, a closer look at the expenditure share of various food and non-food items finds an indication that households that received an in-kind food transfer were constrained to consume food as a higher share of their total expenditure than they would have chosen under an equal-valued cash transfer. This in part is due to the constraints caused by the in-kind wheat transfer, which was both extra-marginal and binding for most households.

The findings in this chapter should ease some paternalistic fears, as cash households who were under no constraint and therefore able to switch the composition of their expenditures from food to non-food more easily chose to increase their spending on children's education, and purchase useful items such as clothing and children's non-food consumption, while at the same time spending no more on vices such as cigarettes. Given the lack of significant difference in calorie consumption, it is not surprising that no differences were found in the health status of children between households under the two programmes and, perhaps more

importantly, no difference compared to non-beneficiary households either.

The structure of this chapter is as follows. Section 2.2 outlines a simple model of consumer demand under in-kind and cash transfers that underlies the empirical analysis. Section 2.3 provides a detailed description of the FFE and PES programmes. Section 2.4 describes the survey design and presents the descriptive statistics from the IFPRI Chronic Poverty and Long-term Impact study in Bangladesh. Section 2.5 discusses the identification of the parameter of interest. Section 2.6 explains the econometric specification for estimating the parameters identified in section 2.5, and for testing the differences between in-kind and cash transfers. Section 2.7 presents the empirical results. Section 2.8 offers a discussion of the policy implications of the findings and concluding remarks.

2.2 Theoretical Model of In-kind vs. Cash Transfer

This section presents the simple theory of consumer demand under cash and in-kind transfers (Moffitt, 1989; Cunha, 2010) that underlies the empirical investigation in later parts of this chapter. To keep the explanation as simple as possible, the case of a single in-kind good that is transferred to households will be considered. This will sufficiently illustrate the differences in outcomes under cash and in-kind transfers without over-complicating the analysis.

Assume a household with a well behaved utility function $U(x_w, x_c)$ over two goods, wheat, x_w, and a composite good, x_c. Ignoring the issue of intra-household allocation of resources, we assume that the household acts as a single unit and maximises its utility with respect to a linear budget constraint $p_w x_w + p_c x_c \leq Y$. Where p_w, p_c and Y are the price of wheat, the price of the composite good, and household income respectively. With a cash transfer of value T, the budget constraint takes the form:

$$p_w x_w + p_c x_c \leq Y + T \tag{2.1}$$

With an equal-valued in-kind transfer of wheat $\overline{x}_w \, (= \frac{T}{p_w})$, the budget constraint takes the form:

$$p_w x_w + p_c x_c \leq \begin{cases} Y + p_w \overline{x}_w = Y + T & if \ \overline{x}_w \leq x_w \\ Y + \overline{p}_w \overline{x}_w & if \ \overline{x}_w > x_w \end{cases} \tag{2.2}$$

The budget constraint in Equation (2.2) is kinked depending on two factors; the amount of the good received in-kind $\bar{x}_w$ relative to the amount of the good consumed under an equal-valued cash transfer x_w, and the resale price $\overline{p}_w$. As shown above, where the in-kind transfer is "infra-marginal", that is the fixed amount of in-kind transfer received is less than the quantity the household chooses to consume of that good under an equal-valued cash transfer ($\overline{x}_w \leq x_w$), the in-kind transfer is identical to the cash transfer. This is because the household could use the income that is saved from not having to spend on the transferred good to spend freely on other goods.

However, where the transfer is "extra-marginal", that is the fixed amount of in-kind transfer is more than what the household would choose to consume with an equal-valued cash transfer ($\overline{x}_w > x_w$), an in-kind transfer will distort the household's consumption. The

extent to which this distortion affects the household's budget constraint depends on the resale price $\overline{p}_w$ of the transferred good. Here, we assume that the resale price lies between zero and the market price for the good ($\overline{p}_w \epsilon [0, p_w]$) which reflects the search and transaction cost involved in the reselling process. Unless $\overline{p}_w = p_w$, the budget constraint under the in-kind transfer is kinked at the point where ($\overline{x}_w = x_w$) and the slope of the kink becomes shallower the bigger the discount in re-sale price from the market price.

Figure 2.1: Cash versus In-kind Transfer

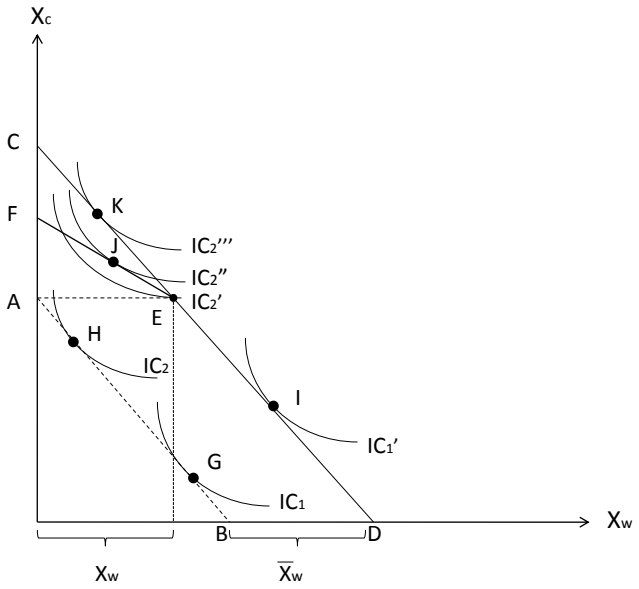


Figure 2.1 illustrates the above discussion. The original budget line before the household receives any type of transfer is given by $\overline{AB}$. The budget line shifts outward to $\overline{CED}$ when the household receives the cash transfer of $T = p_w \bar{x}_w$. With an equal-valued in-kind transfer, the budget line also shifts outward but takes the kinked form ($\overline{FED}$). The kink occurs at point E, where the transferred amount of wheat in-kind ($\overline{x}_w$) is equal to the amount the household chooses to consume under the cash transfer (x_w), meaning the distance is $\overline{AE} = \overline{BD}$. The slope of the kink beyond point E depends on $\overline{p}_w$; the closer the resale price is to the market price, the closer the kinked budget line is to the cash budget line. In the extreme case, where no in-kind good can be resold or reselling is forbidden, the budget line is $\overline{AED}$.

Whether these differences will result in distortions in household consumption depends on three conditions. First is whether the in-kind transfer is extra- or infra-marginal, in other words, whether the tangent point of the household indifference curve to the cash budget line ($\overline{CDE}$) lies above point E or not. From Figure 2.1, the preference of a household for whom

the in-kind transfer is infra-marginal is represented by the indifference curve $IC_1$and IC_1'. For this household, the outcome under cash and in-kind transfers is identical and is represented by point I on the diagram. For a household whose in-kind transfer is extra-marginal, the indifference curve is given by IC_2. Under the cash transfer, the outcome is point K on the indifference curve IC_2'''. Under equal-valued in-kind transfer, K is unattainable and the household is always on a lower indifference curve in all cases except when the in-kind good can be resold at the market price.

The second condition is the degree to which the in-kind transfer is "binding", which links directly to the resale price. An in-kind transfer is considered "fully binding" if a household consumes exactly the amount of good that was given to it, "partially binding" if the household consumes less than the amount of good that was given but more than what it would have chosen to consume under equal-valued cash transfer, and lastly "non-binding" if it consumes the same amount as it would have chosen under equal-valued cash transfer. Similarly, given no restriction on reselling, a transfer is partially binding for any positive resale price except for when $\bar{p}_w = p_w$, where transfer is non-binding. The in-kind transfer is only fully binding when the resale price is zero.

Point E on indifference curve IC_2' represents an outcome when the in-kind transfer is both extra-marginal and fully binding. A household at E who is either not allowed to resell any of their transfer or is facing a zero resale price is constrained to consume all of the transfer received. When the resale price is strictly positive but less than the market price p_w the transfer is only partially binding, and the distortion from an extra-marginal in-kind transfer is less. At point J on IC_2'', the household consumes less wheat than at point E and is better off, but is still constrained to consume more than it would have under an equal-valued cash transfer (point K, where in-kind transfer is non-binding). It must be noted that this model is time-independent. This leaves re-sale as the only explanation for observed consumption level that is not fully-binding whereas, in practice, households may choose to store or consume received goods in a lumpy manner.

The third condition is the degree of substitutability between the in-kind good and the composite good. The more substitutable the in-kind good is with the composite good the lower the welfare loss due to provision of an extra-marginal in-kind transfer. This is not shown in Figure 2.1 but could easily be demonstrated to be true by drawing a shallower

indifference curve to represent the higher degree of substitutability between wheat and the composite good.

The model can be straightforwardly extended to the case of multiple in-kind transferred goods and multiple non-transferred goods. With multiple transferred goods, the results of the model above holds for each good separately (Neary and Roberts, 1980). When more than one non-transferred good is available, the distortion caused by the extra-marginal in-kind transfer is less, because the household would be induced to consume fewer substitutes and more complements of the transferred goods compared to when these goods are not available. For example, in this case, households could respond to extra-marginal transfers of wheat by lowering rice consumption, thus keeping overall consumption of cereals constant.

In conclusion, the theoretical model shows that household utility with cash transfers is always superior to that with equal-valued in-kind transfers, as there are some outcomes under cash transfers that are not attainable with in-kind transfers. The model puts emphasis on the distinction between infra-marginal and extra-marginal in-kind transfers, with the latter resulting in lower utility for households than cash. The welfare loss, however, also depends on the resale price and the degree of substitutability of the in-kind good with other goods. A lower degree of extra-marginality, higher resale price or higher degree of substitutability of in-kind good with other goods will result in a lower welfare loss.

The presented section suggests that a cash transfer is always at least as good as an in-kind transfer. A utility-maximising household would, therefore, prefer to receive transfers in cash and freely choose their consumption bundle. However, as mentioned earlier, paternalism is one of the main justifications for the extensive use of distorting in-kind transfers over cash transfers. In considering the goals of paternalism, it is important to distinguish between whether the society cares about recipients' consumption of the specific item being transferred, in this case wheat and rice, or whether it cares about the outcomes that these in-kind transfers are designed to promote. If we believe that it is the latter rather than the former, one appropriate test for the advantage of paternalistic in-kind food transfers is whether the health and nutritional status of recipients improves significantly compared to when they receive equal-valued cash transfer programmes.

The theoretical model has shown that an extra-marginal in-kind transfer can lead to a constraint in the behaviour of the recipient that would not be the case with an equal-valued

cash transfer. These constraints were argued to be the motivation behind the use of in-kind transfers as a method of redistribution by a paternalistic government that believes this constraint to the behaviour of the recipient will lead to an improved recipient and, hence, social welfare.

There are two main hypotheses to be tested in this chapter. First is whether an in-kind transfer leads to a constraint in behaviour in comparison to an equal-valued cash transfer. This is a test of whether in-kind transfers are extra-marginal and binding. We look at the effect of both types of transfers on *per capita* calorie consumption and will examine changes in the expenditure shares of various food and non-food items to determine the effect of both programmes on the composition of consumption. This can help determine household substitution behaviour and also the difference in household consumption allocation under cash and in-kind programmes.

Second, we test whether the in-kind transfer programme leads to improved health status of household members, especially young children. We look at the effect of both types of transfer on health outcomes among children in each household and test whether the two programmes lead to a significant difference in these outcomes. The next two sections describe in detail the background of the transfer programme and the data set used in the empirical analysis.

2.3 An Overview of the FFE and PES Programmes in Bangladesh

This section presents the origins and the overview of the two conditional cash transfer programmes in Bangladesh studied in this chapter; the FFE and the PES programmes.

2.3.1 The Food for Education Programme

The FFE programme was launched in 1993 after the discontinuation of the Government of Bangladesh's targeted food subsidy programme known as Rural Rationing (Palli). Palli provided unconditional food rations to poor households. However, the high level of leakage and the high cost of the subsidy led the government to end the programme. Concerned about the food security of the 6.1 million households that were formerly entitled to subsidised rural ration, the government started the conditional transfer FFE programme (Ahmed and Khondkar, 2010). It was the aim of the paternalistic FFE programme to effectively distribute food to poor families while increasing children's educational attainment.

The FFE programme provided a free monthly ration of food grain to households conditional on the household's children being enrolled in school. Poor children enrolled in primary school grade 1 to 5 were eligible. The FFE programme was implemented selectively, where unions that were economically disadvantaged and had a low literacy rate were targeted. At its start, the programme only covered 460 unions, one union in each of the rural *upazilas* in Bangladesh[3]. Before the programme was terminated it had expanded to 1,255 unions in 2002, covering 27 percent of all primary schools and enrolling about one-third of all primary-school students in Bangladesh. The FFE programme accounted for a significant share of Bangladesh's expenditure on primary education, increasing from 4.7 percent in 1993/94 to 19.9 percent in 1997/98 (Ahmed and del Ninno, 2002).

Within the unions where FFE programme were available, households with primary-school-age children became eligible for FFE benefits if they met at least one of the following four eligibility criteria:

1. Children from a landless or near-landless household (i.e. one that owns less than half an acre of land);

2. Children of day labourers;

[3]The administrative structure of Bangladesh consists of divisions, districts, upazilas, and unions in decreasing order by size. There are 6 divisions, 64 districts, 489 upazilas (29 are in four city corporations), and 4,463 unions (all rural).

3. Children from female-headed households; or

4. Children from households that earn their living from low-income occupations.

Beneficiary households were chosen by local groups, who, based on the targeting criteria, prepared a list of FFE beneficiary households in every eligible union at the beginning of each year. Households within an FFE union who satisfied these selection criteria should have been eligible for participation, but two important conditions could prevent them from receiving the benefits. The first was that only families with students who enrolled in FFE schools could receive the food subsidy. Some students from eligible household may choose to enrol in non-FFE schools, making them unable to receive the food subsidy.

Second, due to resource constraints, only a maximum of 40 percent of students in each FFE school could receive the FFE rations. If there were more than 40 percent of students who were eligible, decisions on who should receive the subsidy were made by teachers. Teachers were expected to select the poorest households when faced with this problem (Meng and Ryan, 2010). However, due to the lack of clear-cut guidelines or an empirical method for identifying the poorest students, deliberate or unintended bias and distortions in student selection could occur (Tietjen, 2003).

Beneficiary households were entitled to receive a free ration of up to **15 kg of wheat** or **12 kg of rice** per month for sending one child of the appropriate age to primary school. The monthly ration increased to 20 kg of wheat or 15 kg of rice if households send more than one child to school[4]. Children needed to attend 85 percent of the total classes in each month to maintain their eligibility status.

2.3.2 The Primary Education Stipends Programme

The PES programme replaced the FFE programme in July 2002. Instead of subsidised food grain, the PES programme provides cash assistance to poor families conditional on sending their children to primary school. Households who meet the PES programme eligibility criteria are entitled to receive **BDT100** per month for having one child enrolled in primary school and BDT125 per month for having more than one child enrol in primary school[5].

[4]Of the total quantity of FFE food grain distributed from 1997/98 to 1999/00, wheat accounted for about 64 percent and rice, about 36 percent (Ahmed, 2005).

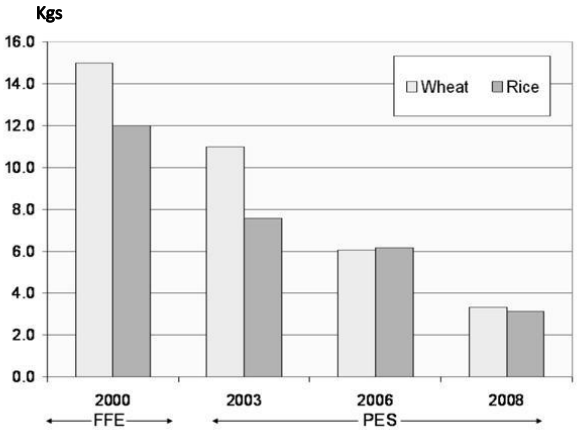
[5]BDT stands for Bangladesh Taka, the currency of Bangladesh. The official exchange rate was BDT69.89 to USD1.00 in 2007 on average.

The PES programme uses the same four targeting criteria that FFE used with one additional criterion; children belonging to a household that derives its income from sharecropping are also eligible for the PES benefits. Unlike its FFE predecessor, the PES programme does not use any geographic targeting mechanism and covers all rural areas. Based on the five targeting criteria, the School Management Committee with the assistance of head teachers prepares a list of potential PES beneficiary households in every union at the beginning of the year. Like with the FFE, the number of beneficiaries may not exceed an upper limit of 40 percent of students in a school. To maintain eligibility, students must attend 85 percent of classes each month and attain a minimum of 50 percent marks on the annual examination (Ahmed 2005).

During the first phase (2002/03 to 2006/07), the Ministry of Primary and Mass Education estimated that 5.5 million children received the stipend at a budgetary cost of BDT2.82 billion (approximately $45 million). The PES programme is now extended for a second phase for a further five years with a budget allocation of BDT2.44 billion ($35 million) and 4.8 million children are expected to benefit from the programme (Baulch, 2010). A detailed account of the historical development of both the FFE and PES programme can be found in the paper by Ahmed and Khondkar (2010).

Finally, in addition to the differences in the mode of transfer and targeting criteria of the two programmes, it is important to note the gradual decline in the value of the transfer. The nominal value of the PES transfer has been fixed since the programme's inception in 2003. Given inflation, this means that the real value of the transfer has been falling over time. Figure 2.2 displays the decline in the real value of the transfer by showing the amount of atta (wheat flour) and rice that the transfer would have been able to purchase compared to the 15 kilograms of wheat and 12 kilograms of rice provided by the earlier FFE programme. Table 2.1 gives a brief summary of the similarities and differences between the two programmes.

Figure 2.2: The Declining Value of the Transfer



Source: "Medium-term impact of the Primary Education Stipend in Bangladesh"; Baulch (2010), Figure 1
Note: Kgs= Kilograms; FFE= Food for Education; PES= Primary Education Stipend

Table 2.1: Programme Descriptions

Programme	Year of Operation	Programme Location	Eligibility Criteria	Transfer
Food for Education (FFE)	1994-2002 (2000 round of survey)	Available in selected disadvantaged union. (Two third of the unions in the survey were FFE unions in the year 2000)	1. Day labour 2. Female Headed household 3. Low Income occupation 4. Own less than 0.5 acre of land	Food grain: 15kg of wheat or 12kg of rice
Primary Education Stipend (PES)	2002-current (2003 and 2006 round of survey)	Available in all rural union. (All unions are PES union since 2003)	1. Day labour 2. Female Headed household 3. Low Income occupation 4. Own less than 0.5 acre of land 5. Sharecropper	Cash: 100 BDT

2.4 Survey Design, Data and Descriptive Statistic

This section describes the Bangladesh panel data set collected by IFPRI and used in this chapter.

2.4.1 The Surveys

The initial survey, carried out in September to October of 2000, was a cross section of households during the period of the Bangladesh FFE programme. In 2003, after the replacement of the FFE programme by the PES programme, IFPRI resurveyed the same households. In late 2006/early2007, the same households that had been included in the 2003 wave were resurveyed again as part of a larger project to study poverty in rural Bangladesh. All three rounds of surveys were conducted in the same months to control for seasonality in factors such as income and agricultural production, which may affect other outcomes of interest such as expenditures. Details on each round of the survey are presented below.

2.4.1.1 The 2000 Survey Round

The 2000 round consists of 600 households in 60 villages in 30 unions in 10 *upazilas*. The sampling process begins by the random selection of 10 *upazilas* using probability proportional to size sampling (PPS) based on the *upazilas*-level population data from the 1991 census. Two FFE unions and one non-FFE union were selected randomly in each *upazilas*. Within each union, two villages were randomly selected using PPS sampling based on village-level population data from the 1991 census. Finally, 10 households that had at least one primary-school age child were randomly selected in each village from the census list of households.

The sample included both beneficiaries and non-beneficiaries of the FFE programme which were in place in 2000. A household that is not a beneficiary of the programme does not receive the transfer for one or more of four main reasons. First, they are not eligible because they did not meet any of the targeting criteria set by the programme. Second, they are eligible for the FFE benefits but are not living in unions where the FFE programme exists. Third, they are eligible to receive the benefits and live in an FFE-union but choose not to enrol their child in primary school or enrolled their child in a non-FFE school. Finally, the eligible child attended a school where more than 40 percent of children were eligible and weren't selected to participate in the programme by the School Management Committee.

2.4.1.2 The 2003 Survey Round

The 2003 follow-up survey was conducted in September to October 2003. Because of budgetary constraints, only eight of the original 10 *upazilas* were resurveyed. One of the *upazilas* was excluded from the 2003 sample frame as FFE food grain distribution was suspended in this *upazilas* because none of the FFE schools were able meet the criteria for maintaining their eligibility for the programme. One additional *upazilas* was randomly dropped for the financial reason stated above.

As a result, the 2003 round consists of 473 households in 48 villages in 24 unions in eight *upazilas*. Even though the household sample was reduced from 600 to 473, only 7 of these lost households should strictly be regarded as being lost as a result of attrition. It is also important to note that, although the same proportion of former FFE and non-FFE union were maintained in 2003, the new PES programme was available in all rural unions.

2.4.1.3 The 2006/07 Survey Round

The 2006/07 round followed the same households in the same eight *upazilas* as the 2003 round. It consists of 511 households in 48 villages in 24 unions in eight *upazilas*. The 511 households are broken down into 455 core households that can be traced back to the original 2000 survey plus 56 households that had split from the core household between 2003 and 2006. Household splits were not followed between the 2000 and 2003 surveys, and in 2006 only those split households who did not migrate outside their home districts were tracked.

Focusing on the core households, the attrition rate over three waves is approximately 5.2 percent, a low rate of attrition in comparison to similar surveys in other low-income countries. However, it is important to test whether attrition occurs randomly or not, and this is dealt with in Section 2.6.5. In addition, anthropometric data were collected for all household members in this survey round, which enables a study of changes in the health status as a result of the programmes[6].

2.4.1.4 The Three-Wave Panel

The Bangladesh three-wave panel data set from 2000 to 2006 used in this chapter consists of 455 non-attrited households in 48 villages in 24 unions in eight *upazilas*. Household splits

[6] Anthropometric measures were taken in the original 2000 survey but was not taken in the 2003 survey

are dropped from the analysis as they were only tracked in the 2006 round.

2.4.2 Variable Construction and Descriptive Statistics

2.4.2.1 Variable Construction

This section discusses the construction of important dependent and independent variables from the raw data. As with most household surveys in low-income countries, there were problems with incomplete sections and missing data when using the IFPRI Bangladesh data set. It is important to deal with these problems correctly and effectively as mistakes in variable construction could lead to misleading and ultimately biased analysis.

Demographic: Household size was calculated by counting those individuals who reported themselves to be current members of the household. This also includes those who were temporary members, as their presence would have had impacts on other variables of interest such as expenditure and calorie consumption. Not including them could lead to an overestimation of expenditure and calorie consumption *per capita*.

Expenditure: Total household expenditure was calculated from food consumption and the non-food consumption section of the survey. The period of recall for all expenditure on food items was either the past seven or three days. The period of recall for non-food expenditure varied including one week, one month, three months, and one year. There were nearly 300 food items and 70 non-food items included in the survey questionnaire. Households were asked to specify the quantity consumed and the unit price at which they purchased the goods. These are converted in to monthly expenditures. When prices were missing for a household, village median prices are used instead[7]. Some food items collected in the first round were not recollected in later rounds. These items were dropped from the expenditure calculation to ensure consistency across all three rounds. All the expenditure-related variables are then converted in to 2006 real expenditures using the overall Bangladesh yearly food and non-food CPI[8]. It is important to note that FFE in-kind transfers received by the household were included in the expenditure calculation. While these in-kind transfers are reported at

[7]When village median prices were missing, upazilas median prices were used. When upazilas median prices were missing, union median prices were used. This go up until the biggest unit which are the median prices over all sampled household.

[8]Inflation source: Economic Trend, Bangladesh Bank and Monthly Statistical Bulletin, Bangladesh Bureau of Statistics

the price of zero, we converted the transfers to expenditures by multiplying the reported quantity received with the village median prices similar to before.

Calories: Household daily calorie consumption was calculated from the same food consumption file, where the quantity of food consumed was converted into kilocalories using the conversion file from Helen-Keller International provided by IFPRI. Again, in-kind food transfers received were converted and included in the final household calorie consumption. Food items in year 2000 that were not recollected in later rounds were eliminated to ensure consistency across all three years.

Health Outcomes: Health outcomes were calculated from anthropometric measurement taken in the 2000 and 2006 rounds. No anthropometric survey data were collected in the 2003 round and only the anthropometrics of young children and pregnant mothers were taken in 2000. Due to data limitations, we examined two health outcomes: the proportion of stunted and underweight children aged 0 to 12 years old in a household. Stunted children are those with height-for-age z-score two standard deviations below the World Health Organization (WHO) new 2005 reference group. Underweight children are those with a Body Mass Index z-score two standard deviations below the same WHO 2005 reference group.

Beneficiary Status: Determining the actual beneficiaries of the two programmes was difficult, as there were no questions in the surveys that directly asked households about their beneficiary status. As a result, we have chosen to report a household as a beneficiary of the programme if they reported receiving any amount of the in-kind food or cash benefits in the past month. Those who received the food grain or cash were given a value of one and classified as a beneficiary while those who did not report getting any transfer were given a value of zero and classified as a non-beneficiary in the beneficiary status indicator variable.

Eligibility Status: Households' eligibility statuses were calculated following the strict eligibility rules of the programme described in Section 2.3. Due to the slight difference in the eligibility criteria between the two programmes, two eligibility indicator variables were created, one for each programme. Eligibility for each programme was calculated for all households in all years regardless of whether the programme was available at those times and locations or not. This yields variables that in fact measured *hypothetical eligibility*. These

hypothetical eligibility variables are then interacted with indicator variables for programme existence in the union (one for each programme) in order to generate *actual eligibility*, i.e. the instrumental variables for beneficiary status of each household used later in the empirical analysis. Eligibility variables take a value of one if a household satisfies any one out of the four and five eligibility criteria for FFE and PES, respectively, and, in addition, has at least one primary-school-aged child as a member of the household. Households that do not satisfy these conditions are classified as ineligible and are given a value of zero.

2.4.2.2 Descriptive Statistics

This subsection presents descriptive statistics calculated from the IFPRI's Chronic Poverty and Long-term Impact study in Bangladesh data set. As the two conditional transfer programmes studied in this chapter are geographically targeted towards rural areas in Bangladesh, the descriptive statistics presented are only representative of the rural population of Bangladesh. Here and throughout the rest of this section we concentrate on households who were present in all three years of the survey, which consists of 455 households in total in each year.

Table 2.2 classifies households according to their eligibility and beneficiary status in all three rounds of survey. The first four rows in both sub-tables are exhaustive categories and add up to 100 percent. In the sample, 46.59% and 56.49% of households change their eligibility and beneficiary status at least once. These changes in status are what allow us to identify the impact of the two programmes. Out of the total sample, 34.73% were eligible for both the FFE and PES programmes, and 24.83% were the beneficiary of both FFE and PES programmes. 12.09% of households changed eligibility status at least once and were eligible for both programmes, while the same statistic for beneficiary status is 18.46%. These households will allow us to identify the effect of the two programmes on the same households. Due to the problem of endogeneity of the beneficiary status, we chose to identify the impact of both programmes through exogenous eligibility status. This will be discussed in more detail in the next section.

Table 2.2: Change in status over years

Eligibility Status	Households (%)	Beneficiary Status	Households (%)
Eligible in all 3 years	22.64	Beneficiary in all 3 years	6.37
Eligible for 2 years	26.59	Beneficiary for 2 years	24.40
Eligible for 1 year	20	Beneficiary for 1 year	32.09
Ineligible in all 3 years	30.77	Non-Beneficiary in all 3 years	37.14
Eligible for both FFE&PES	34.73	Beneficiary of both FFE&PES	24.83

Table 2.3 presents descriptive statistics on the main characteristics of households in the 2000 round when only the FFE programme was in place. Overall, FFE beneficiary households are poorer than non-beneficiary households with smaller total expenditures and expenditures per capita. Parents of the beneficiary households also have fewer years of education and literacy rate. It should be explained that literacy was collected based on a code in the survey, with 2 representing those who "can sign only", and 3 and 4 representing those who "can read" and "can read and write" respectively. Therefore, on average, the rural population in Bangladesh are poorly educated. The number of young children is similar, with beneficiary households having a slightly higher number.

The last five characteristics included in Table 2.3 are the five eligibility criteria stated in Section 2.3. The criterion sharecropper was not one of the eligibility criteria for the FFE programme in the year 2000 but we include it here to allow for comparison with the PES programme later. A higher proportion of beneficiary households have less than 0.5 acres of land, are day labourers and are households led by a female who is either divorced or widowed. Given that the FFE is a poverty-targeted programme, the systematic differences are not surprising. It is important to take these considerations into account once we proceed to the empirical analysis. Finally, the differences between eligible and ineligible households are similar to those observed between beneficiary and non-beneficiary households with a slight difference in that the gap between the two is wider. As the eligible population was constructed following a strict eligibility rule, the smaller gap between beneficiary and non-beneficiary households could point to a targeting problem of the FFE programme. It must be noted that eligibility in this case is the actual and not the hypothetical eligibility discussed in the variable construction section earlier, which is also true for the other tables in this section unless stated otherwise.

Table 2.4 presents the same sets of main household characteristics as before for households

in the 2003 and 2006 round when the PES programme had replaced the FFE programme. The statistics present a similar story to the FFE programme, where beneficiary and eligible households are poorer and less educated than non-beneficiary and ineligible households. Comparing across years, apart from expenditure, the household characteristics of the eligible and beneficiary households hardly change. This reflects the fact that the PES programme was a continuation of the FFE intended to solve the same problem, and target the same group of people, with the only major change being the method by which the transfers were given out.

One important point to note in this study is that because the original 2000 sample was designed to include households with at least one primary-school-age child, the percentage of primary-school-age children declines as these children become older in later survey waves. This is shown in Table 2.5. This means that the number and percentage of beneficiaries also falls, as to be eligible for transfer the household must enrol at least one child in primary school. Thus, the gradual decline in the percentage of beneficiary households does not reflect a change in the targeting criteria but rather the ageing of the panel. This helps provide a natural source of variation for changes in eligibility status. In Table 2.5, we use the hypothetical eligibility to illustrate our point regarding the targeting criteria between the two programmes. There are minimal differences in the percentage of households eligible for both programmes in all three years of the survey.

In addition, as the survey excluded those households with no or older children, these statistics are not representative of the overall rural population in the area but, as the main focus of this study is to compare the effect of in-kind and cash transfers to households, as long as households who received or were eligible for these two programmes had similar characteristics, which seems to be the case from evidence in Table 2.3 and Table 2.4, we can proceed with confidence.

Table 2.3: FFE Summary Statistics (2000)

	Non-Ben	Beneficiary	p	Ineligible	Eligible	p	Total
Household size	5.65	5.52	0.520	5.87	5.26	0.001^{***}	5.60
Household Expenditure	5160.62	3986.91	0.006^{***}	5554.04	3678.96	0.000^{***}	4729.83
Per capita Expenditure	919.30	736.32	0.003^{***}	954.07	722.17	0.000^{***}	852.14
Father Years of Education	2.69	2.14	0.125	3.17	1.62	0.000^{***}	2.49
Mother Years of Education	1.90	1.44	0.095^{*}	2.18	1.15	0.000^{***}	1.73
Father Literacy	2.34	2.16	0.201	2.51	1.97	0.000^{***}	2.27
Mother Literacy	2.31	2.20	0.384	2.41	2.09	0.005^{***}	2.27
Children age 0 to 12	2.36	2.58	0.063^{*}	2.37	2.52	0.191	2.44
Near Landless	0.52	0.70	0.000^{***}	0.34	0.90	0.000^{***}	0.59
Daylabour	0.18	0.29	0.007^{***}	0.12	0.35	0.000^{***}	0.22
Female Headed Household	0.10	0.13	0.258	0.08	0.15	0.015^{**}	0.11
Low Income Occupation	0.02	0.02	0.657	0.02	0.03	0.302	0.02
Sharecropper	0.08	0.07	0.758	0.06	0.10	0.102	0.08
Observations	167	288		200	255		455

Main statistics are the mean.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Author's Calculation.

Table 2.4: PES Summary Statistics (2003 & 2006)

	Non-Ben	Beneficiary	p	Ineligible	Eligible	p	Total
Household size	5.11	5.77	0.000^{***}	5.17	5.48	0.018^{**}	5.32
Household Expenditure	6088.01	5279.50	0.001^{***}	6421.70	5207.88	0.000^{***}	5832.13
Per capita Expenditure	1254.83	933.98	0.000^{***}	1319.70	977.08	0.000^{***}	1153.28
Father Years of Education	2.68	2.08	0.023^{**}	3.04	1.91	0.000^{***}	2.49
Mother Years of Education	1.84	1.49	0.080^{*}	1.93	1.52	0.030^{**}	1.73
Father Literacy	2.32	2.17	0.143	2.47	2.06	0.000^{***}	2.27
Mother Literacy	2.29	2.22	0.434	2.28	2.25	0.679	2.27
Children age 0 to 12	1.15	2.08	0.000^{***}	0.89	2.05	0.000^{***}	1.38
Near Landless	0.54	0.72	0.000^{***}	0.33	0.87	0.000^{***}	0.59
Daylabour	0.17	0.30	0.000^{***}	0.11	0.32	0.000^{***}	0.21
Female Headed Household	0.11	0.13	0.58	0.09	0.14	0.010^{***}	0.12
Low Income Occupation	0.02	0.02	0.877	0.01	0.03	0.043^{**}	0.02
Sharecropper	0.09	0.09	1.000	0.06	0.12	0.001^{***}	0.09
Observations	288	622		425	485		910

Main statistics are the mean.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Author's calculation

Table 2.5: Beneficiary/Eligibility Status by Year

	2000	2003	2006
% Primary School Aged Kid	0.98	0.78	0.60
% FFE Beneficiary	0.37	.	.
% PES Beneficiary	.	0.38	0.25
% Eligible for FFE	0.64	0.53	0.40
% Eligible for PES	0.65	0.55	0.42
Observations	455	455	455

Source: Author's Calculation

Table 2.6 and Table 2.7 present the breakdown of food expenditure for households in the 2000 FFE round and the combined 2003 and 2006 PES rounds, respectively. The main

interest lies in the two in-kind food grain subsidies in the FFE programme; wheat and rice. Wheat expenditure is the main focus as most of the in-kind subsidy provided took the form of wheat. The difference between wheat consumption of beneficiary and non-beneficiary households, and more importantly between the two programmes across years, is apparent. FFE beneficiary and eligible households consumed more wheat than non-beneficiary and ineligible households. Moreover, they consumed more wheat than PES beneficiary and eligible households in later years. Given the fact that the pattern of food consumption between the non-beneficiary and ineligible households hardly changes across the years, this observation on expenditure on wheat gives the first sign of evidence that in-kind transfers may create a constraint in the households consumer demand problem.

Looking at the food share, FFE beneficiary and eligible households spend more on food than non-beneficiaries and ineligibles, while the same is not true for households in the PES programme in 2003 and 2006, in which case the food shares between the two groups are very close to each other. In addition, FFE beneficiary households also spend a higher fraction of their expenditure on food than PES beneficiary households. As total expenditure is divided between food and non-food expenditure by construction, this also suggest that PES households spend a greater share of expenditures on non-food items than FFE beneficiary households do.

These differences may be caused by many factors other than the difference in the nature of the transfer employed by the two programmes. It will be interesting to investigate further whether the difference in the pattern of consumption observed here is a direct result of the change from in-kind food to cash transfer. If these differences were truly a result of the constraint imposed by the in-kind programme does it lead to a difference in short or long-term outcomes that change household welfare?

Table 2.6: FFE Food Expenditure Patterns (2000)

Expenditure Shares (percent)	Non-Ben	Beneficiary	p	Ineligible	Eligible	p	Total
Food groups							
Cereals	32.44	35.63	0.009***	31.32	36.52	0.000***	33.61
Wheat	**0.53**	**4.34**	**0.000*****	**0.68**	**3.51**	**0.000*****	**1.93**
Rice	29.91	29.78	0.916	28.62	31.46	0.016**	29.87
Pulses	1.75	1.93	0.378	1.79	1.84	0.805	1.81
Spices	3.53	3.65	0.482	3.50	3.68	0.253	3.58
Oil	1.99	2.04	0.665	2.01	2.01	0.982	2.01
Fruit & Vegetables	13.52	14.35	0.207	13.57	14.16	0.353	13.83
Meat & Fish	16.99	15.87	0.222	17.63	15.24	0.007***	16.58
Out of Home	1.45	1.14	0.188	1.46	1.18	0.236	1.34
Drink	0.76	0.57	0.101	0.74	0.62	0.277	0.69
Other	3.08	2.77	0.244	3.16	2.72	0.090*	2.97
Total Food	**75.52**	**77.93**	**0.072***	**75.17**	**77.97**	**0.032****	**76.40**
Observations	167	288		200	255		455

Main statistics are the mean.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Author's Calculation

Table 2.7: PES Food Expenditure Patterns (2003 & 2006)

Expenditure Share (percent)	Non-Ben	Beneficiary	p	Ineligible	Eligible	p	Total
Food groups							
Cereals	27.81	31.09	0.000***	27.38	30.40	0.000***	28.85
Wheat	**0.30**	**0.32**	**0.836**	**0.23**	**0.39**	**0.100**	**0.31**
Rice	25.31	28.88	0.000***	24.85	28.12	0.000***	26.44
Pulses	1.73	1.89	0.327	1.69	1.88	0.202	1.78
Spices	3.49	4.16	0.003***	3.64	3.77	0.524	3.71
Oil	2.48	2.57	0.351	2.56	2.45	0.213	2.51
Fruit & Vegetables	14.95	14.54	0.344	15.09	14.54	0.170	14.82
Meat & Fish	15.76	13.86	0.009***	15.81	14.46	0.048**	15.16
Out of Home	4.26	3.18	0.029**	3.80	4.04	0.605	3.92
Drink	0.86	0.64	0.041**	0.93	0.65	0.007***	0.79
Other	2.12	1.65	0.010**	2.18	1.75	0.011**	1.97
Total Food	**73.46**	**73.58**	**0.846**	**73.08**	**73.94**	**0.118**	**73.50**
Observations	288	622		425	485		910

Main statistics are the mean.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Author's Calculation

2.5 Identification

This section discusses the identification issue in this chapter. This involves identifying the quantities of interest that will be estimated in the subsequent sections. The non-experimental nature of the data poses various problems that must be dealt with in order to correctly obtain the objects we want to estimate. An instrumental variable and fixed-effect approach is proposed to deal with these problems. Here we argue that these techniques and assumptions are sufficient for the identification of the parameter of interest.

2.5.1 Parameter of Interest

The main object of interest in this chapter is the difference in the effect of cash versus in-kind transfers on various household outcomes such as calorie consumption and health status. The approach taken is to identify the impact of in-kind transfers by comparing the effects of the programme on outcomes of in-kind beneficiary households to non-beneficiary households. Similarly, the impact of the cash transfer is determined by comparing the outcomes of cash beneficiary households to non-beneficiary households. The in-kind and cash impacts are then compared to each other to investigate whether any significant difference exists. This allows us to test whether the change in the medium of the transfer has any impact on household behaviour. To simplify the discussion without the loss of generality in this case, we will focus on just calorie consumption as the outcome variable.

One of the main quantities that we want to estimate is the causal effect of being a beneficiary of the in-kind transfer programme on household calorie consumption. Allowing for heterogeneous programme effects, we assume two potential outcomes variables:

$$Potential\,outcome\,of\,in-kind\,transfers\; = \begin{cases} Y_{i,t}^{K}(1) & if\; B_{i,t}^{K}=1 \\ Y_{i,t}^{K}(0) & if\; B_{i,t}^{K}=0 \end{cases}$$

$B_{i,t}^{K}$ is an indicator of household i beneficiary status under the ***in-kind*** transfer programme (superscript K) at time t, taking the value 1 when the household is a beneficiary and 0 otherwise. $Y_{i,t}^{K}(1)$ is household i's calorie consumption when it is a beneficiary of the in-kind transfer programme, while $Y_{i,t}^{K}(0)$ is household i's calorie consumption if it is not a beneficiary at time t. The difference between $Y_{i,t}^{K}(1)$ and $Y_{i,t}^{K}(0)$ is the causal effect of being a beneficiary of the in-kind programme on calorie consumption of household i at time t and

is the object that we are interested in. The observed outcome $Y_{i,t}^K$ can be written in terms of potential outcomes as

$$Y_{i,t}^K \;=\; Y_{i,t}^K(1)B_{i,t}^K \;+\; Y_{i,t}^K(0)(1 \;-\; B_{i,t}^K)$$

We get to observe either $Y_{i,t}^K(1)$ or $Y_{i,t}^K(0)$, but never both. We therefore hope to measure the average of $Y_{i,t}^K(1) \;-\; Y_{i,t}^K(0)$. The naive comparison of the outcomes of those who were and were not beneficiaries of the programme is likely to be a poor measure of the causal effect of the programmes given their non-experimental nature.

The conditional naive comparison of beneficiary and non-beneficiary households can be decomposed as

$$\underbrace{E[Y_{i,t}^K|B_{i,t}^K=1,X_{i,t}=x]-E[Y_{i,t}^K|B_{i,t}^K=0,X_{i,t}=x]}_{Naive\,comparison}=\underbrace{\{E[Y_{i,t}^K(1)|B_{i,t}^K=1,X_{i,t}=x]-E[Y_{i,t}^K(0)|B_{i,t}^K=1,X_{i,t}=x]\}}_{ATT^K}$$
$$+\underbrace{\{E[Y_{i,t}^K(0)|B_{i,t}^K=1,X_{i,t}=x]-E[Y_{i,t}^K(0)|B_{i,t}^K=0,X_{i,t}=x]\}}_{Selection\,bias} \tag{2.3}$$

The first term in the curly bracket on the right-hand side of Equation (2.3) is the conditional average treatment effect on the treated of the in-kind transfer programme. ATT^K measures the effect of the in-kind programme on those who are the beneficiary of the programme. In this case we are particularly interested in ATT^K conditional on observables $X_{i,t}$ as there are various observable differences such as income and household structure between beneficiary and non-beneficiary households, due to the non-random assignment of treatment, that also have a direct impact on the outcome of interest. As it is the direct impact of the programme on beneficiary households that we want to know, it is necessary to control for these observable differences. Unconditional ATT^K is then obtained by taking the expectation over all conditional ATT^K[9]. ATT^K is a parameter of interest for policymakers implementing the programmes, and is the object we look to estimate in this chapter.

Another object of interest, as stated before, is the impact of the cash programme on households' calories consumption. Similar to before, we make use of the potential outcome framework:

$$Potential\,outcome\,of\,cash\,transfers \;=\; \begin{cases} Y_{i,t}^C(1) & if\; B_{i,t}^C = 1 \\ Y_{i,t}^C(0) & if\; B_{i,t}^C = 0 \end{cases}$$

[9]Making use of the law of iterated expectation $E_X\{E[Y_{i,t}^K(1)|B_{i,t}^K = 1, X_{i,t} = x] - E[Y_{i,t}^K(0)|B_{i,t}^K = 1, X_{i,t} = x]\} = E[Y_{i,t}^K(1)|B_{i,t}^K = 1] - E[Y_{i,t}^K(0)|B_{i,t}^K = 1]$

Here, $B^C_{i,t}$ is an indicator of household i beneficiary status under the ***cash*** transfer programme (superscript C) at time t, taking the value 1 when household is a beneficiary and 0 otherwise. $Y^C_{i,t}(1)$ is household i's calorie consumption when it is a beneficiary of the cash programme, and $Y^C_{i,t}(0)$ is household i's calorie consumption if it is not a beneficiary of the cash programme at time t. The difference between $Y^C_{i,t}(1)$ and $Y^C_{i,t}(0)$ is the causal effect of being a beneficiary of the cash programme on calorie consumption of household i at time t and is the object that we are interested in. The observed outcome $Y^C_{i,t}$ can be written in terms of potential outcomes as

$$Y^C_{i,t} \;=\; Y^C_{i,t}(1)B^C_{i,t} \;+\; Y^C_{i,t}(0)(1 \;-\; B^C_{i,t})$$

As before, we get to observe either $Y^C_{i,t}(1)$ or $Y^C_{i,t}(0)$, but never both. We therefore hope to measure the average of $Y^C_{i,t}(1) \;-\; Y^C_{i,t}(0)$. The naive comparison of the outcome for those who were and were not beneficiaries of the cash programme is

$$\underbrace{E[Y^C_{i,t}|B^C_{i,t}=1, X_{i,t}=x] - E[Y^C_{i,t}|B^C_{i,t}=0, X_{i,t}=x]}_{Naive\,comparison} = \underbrace{\{E[Y^C_{i,t}(1)|B^C_{i,t}=1, X_{i,t}=x] - E[Y^C_{i,t}(0)|B^C_{i,t}=1, X_{i,t}=x]\}}_{ATT^C}$$
$$+\underbrace{\{E[Y^C_{i,t}(0)|B^C_{i,t}=1, X_{i,t}=x] - E[Y^C_{i,t}(0)|B^C_{i,t}=0, X_{i,t}=x]\}}_{Selection\,bias} \tag{2.4}$$

The first term on the right-hand side of Equation (2.4) is the conditional average treatment effect on the treated of the cash programme (ATT^C), which measures the effect of cash programmes among those who were treated conditional on a set of observable differences $X_{i,t}$. Again, making use of the law of iterated expectation, we can obtain the unconditional ATT^C by taking expectation over the conditional ATT^C. This unconditional ATT^C is another parameter of interest in addition to ATT^K mentioned earlier, and it is the comparison between ATT^K and ATT^C that is the ultimate quantity of interest.

However, the major obstacle to the correct identification of both ATT^K and ATT^C from the naive estimator in Equation (2.3) and (2.4) is the second term on the right-hand side of both equations. These two terms are the *selection bias*, the differences in the non-treatment outcome among those who are the beneficiaries and those who are not. Given the poverty-targeting nature of both the FFE and PES programmes, selection bias will be a major problem. For example, households that are beneficiaries of the programme will be poorer on average and it is likely that they would have been able to consume less calories in the first place. The selection bias in this case is expected to be negative and the naive comparison

would likely underestimate the true impact of the programme. However, selection based on households' observable characteristics, such as income as above, could be accounted for by conditioning on such observables. If it is only selection on observables that is driving the bias, the conditional selection bias term in both equations would be zero and the conditional naive comparisons would give us consistent estimates of the conditional ATT^K and ATT^C[10].

However, it must also be noted that there are also other unobserved factors that are correlated with treatment status and could contribute to the selection bias in one direction or the other. In an ideal randomised experiment these unobserved factors will tend to balance out between treatment and control groups. Given that this is not the case with the FFE and PES programmes, the conditional selection bias term will be non-zero and these unobserved factors will exacerbate the conditional selection bias and cause the conditional naive estimators to be inconsistent estimates of the true impact of the programmes.

Selection on unobservables in this case often presents itself as the problem of partial compliance. Partial compliance is a situation when not all those who are eligible will choose to take up the benefits. Compliance is a choice made by a household that is likely to be correlated with unobserved characteristics of the household. If these characteristics affect the outcome variable, partial compliance will affect the selection bias. Partial compliance almost always leads to overestimation of the true programme impact as those who are likely to benefit most from the programmes are those who take it up while those who are unlikely to benefit will not.

There are two types of unobservable characteristics, those that are fixed across time and those that vary across time. One example of fixed unobservable characteristics is household preference for food consumption. In a partial compliance scenario, it is possible that households that choose to become beneficiaries of the in-kind programme are those that have a high preference for the consumption of food and would have benefitted more from the in-kind transfer than non-beneficiary households. It is possible that these households would have consumed more food than non-beneficiary households in the non-treatment state. Non-time-varying unobservables that cause selection bias can be dealt with by the fixed-effect approach utilised later.

Time-varying unobservables such as transitory shock to individual households could also

[10]This is equivalent to making the conditional independence assumption $\{Y^K_{i,t}(0), Y^K_{i,t}(1),\} \perp\!\!\!\perp B^K_{i,t}|X_{i,t}$ and $\{Y^C_{i,t}(0), Y^C_{i,t}(1),\} \perp\!\!\!\perp B^C_{i,t}|X_{i,t}$ for both programmes

affect selection bias. For example, a household may be hit with a negative shock in one period that increases the opportunity cost of sending children to school, which ultimately affects its beneficiary status and also its calorie consumption. Not observing the negative shock may cause us to think that this is one of the better-off non-beneficiary households and could result in an incorrect comparison between beneficiary and non-beneficiary households. Here again, partial compliance, where those who are unlikely to benefit select themselves out of the programme, will cause overestimation of programme impact.

In addition, as discussed earlier, the data set could suffer from a measurement error problem. This results from the lack of clarity and inconsistencies in the questionnaire in determining the actual beneficiary status of households. Measurement error does not affect selection bias but, assuming classical measurement error, will cause attenuation bias in the naive estimator. We later used the instrumental variable approach to deal with selection based on unobservables and the measurement error problem.

2.5.2 Dealing with Selection Bias

In this subsection we present the methods and assumptions made to deal with the problems specified in subsection 2.5.1 above. The problems boil down to finding a credible estimate of the unobserved counterfactuals $E[Y^K_{i,t}(0)|B^K_{i,t} = 1, X_{i,t} = x]$ and $E[Y^C_{i,t}(0)|B^C_{i,t} = 1, X_{i,t} = x]$, and getting rid of the selection bias term in order to obtain consistent estimates of the conditional ATT of both in-kind and cash transfer programmes.

Given the poverty-targeting nature of both programmes, the assignment of households into treatment and control groups is non-random and is based on the eligibility rule stated in section 2.3. We propose to identify the counterfactual using *actual eligibility* for both programmes *($D^K_{i,t}$, $D^C_{i,t}$)*, computed from the interaction of eligibility rules and the indicator of programme availability, as an instrumental variable that is correlated with beneficiary status, but conditionally exogenous and uncorrelated with the outcome variables. We termed this the fixed-effect instrumental variable (FE-IV) approach, which will solve the problem of partial compliance and non-random assignment and placement of the programmes.

Here we make use of the potential outcomes notation for the receipt of treatment ($B^K_{i,t}$, $B^C_{i,t}$) and the outcome ($Y^K_{i,t}$, $Y^C_{i,t}$). The instruments in this case are binary: $D^K_{i,t} \epsilon \{0, 1\}$ and $D^C_{i,t} \epsilon \{0, 1\}$. Denote the level of the treatment received if the instrument takes on the value

0 and 1 respectively by $\{B^K_{i,t}(0),\ B^K_{i,t}(1)\}$ for the in-kind programme, and $\{B^C_{i,t}(0),\ B^C_{i,t}(1)\}$ for the cash programme. $B^K_{i,t}(1)$ and $B^C_{i,t}(1)$ tells us whether household i would become a beneficiary household given that it is eligible, while $B^K_{i,t}(0)$ and $B^C_{i,t}(0)$ tells us whether household i would become a beneficiary household if it is ineligible. As before, we only observe one of the potential receipts of treatment for each household depending on the instruments $D^K_{i,t}$ and $D^C_{i,t}$[11].

We made the following assumptions to identify the parameters of interest using the FE-IV approach.

- Assumption (1): Stable Unit Treatment Value Assumption (Rubin, 1980). This means that programme participation and the potential outcome of each household unit depends on the treatment received and are unrelated to the treatment status of other households. This rules out any spillover or general equilibrium effect from treated to non-treated households. However, as shown in the Table 2.A.1 in Appendix 2.A.1, the price of wheat, the main FFE in-kind transfers, is statistically significantly lower in FFE union than non-FFE union. This suggests the possibility that there may be a spillover effect as non-beneficiary households in FFE union also enjoy the benefit of the programme through lower wheat price. This spillover effect will cause the impact of the FFE programme to be underestimated. The price impact of the PES programme is harder to measured as all unions are PES union by 2003 and the increase in price across year could just be a general price increase instead of the programme effect. Nevertheless, given the possible underestimation of the FFE programme, the final estimate of the difference between the two programmes could also be underestimated and this is one of the limitations of the study. To alleviate some of the general equilibrium effect, we will include price of wheat as one of the control variables going forward.

- Assumption (2): Conditional Independence of the Instruments Assumption. In this case, we assume that conditional on time-varying observables ($X_{i,t}$) and fixed unobservables (U_i), potential outcomes and potential treatment assignments are independent of the instruments, which are the household's actual eligibility status for the FFE and PES programmes ($D^K_{i,t},\ D^C_{i,t}$).

[11]The observed treatment status is therefore $B^K_{i,t} = B^K_{i,t}(1)D^K_{i,t} + B^K_{i,t}(0)(1 - D^K_{i,t})$ for the in-kind programme and $B^C_{i,t} = B^C_{i,t}(1)D^C_{i,t} + B^C_{i,t}(0)(1 - D^C_{i,t})$ for the cash programme.

$$\{Y_{i,t}^{K}(0),\, Y_{i,t}^{K}(1),\, B_{i,t}^{K}(0),\, B_{i,t}^{K}(1)\} \perp\!\!\!\perp D_{i,t}^{K} | X_{i,t},\, U_i$$

$$\{Y_{i,t}^{C}(0),\, Y_{i,t}^{C}(1),\, B_{i,t}^{C}(0),\, B_{i,t}^{C}(1)\} \perp\!\!\!\perp D_{i,t}^{C} | X_{i,t},\, U_i$$

- Assumption (3): Instrument relevance. There is a positive correlation between the instruments and treatment status. In this case, this means that there is a non-zero relationship between eligibility status and beneficiary status of household in the programmes.

$$E[D_{i,t}^{j},\, B_{i,t}^{j} | X_{i,t},\, U_i] \neq 0 \, where \, j = K \, , \, C$$

- Assumption (4): Exclusion restriction. Formally, the exclusion restriction is

$$B_{i,t}^{j} = B_{i,t}^{j}(1) D_{i,t}^{j} + B_{i,t}^{j}(0)(1 - D_{i,t}^{j}) \mid X_{i,t,} U_i$$

$$Y_{i,t}^{j} = Y_{i,t}^{j}(1) B_{i,t}^{j} + Y_{i,t}^{j}(0)(1 - B_{i,t}^{j}) \mid X_{i,t,} U_i$$

$$where \, j = K \, , \, C$$

The instrument $D_{i,t}^{j}$ only appears in the first equation above, which determines the beneficiary status $B_{i,t}^{j}$, but is absent from the second equation which determine the outcome, $Y_{i,t}^{j}$. This means that conditional on time-varying observables and time-invariant unobservables, the instruments only affect the outcome variable through their effect on the treatment status. Specifically for this case, eligibility status affects the beneficiary status of an individual household but it does not have a direct effect on a household's calorie consumption, controlling for $X_{i,t}$ and $U_{i,t}$. A household does not consume more or fewer calories just because it is eligible for either of the programmes. Notice that unlike the case with exogenous instruments, the instrument here is only exogenous and satisfies the exclusion restriction upon conditioning on observables and the unobservables fixed effect.

- Assumption (5): Monotonicity. The instrument moved every household's responses in the same direction. In other words there is no household that behaves as a defier, i.e. a household that would have participated in the programme if it is not eligible, and not participated in the programme if it is eligible. In this case monotonicity is assumed to be

$$B_{i,t}^{K}(1) \geqslant B_{i,t}^{K}(0)$$

$$B^C_{i,t}(1) \geqslant B^C_{i,t}(0)$$

If all of these assumptions hold, the FE-IV will identify the conditional Local Average Treatment Effect (LATE) instead of the average treatment on the treated. Unlike the ATT, LATE is the average treatment effect for those households that are moved by the instrument, and is written formally as:

$$\begin{aligned} LATE^j &= \frac{E[Y^j_{i,t}|X_{i,t}=x,U_i\, D_{i,t}=1] - E[Y^j_{i,t}|X_{i,t}=x,U_i, D_{i,t}=0]}{E[B^j_{i,t}|X_{i,t}=x,U_i, D_{i,t}=1] - E[B^j_{i,t}|X_{i,t}=x,U_i, D_{i,t}=0]} \\ &= E[Y^j_{i,t}(1) - Y^j_{i,t}(0)|B^j_{i,t}(1) > B^j_{i,t}(0),\, X_{i,t}=x,\, U_i] \end{aligned}$$

$$where\ j = K\ ,\ C$$

As we needed the conditional independence of the instrument assumption for the strategy to work, the LATE we estimated is in fact a covariate-averaged LATE[12]. LATE will be explained in more detail in the next subsection.

2.5.3 Local Average Treatment Effect (LATE)

In this subsection we gave formal interpretations of the LATE to help increase our understanding of the parameters we identified in the last subsection. Though we drop the superscripts K and C, the discussion here applies to the parameters of both in-kind and cash programmes ($LATE^K$ and $LATE^C$). Finally, we considered how the LATE that we identified in this case will also be the ATT if a certain condition is satisfied.

We start by looking at the different types of population in the binary outcome and binary instrument setting. Table 2.8 below shows all the possible sub-groups of the population. Each group is classified by how the population moved with the instrument. Always-takers are those that would have received treatment whether they are eligible or not. Never-takers are those that will never take up the treatment regardless of their eligibility. Compliers are those that behave according to the instrument they were assigned to. They will take up the treatment when they are eligible and will not when they are ineligible. Lastly, we have defiers, who behave in the opposite direction of the instrument they were assigned to, taking up the treatment when they are not eligible and not taking up the treatment when they are.

[12] Abadie (2003) provided the exact method to compute the covariate-averaged LATE, which involves weighting the least square equation by "Abadie Kappa". Angrist (2001), however, showed that estimates based on Abadie's method are indistinguishable from two-stage least square estimates, which are the estimates we will use in this chapter.

Table 2.8: Types of Population

	$D_{i,t} = 1$	$D_{i,t} = 0$
Always-takers	$B_{i,t}(1) = 1$	$B_{i,t}(0) = 1$
Compliers	$B_{i,t}(1) = 1$	$B_{i,t}(0) = 0$
Defiers	$B_{i,t}(1) = 0$	$B_{i,t}(0) = 1$
Never-takers	$B_{i,t}(1) = 0$	$B_{i,t}(0) = 0$

LATE measures the treatment effect on those households that are moved by the instrument. The effect of the programme on always-takers and never-takers cannot be determined as the two sub-populations' treatment status never changed regardless of the value of the instrument. In addition, the monotonicity assumption made in the last section ruled out the presence of defiers in the population. Therefore, the LATE is the average treatment effect of the programme on compliers, those who became beneficiaries when they were eligible and would not have otherwise.

If we are willing to make a further assumption of one-sided partial compliance, that only those who are eligible can choose to not take up the programme but those who are not eligible for the programme will never be able to gain access to the programme (cannot be a beneficiary), we will also be able to rule out the existence of always-takers as well. This is because one-sided compliance implies that $B_{i,t}(0) = 0$ always, and as always-takers have $B_{i,t}(0) = 1$, it is impossible to have always-takers in this case. This seems like a plausible assumption to make in places where programmes are well and strictly implemented.

As the ATT can be decomposed into the programme effect on always-takers and compliers[13], ruling out the presence of always-takers means that those who are treated consist entirely of compliers. Therefore, If this assumption of "$B_{i,t}(0) = 0$ always" holds, the conditional LATE we identified in the last subsection using the FE-IV approach will, in this special case, be the same as the conditional ATT, which is the parameter of interest we originally wanted to identify[14].

[13] $E[Y_{i,t}(1) - Y_{i,t}(0)|B_{i,t} = 1, X_{i,t} = x] = E[Y_{i,t}(1) - Y_{i,t}(0)|B_{i,t}(0) = 1, X_{i,t} = x]P[B_{i,t}(0) = 1|B_{i,t} = 1] + E[Y_{i,t}(1) - Y_{i,t}(0)|B_{i,t}(1) > B_{i,t}(0), X_{i,t} = x]P[B_{i,t}(1) > B_{i,t}(0), D_{i,t} = 1|B_{i,t} = 1]$, The effect of conditional average treatment on the treated is equal to the conditional average effect of treatment on always-takers times the probability of the subjects being always-takers, given that they are treated, plus the conditional average effect of treatment on compliers times the probability of the subjects being compliers with the instruments taking a value of 1 (compliers will not be treated if the instrument takes a value of 0), given that they are treated.

[14] For a simple and direct proof of this result following from Bloom (1984), please refer to Angrist and Pischke (2009)

$$
\begin{aligned}
LATE &= \frac{E[Y_{i,t}|X_{i,t}=x,U_i\, D_{i,t}=1]-E[Y_{i,t}|X_{i,t}=x,U_i,D_{i,t}=0]}{E[B_{i,t}|X_{i,t}=x,U_i,D_{i,t}=1]-E[B_{i,t}|X_{i,t}=x,U_i,D_{i,t}=0]} \\
&= \frac{E[Y_{i,t}|X_{i,t}=x,U_i\, D_{i,t}=1]-E[Y_{i,t}|X_{i,t}=x,U_i,D_{i,t}=0]}{E[B_{i,t}|X_{i,t}=x,U_i,D_{i,t}=1]}\ as\ E[B_{i,t}|D_{i,t}=0]=0 \\
&= E[Y_{i,t}(1)-Y_{i,t}(0)|B_{i,t}=1,\, X_{i,t}=x]\ =\ ATT
\end{aligned}
$$

However, this assumption is questionable in the Bangladesh dataset used here. In this dataset we do have households that are beneficiaries despite being ineligibles due to problems with mis-targeting, which is common in this type of programme in under-developed countries. The rate of mis-targeting is about 8% for both the FFE and PES programmes. These figures are non-negligible, hence making the one-sided compliance assumption an invalid one. We therefore have to be content with the LATE.

ATT, which measures the effect of the programme of all those who received the benefits, is more externally valid as a parameter than LATE. This means that the ATT is more useful in providing policymakers with an idea of the effect that similar programmes, implemented under different conditions, may generate based on the impact found from this programme. Although this is a very useful property, we cannot consistently estimate the ATT of both programmes due to data limitations. The ATT, therefore, is not *internally valid*.

LATE, although not the original parameter we set out to estimate, is still useful. LATE measures the effect of the programme on the sub-population of compliers. Although this is less externally valid than the ATT, as the compliant sub-population may differ under different programme conditions or different instruments, making the resulting effect of the LATE less generalisable, it is internally valid. We take the view here that an internally valid parameter with a lower level of external validity is better than a more externally valid parameter that is not internally valid[15].

In this section, we identify two parameters that we are interested in estimating and comparing to each other, $LATE^K$ and $LATE^C$. We could obtain the estimates of both parameters by conducting two separate regressions, one for each programme. However, pooling the two regressions together is a more efficient way to both estimate and compare the two quantities. The next section discusses the issue of estimating the two parameters and provide details on the method.

[15]Imbens (2010) provides a good discussion of the usefulness of LATE.

2.6 Econometric Specification and Estimation

This section discusses the econometric specifications used in this chapter to consistently estimate the parameters identified in section 2.5 and test the hypothesis of whether there is a difference in household outcomes under in-kind and cash transfers. The outcomes of interest are: *per capita* calorie consumption, expenditure shares of various food categories, and finally, measures of health outcomes. Due to the non-experimental nature of the data, there are issues of endogeneity for programme beneficiary status. These will be corrected for using an instrumental variables (IV) approach. Importantly, the panel structure of the data make it possible to control for time-invariant unobserved households characteristics that may be correlated with the outcomes of interest.

2.6.1 The Basic Specification: In-kind vs. Cash

In this section, we present the basic specifications that form the basis of the analysis of the impact of the FFE and PES programmes. The ultimate goal is to test whether there is a significant difference between the two programmes and hence, the two types of transfer. We also present discussions of the problems that arise with this basic specification.

We are specifically looking at three major sets of outcomes; food consumption, expenditure shares and health status of children. The general specification is the same for all three outcomes. First, the general basic specification is given by Equation (2.5):

$$y_{it} = \alpha + InKind_{i,2,t} \cdot \beta_{InKind} + Cash_{i,3,t} \cdot \beta_{Cash} + x_{it}\gamma + \epsilon_{it} \tag{2.5}$$

Where y_{it} is one of the three outcomes of interests; *per capita* calorie consumption, expenditure share of various food categories or health status of children for household $i = 1, ..., N$ at time $t = 1, ..., T$. $InKind_{i,2,t}$ is an indicator variable taking the value of 1 if household i is a beneficiary of the FFE programme at time t and zero otherwise. $Cash_{i,3,t}$, is similarly an indicator variable taking the value of 1 if household i is a beneficiary of the PES programme at time t and zero otherwise. x_{it} is a $(1 \times k)$ vector of both time-variant and time-invariant control variables, including household characteristics, expenditures per capita, prices, occupational status and year; γ is the $(k \times 1)$ vector of corresponding coefficients; and ϵ_{it} is a mean zero error term, which is assumed to be orthogonal to the included regressors

and has a variance of σ_ϵ^2.

Equation (2.5) is a reduced-form equation of the impacts of the two programmes. The two main coefficients of interest are β_{InKind} and β_{Cash} which respectively identify the impact on household i at time t that received the in-kind and cash benefit relative to receiving no benefits at all. As the PES programme is a continuation of the FFE programme, a household can never be a beneficiary of both programmes in the same period. β_{InKind} and β_{Cash} in this specification correspond to the naive estimator discussed in the last section. We are interested both in the individual magnitude of the coefficient β_{Inkind} and β_{Cash}, and also their magnitudes relative to each other.

As both programmes target poverty, it is normal for the observed characteristics of beneficiary and non-beneficiary households to be different. The beneficiaries are likely to be poorer, less educated and have larger households. These characteristics ultimately affect the outcomes of interest and, therefore, by omitting them, the estimated impacts of the programmes are going to be biased. It is therefore necessary to include these observable factors as control variables x_{it} in Equation (2.5). We will discuss this in detail below.

2.6.1.1 Dependent Variables

There are three main sets of regression in this chapter: food consumption regression, expenditure share regression and health regression. In this subsection we describe the dependent variables that we use in each set of regressions.

Food Consumption Regression: the outcome variable y_{it} is daily per capita consumption of food measured in kilocalories.

Expenditure Share Regression: the outcome variables y_{it} of interest are the expenditure of various food and non-food items/categories as a share of household total expenditures.

Health Regression: The outcome variables y_{it} is a measure of the health status of children in a household. The two measures we concentrate on are the proportion of stunted and underweight children age 0 to 12 within a household.

2.6.1.2 Income Effect vs. Compositional Effect

The impact of both transfer programmes can be broken down into an income effect and a compositional effect. Referring back to the theory in Section 2.2, the income effect of both programmes is the impact of increasing income by T, causing households to spend more on food (foods are normal goods), regardless of the type of transfer. The compositional effect occurs when the food subsidy is extra-marginal, after controlling for any general equilibrium price effects of the programme. This compositional effect of the transfer causes changes in the composition of household food consumption resulting in households consuming more of some type of foods than others. If the transfer is both fully binding and extra-marginal as in point E or partially binding at any point along the kinked part of the budget line in Figure 2.1, its effect is not only to increase household wheat consumption but also to cause the household to consume more wheat than they would have otherwise, at the expense of other categories of food.

Both types of transfer have an income effect that will also be equivalent when both transfers are of equal value, while only the in-kind transfer has a compositional effect. To test for the difference between the impact of the two types of transfer on food, and hence, on calorie consumption, it is important that we control for the income effect of both programmes. This allows us to focus solely on the compositional effect of the transfer programme and determine the effect of an extra-marginal and binding transfer on various outcomes. For this purpose, the logarithm of per capita monthly expenditure is an appropriate wealth control for all specifications. By using expenditure that would have already incorporated changes in total spending out of the benefits received from the FFE and PES programmes, we can effectively control for income effects and focus on the compositional impact of both programmes.

2.6.1.3 Control Variables

In addition to the logarithm of real *per capita* monthly expenditure as a control for different levels of wealth between beneficiary and non-beneficiary households, there are several other factors for which we have controlled.

A logarithm of household size and the proportion of members in each age group is added as demographic controls that control for the compositional effect of household structure.

Year dummies are added as control variables to capture all the effects that change with time. Prices of the in-kind transferred goods are also added to capture some of the general equilibrium effect that could result from the in-kind programme relative to the cash programme, which can affect outcome $y_{i,t}$ variables. The underlying variables that affect the eligibility status of households for both the FFE and PES programmes such as the occupational status of the head and whether the household is female-headed are added to control for the difference in characteristics of eligible and ineligible households.

The control variables are the same for all three sets of regression except that the price variables are excluded from the set of control variables in the health regression, as they should not matter in that context.

2.6.1.4 Conditionality of FFE and PES programmes

Both FFE and PES are conditional on children's enrolment in primary school. Therefore, theoretically, both programmes are in fact in-kind programmes. However, the cash and in-kind comparison between the two programmes was made possible by the fact that apart from the medium of transfer, the two programmes are nearly identical in all other aspects. By controlling for all these aspects a comparison of differences in household behaviour under the two programmes will allow identification of the difference between the two types of transfers.

However, as both programmes were conditional on children's enrolment, they are likely to have a positive impact on children's enrolment in primary school among beneficiaries when compared to non-beneficiaries. It is possible for these enrolment effects to be correlated with the outcome of interest and, by ignoring this effect, the estimates of programme impact between beneficiary and non-beneficiary households will be biased. In addition, if the effect on enrolment is different between the two programmes, the estimate of the comparison between the two programmes will also be biased.

Estimation of both programmes effect on enrolment has found that both programmes have positive and statistically significant impacts on enrolment. In addition, the t-test of the two programmes found that they are statistically different, with the PES having bigger impact than the FFE programme. Failing to account for this means that the estimates of both the impact of in-kind and cash transfers on beneficiary households compared to non-beneficiary households, and compared to each other, will be biased. To solve this problem,

we add the proportion of children enrolled in primary school for each household as a control variable in all three sets of regressions. The enrolment regression results can be found in Appendix 2.A.2.

2.6.1.5 Controlling for Difference in the Values of Transfer

As shown in Figure 2.2, the real value of the transfer has been gradually declining since the year 2000. Therefore, differences observed in the effect between the two programmes, if there are any, could be due to differences in the size of the transfer under the two programmes.

To control for the differences in the value of the transfer, Equation (2.5) with the appropriate control variables is estimated with changes made to the variable $InKind_{i,2,t}$ and $Cash_{i,3,t}$; these two variables will now measure the value of the in-kind and cash transfer received by the beneficiary household respectively, as opposed to being indicator (0/1) variables as before. There are two measures of these values for both types of transfer: the actual and the hypothetical value.

The actual values of the transfers are obtained directly from the questionnaire. It is the monthly quantity/value of the transfers that the household itself reports to be receiving from the programme. In-kind food quantity transfers were converted to their monetary value by multiplying by the price variable constructed from the food expenditure data. All monetary value variables are then converted into their real 2006 value. The hypothetical values are constructed in a similar manner, with the only difference being that the values of the transfers received are assumed to be equal to what the programme promised. The assumed value for in-kind transfer is 15kg of wheat and the assumed value for cash transfer is 100 BDT for each household[16]. Both real actual and hypothetical values are then converted to *per capita* values, so as to generate sufficient exogenous variation for identification[17].

It is important to note that by controlling for expenditures per capita, the regressions already control for the differences in the income effect on consumption and so to a certain extent for the differences in the value of the two types of transfers. Therefore, it will not be necessary to use the value regression specification unless we believe that the size of the compositional effects of the programmes is proportional to the size of their income effects.

[16] Some households will be assign with a hypothetical 20kg of wheat or 125BDT if the actual quantity/value reported in the questionnaire exceeds 15kg or 100BDT. These are what the FFE/PES programmes promised to the household who send more than one child to school.

[17] This is valid as long as household size is not endogenous to programme availability, which it is not.

In this case, the size of the transfers would matter. We will run the value specification for all three sets of outcomes for robustness check.

2.6.1.6 Estimating the Basic Specification

To simplify the exposition we rewrite Equation (2.5) in another form stacking all of the independent variables $InKind_{i,2,t}$, $Cash_{i,3,t}$, x_{it}, and the intercept together to form a single $(1 \times K)$ vector, where δ is a $(K \times 1)$ vector of corresponding coefficients. The number of parameters in X_{it} and δ are equal to $K = k + 3$, where k is the number of parameters in x_{it} and γ is as defined earlier[18].

$$y_{it} \; = \; X_{it}\delta \; + c_i \; + \; \mu_{it} \tag{2.6}$$

Equation (2.6) decomposes the original error term from Equation (2.5) in to $\epsilon_{it} \; = c_i \; + \; \mu_{it}$, where c_i is the unobserved household-specific time-invariant component, and μ_{it} is the unobserved household-specific time varying component. Equation (2.6) cannot be estimated directly, as the number of household fixed effects c_i increases with the number of sample size N. The estimation technique used to identify the estimates of interest therefore rest on how we deal with the household fixed effect.

The most basic estimation with panel data is the pooled ordinary least squares (OLS), which treats each observation as an independent unit ignoring the panel feature of the data. The pooled OLS regression will give consistent estimates of β_{InKind} and β_{Cash}, provided that the assumption $Cov[\epsilon_{it}\,, X_{it}] = 0$ holds as either $N \to \infty$ or $T \to \infty$. We estimate the pooled OLS regressions for all three outcomes set out previously. Given the assumption of exogenous regressors, we should obtain consistent estimates for statistical inference. In this particular case, the pooled OLS estimator $\hat{\beta}^{POLS}_{InKind}$ and $\hat{\beta}^{POLS}_{Cash}$ give consistent estimates of ATT^K and ATT^Cdiscussed in section 2.5 respectively[19].

2.6.1.7 Problems with the Basic Specification

We now look at the validity of the crucial assumptions that were made to obtain consistent estimators of β_{inkind} and β_{cash}. The assumption $Cov[\epsilon_{it}\,, X_{it}] \; = \; 0$ can be rewritten as

[18]More precisely $X_{it} = (1, InKind_{i,2,t}, Cash_{1,3,t}, x_{it})$ and $\delta' = (\alpha, \beta_{InKind}, \beta_{Cash}, \gamma)$

[19]In fact, β_{InKind} and β_{Cash} identify a variance weighted average $ATT^K$and ATT^C respectively as supposed to the unconditional $ATT^K$and ATT^C.

$Cov[c_i + \mu_{it}, X_{it}] = 0$. To obtain a consistent estimate, it is crucial that the included regressors (especially $InKind_{i,2,t}$ and $Cash_{i,3,t}$) are not correlated with either the household fixed effect or the household-specific error term.

However, due to the nature of both the FFE and PES programmes, there are significant differences between the characteristics of beneficiary and non-beneficiary households. We have controlled for some of them, such as income and household demographic structure, but it is possible that there are other characteristics that we could not control for, such as household preferences for food consumption, which are correlated with the beneficiary status and also the outcomes of interest. These characteristics are unobserved, but are arguably fixed over time for each household, and are captured by c_i in Equation (2.6). As a result, we have that $Cov[c_i, X_{it}] \neq 0$, and hence $Cov[\epsilon_{it}, X_{it}] \neq 0$. Failing to account for this correlation will lead to inconsistent estimation of the parameter of interest.

Although we may be able to control for the observable and time-invariant unobservable household characteristics, the assumption of strict exogeneity may still be violated. This occurs when the programme treatments are endogenous. The FFE programme has been intentionally placed in a relatively poorer region, both programmes also target poverty and suffer from issues of selection into treatment. In addition, the current data set also posed a problem with the determination of the actual beneficiary of the programme as discussed in Section 2.4, which results in measurement errors. For these reasons, it is very likely that the programme treatment variable will also be correlated with the unobserved error term μ_{it} and, hence, $Cov[\epsilon_{it}, X_{it}] \neq 0$. Ignoring this will also cause the estimates of interest to be inconsistent.

To deal with both problems, we designed a framework using both households fixed effects and instrumental variables. Due to the panel structure of the data set, it will be possible to control for the correlation between the unobserved household fixed effect and the regressor of interest. Using valid instruments will also solve the problem of endogeneity in programme treatment. Both methods are discussed in detail in the next two subsections.

2.6.2 Fixed Effect Estimation

As mentioned in the previous subsection, the correlation between unobserved household fixed effect and the regressor (especially the treatment) would lead to estimates of the programme

effect being inconsistent. To solve this problem, we make use of the panel structure of the data, which allows the transformation of the basic specification that will help control for this correlation.

Here we will move on from the basic specification in Section 2.6.1 and make use of the fixed effect (or "within") model. The model estimates a transformed version of the basic Equation (2.6). The transformed model removes the household fixed effect, which is thus treated as a nuisance parameter. The coefficient vector is then estimated by performing OLS on the transformed equation. The transformed equation can be written as follows:

$$y_{it} - \bar{y}_i = (X_{it} - \bar{X}_i)\delta + (\mu_{it} - \bar{\mu}_i) \tag{2.7}$$

for $i = 1, ..., N$ and $t = 1, ..., T$ where $\bar{y}_i = \sum_{t=1}^{T} y_{it}$, $\bar{X}_i = \sum_{t=1}^{T} X_{it}$, and $\bar{\mu}_i = \sum_{t=1}^{T} \mu_{it}$. As $\bar{c}_i = \sum_{t=1}^{T} c_i = c_i$, the fixed effect model transformation removed the individual household fixed effect parameter c_i. In addition to removing the individual household fixed effect, the transformation in Equation (2.7) will also remove all time-invariant observable characteristics. As a result, coefficients on time-invariant regressors such as sex or adult education cannot be estimated in contrast to the pooled OLS perform on the basic Equation (2.6).

The fixed effect estimator will be consistent as either $N \to \infty$ or $T \to \infty$ and $E[\mu_{it} - \bar{\mu}_i | X_{it} - \bar{X}_i] = 0$. Due to the presence of the averages $\bar{X}_i$ and $\bar{\mu}_i$, this condition is stronger than $E[\mu_{it}|X_{it}] = 0$ required for consistency of pooled OLS estimators. A sufficient condition for $E[\mu_{it} - \bar{\mu}_i | X_{it} - \bar{X}_i] = 0$ is the strong exogeneity condition that $E[\mu_{it}|X_{i1}, ..., X_{iT}] = 0$, which means that the present error term cannot be correlated with past, present or future regressors. This precludes a fixed effect (within) estimation with lagged endogenous variables as regressors.

Assuming that strong exogeneity holds, estimating Equation (2.7) will help solve the problem of endogeneity in the treatment. By controlling for individual household fixed effects we partly deal with endogeneity due to household self-selection into treatment as a result of time-invariant unobservable characteristics. Given that none of the households in this study moved between unions, the fixed effect model also deals with endogeneity that is a direct result of the non-random placement of the two programmes. The fix effect estimators $\hat{\beta}^{FE}_{InKind}$ and $\hat{\beta}^{FE}_{Cash}$ provide consistent estimates for ATT^K and ATT^C in this case.

Another possible method to control for household fixed effect is by using the first-

differences estimator. This is obtained by subtracting one period lagged equation from the original Equation (2.6). This first-differences equation is written as

$$y_{it} - y_{i,t-1} = (X_{it} - X_{i,t-1})\delta + (\mu_{it} - \mu_{i,t-1}) \tag{2.8}$$

The fixed effect c_i is removed by this transformation and an OLS estimation of Equation (2.8) yields the first-differences estimator. By estimating the equation in first-differences, there is no equation for the first time period, therefore we only have $N(T-1)$ observations in this regression. Consistency of the first-differences estimator requires that $E[\mu_{it} - \mu_{i,t-1} | X_{it} - X_{i,t-1}] = 0$. This is a stronger condition than $E[\mu_{it}|X_{it}] = 0$, but is weaker than the strong exogeneity condition needed for consistency of the within estimator discussed above. If this assumption holds, The first difference estimators $\hat{\beta}^{FD}_{InKind}$ and $\hat{\beta}^{FD}_{Cash}$ will also give consistent estimates for $ATT^K$and ATT^C. This approach is less efficient than the fixed effect estimator if μ_{it} is iid. However, it does have advantages when we allow for the endogenous regressor. This will be discussed in more detail in the next section.

2.6.3 Instrumental Variable Approach

The specification in the previous section helps control for unobserved household fixed effect. However, in the case where the included regressors (especially the treatment status regressor) are still endogenous, possibly due to measurement error or selection into treatment based on time-varying unobservables, the fixed effect estimator $\hat{\beta}^{FE}_{InKind}$ and $\hat{\beta}^{FE}_{Cash}$would still be inconsistent.

Participation in both programmes is affected by selection. There are two types of selection going on here, as discussed in the programme description in Section 2.3. First, an individual could self-select whether they want to participate in the programme by making a decision whether to enrol their child in school or not, and, if they do, whether to enrol in the school that participates in the programmes or not. Households may choose not to send their children to school for many reasons, including that the opportunity cost of enrolment may be too great as children are able to work otherwise. Households may choose to not enrol their children at a programme school, such a school could be of lower quality or too costly to travel to.

Secondly, if the programme is oversubscribed in some schools (i.e. over 40 percent of all

its students are eligible), the school committee is responsible for choosing who will get the benefit and who does not. As there were no strict guidelines for selection of eligible pupils in this case, selection decisions are at the discretion of the school committee, which could be influenced by wealthier and more powerful parents. Contaminated by these selection effects, this beneficiary status will be correlated with the error term and is no longer an ideal regressor.

Another problem is measurement error, which in this case refers to the problems in identifying actual beneficiaries of both the FFE and PES programmes due to inconsistencies within the questionnaire (as discussed in Section 2.4). In the value regression, it is also very likely that the self-reported value of transfer received are reported with errors as well. Even if we have taken care of the fixed effect, the inconsistency of $\hat{\beta}^{FE}_{InKind}$ and $\hat{\beta}^{FE}_{Cash}$ due to measurement error would still remains.

IVs can help solve these problems as a result of selection bias and measurement errors. In general, a valid instrument Z_{it} is an instrument that satisfies the following two conditions:

1. Instrument Relevance: $E(X_{it}\,, Z_{it}) \neq 0$
2. Instrument Exogeneity: $E(Z_{it}\,, \mu_{it}) = 0$

By using a valid instrument that satisfies both the relevance and the exogenous assumption, we can solve the endogeneity problem.

Choosing a valid instrument in a panel data setting is different to the same action in a cross-sectional setting and it depends heavily on the estimator used. If we use the first-differences estimator as in Equation (2.8), we could use an internal instrument in the form of lagged value or lagged difference. The difference-GMM and the System-GMM estimator, proposed by Arellano and Bond (1991) and Blundell and Bond (1998) respectively, make use of these instruments to correct for endogeneity in the first-differences equation. However, this requires that we have at least three years of data. Due to the constraint of the data set this is not possible here.

Even though we may have three years of data in total, we only have one year of data from when the FFE programme was active and only two years of data for the PES programme. Given that it is the impact of both programmes relative to each other that we are interested in, we need to have both as a separate regressor ($InKind_{i,2,t}$ and $Cash_{i,3,t}$). This means that we will not have a sufficient set of lagged values to act as instruments. In addition,

the first-differences coefficient of treatment effect would not allow the separation of in-kind and cash programme as we want, and all these reasons together are why we cannot use the first-differences specification.

The fixed effect (within) estimator is, therefore, the preferred estimator in this chapter. However, the nature of the transformation as in Equation (2.7) means that for an instrument Z_{it} to be a valid instrument it must satisfy the exogeneity assumption that $E(Z_{it}, \tilde{\mu}_{it}) = 0$, where $\tilde{\mu}_{it} = \mu_{it} - \bar{\mu}_i$. This requires stronger condition of strong exogeneity $E(Z_{is}, \mu_{it}) = 0 \, for \, all \, t$, which means that the instrument cannot be correlated with any past, present or future period error. This is all due to the presence of $\bar{\mu}_i$ in the transformed error term. Therefore, the within transformation can only be used when the instrument is strongly exogenous.

The equations below illustrates the IV approach used in this chapter:

$$\tilde{InKind}_{i,t} = (e_{i,FFE,t} \cdot b_{i,FFE,t})\delta_1 + (e_{i,PES,t} \cdot b_{i,PES,t})\delta_2 + \widetilde{x}_{i,t}\rho + \widetilde{\eta}_{it} \tag{2.9}$$

$$\tilde{Cash}_{i,t} = (e_{i,FFE,t} \cdot b_{i,FFE,t})\theta_1 + (e_{i,PES,t} \cdot b_{i,PES,t})\theta_2 + \widetilde{x}_{i,t}\pi + \widetilde{\nu}_{it} \tag{2.10}$$

$$\widetilde{y}_{it} = (\widehat{InKind}_{i,t})\beta_{InKind} + (\widehat{Cash}_{i,t})\beta_{Cash} + \widetilde{x}_{i,t}\gamma + \widetilde{\mu}_{it} \tag{2.11}$$

We suspect that there are only two endogenous regressors, those indicating the beneficiary (treatment) status in both FFE and PES programme. The two regressors could be endogenous as a result of measurement error or selection into treatment. Equation (2.9) and (2.10) are the first-stage regression of the two within transformed endogenous regressors $InKind_{i,t}$ and $Cash_{i,t}$ respectively. Equation (2.11) shows the fixed effect IV specification used to obtain a consistent estimate of β_{InKind} and β_{Cash} via two-stage least squares. Fitted value of $\widehat{InKind}_{i,t}$ and $\widehat{Cash}_{i,t}$, obtained from the first-stage equations, are substituted in to the within transformed version of Equation (2.5), where $\widetilde{y}_{it} = (y_{it} - \bar{y}_i)$ with $\tilde{x}_{i,t}$, and $\tilde{\mu}_{i,t}$ defined in the same way. The household fixed effect c_i has already been removed by the within transformation.

The instrument used in this fixed effect IV regression is the interaction between the eligibility status of the household in both programmes, regardless of whether the programme was

available or not ($e_{i,FFE,t}$ and $e_{i,PES,t}$) and an indicator of programme availability ($b_{i,FFE,t}$, $b_{i,PES,t}$). The creation of these variables was discussed in Section 2.4. The instruments ($e_{i,FFE,t}$·$b_{i,FFE,t}$) and ($e_{i,PES,t}$·$b_{i,PES,t}$) provide an alternative way to measure the beneficiaries of the programme. By strictly following the eligibility rule provided by the programme and identifying the eligibles, we managed to circumvent the selection problems previously mentioned, provided that the eligibility rule is exogenous. The focus is thus on measuring the impact of eligibility rather than participation (Morduch, 1998). Both instruments are definitely relevant as $E[InKind_{i,2,t}, (e_{i,FFE,t} \cdot b_{i,FFE,t})] \neq 0$ and $E[Cash_{i,3,t}, (e_{i,PES,t} \cdot b_{i,PES,t})] \neq 0$.

Given the poverty-related nature of the eligibility rule, at first it seems unlikely that the instruments based on it will be exogenous. However, given that we have controls for many relevant household observables and fixed unobservable characteristics, the eligibility variable should now be exogenous from the transformed error term. We are making the assumption that both of the instruments are strongly exogenous conditional on $\tilde{x}_{it}$, i.e. $E[(e_{i,j,t} \cdot b_{i,j,t}), \tilde{\mu}_{it}|\tilde{x}_{it}] = 0$ where $j = FFE\, or\, PES$. These assumptions corresponds to those made in subsection 2.5.2.

There will only be a reason to believe that the instrument is not exogenous if households respond to the availability of both programmes and try to change their eligibility status as a result. Recall the four and five eligibility rules of the FFE and PES programmes respectively (see Section 2.3). It is very unlikely that households would or could change their land holding or become a female-headed household (through divorce, or death of husband) in order to be eligible for a welfare programme. There might be more concerns that heads of households may change their occupations, or manage to get the official to misreport their occupation in order to change the eligibility status. If households do manipulate their eligibility status in response to the programme, this would lead to a potential problem, but only if the selectivity were linked to factors affecting outcomes of interest. However, given the small size of the transfer (which only represents around 5% of household expenditure on average), it is very unlikely that this will happen for enough people to cause the instrument to become endogenous and lose its validity entirely.

If all the assumptions stated above hold together with the monotonicity assumption, the fixed effect IV estimator, $\hat{\beta}^{FE-IV}_{InKind}$ and $\hat{\beta}^{FE-IV}_{Cash}$, will give consistent estimates of $LATE^K$ and $LATE^C$ discussed in section 2.5. Although, these parameters may be less externally valid

than ATT^K and ATT^C, given by the pooled OLS and fixed effect estimator (if the required assumptions hold), they are the parameters that can be estimated with more credibility and more plausible assumptions needed for consistency than both the pooled OLS ($\hat{\beta}^{POLS}_{InKind}$ and $\hat{\beta}^{POLS}_{Cash}$) and the fixed effect estimator ($\hat{\beta}^{FE}_{InKind}$ and $\hat{\beta}^{FE}_{Cash}$). This makes the fixed effect IV estimator the most internally valid one. Here, due to data limitations, we sacrifice some degrees of external validity in favour of higher degree of internal validity.

A household fixed effect IV regression will therefore be our preferred specification. We will test the result of this specification with the pooled OLS of the basic specification, regression with just the fixed effect estimator, and regression with just the IV estimator.

2.6.4 Standard Errors

The default assumption for all the estimators above is that the random error term μ_{it} is independent and identically distributed, with a mean of zero and a homoskedastic variance $\mu_{it} \sim iid\left[0\,,\,\sigma^2_{\mu}\right]$. Under this assumption, all the estimators discussed in the previous section will provide correct standard errors and, provided that we have consistent estimates, will allow us to conduct correct statistical inferences. It is reasonable to assume independence over individual i, but in a panel data setting the errors face two potential problems; heteroskedasticity and serial correlation .

Heteroskedasticity occurs when the variance is not constant across the observations; bias caused by heteroskedasticity could be in any direction. Using the White heteroskedastic consistent estimator allows for correction of certain forms of heteroskedasticity, but it still fails to take account of the serial correlation. Serial correlation occurs when the errors are correlated over t for given individual i, $Cov\left[\mu_{i,t},\,\mu_{i,s}\right] > 0 \;for\; t \neq s$. Ignoring this serial correlation can lead to greatly underestimated standard errors, which lead in turn to over-estimated t-statistics. Controlling for fixed individual-specific effects can reduce serial correlation in the composite error term $\epsilon_{i,t}$, but it may not be completely eliminated.

Cluster-robust standard errors as proposed by Arellano (1987) are used for all the estimation set out in previous sections to account for both heteroskedasticity and serial correlation. The individual household has been chosen as the clustering unit, which is the appropriate choice here, given that we are looking for variations at the household level.

2.6.5 Attrition

As set out in Section 2.4, 25 households were lost from an original sample of 480 household due to attrition. This is an attrition rate of approximately 5.2 percent, which is relatively low in comparison to similar studies in low-income countries. However, non-random attrition, where households with certain characteristics which are correlated with the outcome of interest systematically leave the survey, will cause bias in the estimates.

To test for randomness of attrition, we run the attrition probit (Fitzgerald et al., 1998), in which the dependent variables take the value of one for households who had left the sample any time after the first wave and zero otherwise. Explanatory variables are baseline values for all variables that are believed to have affected the outcome of interest. Significance in some of the regressors may suggest that attrition is non-random.

The attrition probit result is presented in Table 2.A.3 in Appendix 2.A.3. Age of household head is the only important variable that explains attrition, with households with younger head more likely to leave the sample. This could be for many reasons, including that the younger heads are more likely to migrate to find a job. All other household characteristics strongly related to food consumption and health status of young children such as household size, household composition, level of education and income (proxy by asset and expenditure) are not significant predictors of attrition.

There are methods to correct for attrition bias depending on the assumption regarding the nature of attrition. If attrition occurs as a result of selection on observables, we can use the method of inverse probability weight (IPW) (Wooldridge, 2002) to correct for it. If attrition occurs as a result of selection on unobservables, a less restrictive assumption, an exclusion restriction which involves an exogenous instrument that is correlated with attrition but not the outcome of interest is needed. A common instrument for this is normally the interviewer's characteristics. Given that this information is not available, it is not possible to perform a correction for attrition based on selection on unobservables. IPW makes strong assumptions, which may create more noise due to the restrictive assumptions imposed without much gain. Given a small attrition rate and the evidence from the attrition probit, we concluded that attrition bias is small and it was therefore ignored.

2.7 Empirical Results

This section presents the results of our estimations following the strategy set out in Section 2.6. We are interested in three sets of outcomes; calorie consumption by household, expenditure share of various food and non-food items, and health outcomes. The difference in the effect of in-kind versus cash transfer on these outcomes will be the major focus of this section.

2.7.1 Effect on Consumption

Household calorie consumption is the starting point of comparison. This is because in-kind transfers in this case are given out as food grains. Therefore, if households were to face a constraint on their consumption as a result of the in-kind programme, it would most likely show in their calorie consumption. As the theoretical model in Section 2.2 predicted, households who face extra-marginal and binding in-kind transfers will end up having to consume more of that transfer than they would have liked to do under an equal-valued cash transfer.

Table 2.9 presents the impact of the FFE and PES programmes on household calorie consumption. The dependent variable in all specifications in Table 2.9 is the daily *per capita* calorie consumption by household measured in kilocalories. In regressions (1) to (4) we do not scale the impacts by the value of subsidies under the FFE and PES programmes. We relax this and allow for the difference in the size of transfers between the two programmes in the value regressions (5) and (6) to check for robustness of the results.

Results for pooled OLS regression on the basic specification set out in Equation (2.5) is presented in Table 2.9, Column (1). Households under the in-kind programme consume significantly more calories than households who received no benefit, the omitted category in this case. The coefficient for the cash programme is negative, implying that a beneficiary household of the cash programme consumes less food than a household who received no benefit. However, the coefficient is not statistically significant. $\hat{\beta}^{POLS}_{InKind}$ and $\hat{\beta}^{POLS}_{Cash}$ are the naive estimators of programmes effect which will give ATT^K and ATT^Conly when both can be estimated consistently. Simple t-test of equality between in-kind and cash indicates that the two programmes are statistically different in their effects on household calorie consumption. This is the first evidence that there may be a constraint on the consumption behaviour

Table 2.9: Daily Per Capita Calorie Consumption Regression

	(1)	(2)	(3)	(4)	(5)	(6)
Calorie per capita	Pooled	FE	IV	FE-IV	FE-IV	FE-IV
In-Kind	**97.167****	**16.260**	**129.319**	**-33.502**		
	(47.989)	(57.854)	(113.660)	(188.356)		
Cash	**-20.043**	**-98.776****	**-671.608***	**-670.306**		
	(43.659)	(42.702)	(343.659)	(451.212)		
PC value FFE					**-0.839**	
					(4.304)	
PC value PES					**-25.628**	
					(15.911)	
PC hypo.value FFE						**-0.517**
						(3.694)
PC hypo.value PES						**-23.511**
						(14.663)
Log Household Size	-225.393***	-365.654	-164.752*	-330.906	-416.712*	-430.393*
	(84.602)	(256.700)	(92.923)	(255.790)	(248.216)	(248.979)
Log PC Expenditure	1143.939***	1297.789***	1098.321***	1296.698***	1290.159***	1294.261***
	(88.921)	(109.822)	(93.177)	(108.629)	(107.665)	(107.135)
Children 0-4	-561.798***	315.238	-482.986**	287.446	331.211	347.770
	(210.646)	(679.870)	(222.298)	(684.649)	(667.422)	(666.842)
Children 5-9	-374.999**	275.364	-163.530	315.079	306.426	312.707
	(177.774)	(615.785)	(214.850)	(615.179)	(602.724)	(603.592)
Children 10-14	76.111	596.396	302.439	793.258	727.365	781.018
	(155.221)	(593.426)	(200.649)	(593.565)	(580.351)	(584.511)
Adult 15-54	214.616*	607.628	211.106	694.952	655.768	682.748
	(127.175)	(536.337)	(129.078)	(532.905)	(524.715)	(524.706)
Wheat Price	-16.712*	2.966	-17.646*	3.111	-0.232	0.913
	(9.415)	(10.817)	(10.562)	(11.815)	(11.348)	(11.294)
Rice Price	-87.169***	-40.415	-88.340***	-51.165	-48.035	-45.198
	(13.694)	(26.670)	(14.798)	(32.403)	(29.884)	(29.462)
year==2003	254.699***	116.582	573.926***	358.673*	321.674**	306.252**
	(57.977)	(78.407)	(166.831)	(195.878)	(161.507)	(153.566)
year==2006	396.725***	-135.447	699.663***	109.614	36.669	30.388
	(149.327)	(215.468)	(225.045)	(319.521)	(260.977)	(258.193)
Daylabour	81.282**	33.758	130.161***	95.723	82.099	76.153
	(35.234)	(49.021)	(47.333)	(76.154)	(66.199)	(65.088)
Low Income Occupation	-245.877**	-122.693	-225.147**	-126.510	-111.309	-103.865
	(106.297)	(146.039)	(109.573)	(158.043)	(151.048)	(152.166)
Female-headed Household	-42.764	-13.371	-27.457	58.271	56.678	46.615
	(58.962)	(89.713)	(64.156)	(110.783)	(106.460)	(104.161)
Child Enrolment	-106.249	68.162	307.846	476.790	390.446	398.569
	(145.154)	(158.185)	(294.174)	(411.049)	(336.265)	(341.562)
Constant	-3750.053***	-5757.510***	-3748.015***			
	(570.154)	(879.207)	(588.195)			
Observations	1365	1365	1365	1365	1365	1365
InKind=Cash	**0.049****	**0.112**	**0.008*****	**0.059***	**0.062***	**0.065***

Robust-Clustered standard error in parentheses; main statistics for InKind=Cash are the p-values
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of households when transfers are given in-kind.

The coefficients on other regressors included as control variables have reasonable interpretations in the pooled OLS setting. An increase in household size results in lower per capita calorie consumption. The wealthier the household (measured by higher real per capita expenditure), the higher their calorie consumption (holding other factors constant). Both household size and income are significant in explaining calorie consumption. The composition of households also plays a major role, with children eating less calories while young and middle age adults eating more calories relative to the omitted categories (i.e. adults of the age over 55 years). The rice price is also significant in explaining calorie consumption, which mainly stems from the fact that rice is a staple and makes up a major part of the household diet (as shown in Table 2.6). The indicator variables for 2003 and 2006 also show positive signs, suggesting that households in this time period tend to consume more calories than households in the year 2000.

However, as previously discussed, there are numerous problems with the endogeneity of the regressors in the basic specification estimated by pooled OLS. These include endogeneity as a result of failing to control for unobserved household fixed effects, selection into treatment and measurement errors, which could all cause inconsistencies in the pooled OLS estimates.

As explored earlier, these problems can be corrected for using the fixed effect (within) estimator that exploits the panel structure of the data and an IV approach. Both methods are attempted separately and then together as the FE-IV estimator. Regression (2) in Table 2.9 shows the result of the fixed effect estimation as shown in Equation (2.7) in Section 2.6. In the fixed effect estimation, coefficients on both in-kind and cash still have the same sign as the pooled OLS estimates, but both decrease in magnitude. This is consistent with the selection into treatment based on household unobservable characteristics, which could result in an upward bias in the pooled OLS estimates. The coefficient for the in-kind programme also loses significance, while the coefficient for the cash programme is now more negative and statistically significant, suggesting that households under the cash programme consume less calories than non-beneficiaries. However, the t-test for difference between in-kind and cash transfer is no longer significant. This suggests that, when we allow for the household fixed effect, the apparent difference between the two programmes disappears. It could be that eligible households that select themselves for the food transfer are those who have

a preference for food consumption and are therefore likely to benefit the most from food transfers. $\hat{\beta}^{FE}_{InKind}$ and $\hat{\beta}^{FE}_{Cash}$ here are consistent estimators of ATT^K and ATT^Crespectively only when selection bias are caused by observable $X_{i,t}$ and time-invariant unobservables c_i only which, as argued earlier, is not the case in this setting.

The IV specification in regression (3) provides different results to the fixed effects specification. In comparison to the pooled OLS regression (1), the coefficients of both the in-kind and cash programmes increase in magnitude, but in an opposite direction. The coefficient for the in-kind programme is more positive while the coefficient of the cash programme is more negative with both coefficients becoming statistically significant. The t-test for differences between the in-kind and cash programmes is significant once more. The standard error on both estimates increased, which is a usual characteristic of the IV specification. $\hat{\beta}^{IV}_{InKind}$ and $\hat{\beta}^{IV}_{Cash}$ are estimators of $LATE^K$ and $LATE^C$. Both are only consistent if the instruments are relevant and exogenous[20]. Applying the Kleibergen and Paap rk statistic (a generalisation of the Cragg-Donald Wald statistic) to the case where the errors are not iid, and comparing it to the Stock-Yogo critical value (Stock and Yogo, 2005), we were able to reject the null hypothesis that the instrument is weak. As the IV regression in this case is exactly-identified, it is not possible to perform an over-identifying restriction test to look for instrument exogeneity[21].

As discussed earlier, it is possible for the instrument to fail the exogeneity assumption without controlling for the unobserved household fixed effect. Despite the instrument being calculated from the exogenous eligibility rule that households have no control over, the poverty-targeting nature of the programme means that these eligibility criteria will be correlated with many of the poor household characteristics. In this case, the instruments not only act through the endogenous variable but also directly affect the outcome as well. Depending on the nature of the correlation of the instruments with the unobserved household fixed effect, the IV estimates could be positively or negatively biased.

The FE-IV is used to solve the above problem. By controlling for all the observable and unobservable fixed effects, the exogeneity assumption of the IV can arguably be satisfied. The FE-IV, which is the preferred specification, should give consistent estimates of both

[20]In addition, we also need the monotonicity assumption for the identification of $LATE$ as well.

[21]The critical values due to Stock and Yogo were only originally computed for using with the Cragg-Donald Wald statistics where the error are assumed to be iid. However, there are no other critical values computed for non-iid error as is the case here. Using this critical value however, is still informative and is better than nothing at all but the test cannot be taken to be 100% accurate.

in-kind and cash programme variables. $\hat{\beta}_{InKind}^{FE-IV}$ and $\hat{\beta}_{Cash}^{FE-IV}$ here are consistent estimators of $LATE^K$ and $LATE^C$. However, due to the fact that the fixed effect estimation requires a transformation of the original equation, it could potentially reduce the first stage correlation between the instrument and the transformed regressor, resulting in a problem of weak instruments. However, the Kleibergen and Paap rk statistics reject the null hypothesis that the instrument is weak. From regression (4) in Table 2.9, the coefficient of the in-kind programme decreases in magnitude, and has become negative though it is not statistically different from zero. The coefficient on the cash transfer is more negative than with the pooled OLS estimate and also is not statistically significant. More importantly, with the preferred FE-IV estimation, the t-test still only found statistically marginal difference between the two programmes.

As mentioned in Section 2.6, the size of transfers under both programmes are not equal, with the cash programme receiving both less stipend and also facing the effect of rising inflation eroding the value of the transfer. If the compositional effects of the programmes were proportional to the size of the transfers, our comparison of the two programmes' estimates in regression (4) would be incorrect. To solve this, we use the transfer value of each programme as regressors instead of the indicator variable used in regressions (1) to (4). Regression (5) uses the actual values received by household, while regression (6) uses the hypothetical value that the household should receive based on the promise of the programmes. Both regressions were estimated using the FE-IV approach and both provide similar results to regression (4). The positive coefficient of the FFE transfer means that, as the value of the in-kind transfer increases, beneficiaries will consume more calories, while the negative coefficient for the PES transfer suggests that beneficiaries will spend less on food calories as the value of their cash transfer increases. However, neither coefficient is statistically different from zero, and again the t-test for the test of difference between the two programmes fails to reject the null hypothesis that there are no difference in both estimates at 5% significance level.

The first-stage results for the IV regression (3) to (6) discussed above are shown in Table 2.A.4 in Appendix 2.A.4. The in-kind and cash instruments both exhibit statistically significant positive relationship with their respective endogenous variables in all the first-stage regressions. This shows that both instruments are relevant. However, the main caveat is that the both instruments also exhibit statistically significant negative relationship with

the beneficiary status of alternative programmes (in-kind instrument with PES beneficiary status and cash instrument with FFE beneficiary status). The negative relationship arises mainly as a result of the fact that a household can never be a beneficiary of the other programme within the same period as the FFE and PES take place in different period. In this case, the binary in-kind instrument, which represent the FFE beneficiary status will only take positive values in the year 2000 where PES beneficiary has yet to exist, and will always be zero when PES programme starts in the year 2003 and 2006 when FFE programme has already stopped.

The negative relationship between the in-kind instrument is stronger and remains statistically significant in all first stage regressions as a result of the within transformation introducing more negative bias to the in-kind instrument relationship with cash beneficiary status than the other way around. While the mechanical nature of this negative correlation could result in a problem with the FE-IV estimates of the cash programme, the impact should be insubstantial as the number of hypothetical cash beneficiary (cash instrument = 1) will greatly outweigh the number of hypothetical in-kind beneficiary when computing the fitted value in the first stage. In addition, the F statistics in all first-stage regressions exceed 10, indicating that the instruments are not weak.

Apart from the pooled OLS and IV regression, which both suffer from endogeneity problems, there is no statistical difference between the level of daily *per capita* consumption of calories among beneficiaries of the in-kind and cash transfer programmes at 5% significance level. Even when the difference in the value is taken in to account, both programmes still have statistically the same effect on food consumption. The lack of difference between the two programmes on calorie consumption is robust to different specifications.

There are three possible explanations for these results. The first is the issue mentioned in Section 2.6.1.2 regarding income and the compositional effect of the transfer. By only looking at the compositional effect of the transfer, we can conclude that both in-kind and cash programmes do not have any effect on the calorie composition of beneficiary compared to non-beneficiary households, and that the compositional change between the two programmes was not strongly significant either. This means that, while either programme could have led to increases in overall consumption of calories as a result of an income effect, neither have led to compositional changes in spending that shift calorie consumption enough to

be identified as strongly statistically significant from non-beneficiary households and each other. This explanation helps to reconcile the small and even negative coefficient observed for households under both programmes, but it still fails to explain why there is no difference in effect between the two types of transfer from each other.

A second explanation for the lack of difference between the two programmes both from non-beneficiaries and each other would be the small size of the transfer given out in both programmes. On average, the in-kind and cash transfers are only approximately 5% of household total expenditure. Though the programme may have an impact, and outcomes under cash and in-kind transfers could be significantly different, we were unable to observe such an outcome, as the initial transfers to households were too small to have a noticeable impact. Lastly, we cannot rule out the possibility that there is really no difference between the two programmes, and, hence the two types of transfer. Nevertheless, an inspection of the direction of the coefficient suggests that in-kind beneficiary households shift more of their consumption towards food, hence the positive coefficient in comparison to the cash programme, where households shift more of their consumption away from food resulting in the observed negative sign.

2.7.2 Effect on Expenditure Share

Evidence from the food calories regression suggests that, once we controlled for the household-specific fixed effect and the endogeneity of beneficiaries' status, neither the FFE nor PES programme beneficiaries consume more calories than those who received no transfer or in comparison to the other programme net of income effects of the programmes.

Despite the result above, the direction and magnitude of the coefficients suggest that the in-kind food transfer did lead beneficiary households to shift more of their consumption towards food than under the cash programme. It is interesting to investigate this further in order to determine where the opposing shift seen in two the programmes may have originated from and, if there is a constraint, why it did not show up more strongly in the overall food calorie consumption. This is done by looking deeper into various expenditures categories.

Table 2.10 shows the result of the regression of the impact of both the in-kind and cash programmes on various expenditure categories. The specification in all of the regressions in Table 2.10 is the preferred FE-IV regression. Similar to before, $\hat{\beta}_{InKind}^{FE-IV}$ and $\hat{\beta}_{Cash}^{FE-IV}$

Table 2.10: Expenditure Shares Regression

	(1) Wheat	(2) Rice	(3) Cereal	(4) Meat & Fish	(5) Fruit & Veg	(6) Food	(7) Cloth	(8) Education	(9) Kid	(10) Smoke
In-Kind	**0.062*****	**-0.039**	**0.023**	**0.012**	**0.010**	**0.046**	**0.010**	**-0.000**	**0.019**	**-0.008**
	(0.011)	(0.029)	(0.029)	(0.033)	(0.022)	(0.040)	(0.021)	(0.011)	(0.013)	(0.008)
Cash	**0.007**	**-0.056**	**-0.058**	**0.030**	**-0.021**	**-0.019**	**0.075**	**0.026**	**0.059***	**-0.036****
	(0.021)	(0.066)	(0.065)	(0.075)	(0.049)	(0.085)	(0.046)	(0.025)	(0.031)	(0.017)
Log Household Size	0.001	-0.007	0.005	0.036**	0.007	0.026	0.019**	-0.006	0.014**	-0.000
	(0.004)	(0.013)	(0.014)	(0.015)	(0.010)	(0.018)	(0.009)	(0.007)	(0.006)	(0.003)
Log PC Expenditure	-0.004*	-0.139***	-0.132***	0.106***	-0.003	-0.005	0.007	-0.004	0.002	-0.006***
	(0.003)	(0.008)	(0.008)	(0.013)	(0.008)	(0.017)	(0.009)	(0.003)	(0.004)	(0.002)
Children 0-4	0.006	0.014	-0.009	0.006	-0.030	0.001	-0.021	-0.003	0.009	-0.003
	(0.013)	(0.049)	(0.050)	(0.050)	(0.036)	(0.058)	(0.032)	(0.019)	(0.020)	(0.011)
Children 5-9	0.001	0.025	-0.010	0.074	-0.053	0.044	-0.018	-0.006	0.014	0.006
	(0.012)	(0.047)	(0.047)	(0.050)	(0.034)	(0.061)	(0.032)	(0.018)	(0.021)	(0.010)
Children 10-14	-0.004	0.100**	0.069	0.021	-0.028	0.051	-0.052*	0.028	-0.002	0.014
	(0.011)	(0.044)	(0.045)	(0.050)	(0.032)	(0.054)	(0.030)	(0.018)	(0.020)	(0.010)
Adult 15-54	-0.001	0.110***	0.079**	-0.020	-0.031	0.022	-0.035*	0.026	-0.013	0.015**
	(0.008)	(0.035)	(0.036)	(0.039)	(0.026)	(0.041)	(0.021)	(0.016)	(0.014)	(0.007)
Wheat Price	-0.001	-0.001	-0.000	0.002	0.001	0.006***	-0.000	-0.001	-0.000	-0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Rice Price	0.003**	0.001	0.005	-0.010**	0.010***	0.012**	-0.005*	0.003*	-0.001	-0.001
	(0.002)	(0.005)	(0.004)	(0.005)	(0.003)	(0.006)	(0.003)	(0.002)	(0.002)	(0.001)
year==2003	0.000	0.011	0.016	-0.027	-0.008	-0.047	-0.028	-0.017	-0.012	0.015**
	(0.009)	(0.027)	(0.026)	(0.029)	(0.021)	(0.035)	(0.021)	(0.011)	(0.013)	(0.007)
year==2006	-0.009	0.007	-0.021	-0.012	-0.058*	-0.152***	0.003	-0.013	0.001	0.021**
	(0.016)	(0.044)	(0.043)	(0.047)	(0.034)	(0.055)	(0.032)	(0.018)	(0.020)	(0.011)
Daylabour	0.001	0.022**	0.023**	-0.000	-0.006	0.002	-0.011	-0.003	-0.014***	-0.000
	(0.004)	(0.011)	(0.011)	(0.012)	(0.008)	(0.013)	(0.007)	(0.004)	(0.005)	(0.003)
Low Income Occupation	-0.018*	0.015	-0.010	-0.011	-0.009	-0.046**	0.016	-0.002	0.011	0.007
	(0.010)	(0.018)	(0.022)	(0.019)	(0.014)	(0.022)	(0.013)	(0.007)	(0.009)	(0.006)
Female-headed Household	-0.000	0.013	0.020	-0.022	0.013	-0.006	0.008	0.003	0.001	-0.006
	(0.006)	(0.015)	(0.015)	(0.017)	(0.013)	(0.018)	(0.011)	(0.007)	(0.008)	(0.004)
Child Enrolment	-0.003	0.066	0.077	-0.091	0.031	-0.017	-0.057	-0.012	-0.041	0.031**
	(0.018)	(0.057)	(0.057)	(0.067)	(0.046)	(0.075)	(0.040)	(0.020)	(0.027)	(0.014)
Observations	1365	1365	1365	1365	1365	1365	1365	1365	1365	1365
InKind=Cash	**0.0031*****	**0.737**	**0.094***	**0.737**	**0.401**	**0.323**	**0.080***	**0.136**	**0.096***	**0.0184****

Robust-Clustered standard error in parentheses; main statistics for InKind=Cash are the p-values

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

here are consistent estimators of $LATE^K$ and $LATE^C$. The only difference is that the dependent variable is now the expenditure on various food, non-food items/categories, all computed as a share of household total expenditure. Though we do not report the value regressions, the results in this section are robust to using the value specification as with the calorie regressions earlier.

Only a few of the more relevant items and categories are presented in Table 2.10. Wheat and rice share were selected, as the in-kind food transfers were given out in terms of these two items. Cereals, meat, fish, fruit and vegetables are the other food categories presented in Table 2.10; these are not an exhaustive list of categories but represent around 80% of all food expenditure and are sufficient for the analysis. The total food share is also presented, along with other non-food categories. Selected categories are clothing expenditure share for the household (adults and children), expenditure share on education, expenditure share on cigarette, and finally all expenditures on children as a share of total expenditure. It must be noted that expenditures on children's clothing are also a part of expenditure on children (though expenditures on education are not), but we chose to present the overall spending on children in addition to the clothing expenditure so as to illustrate a possible intra-household allocation story.

From Table 2.10 it is evident that the FFE programme leads to more consumption of wheat than the PES programme, with beneficiaries who received the in-kind food transfer ending up with a wheat share 6.2% higher than households who received no benefits. The wheat share for beneficiaries of the cash programme is not statistically significant from the non-beneficiary households. The t-test for the difference between the two programmes is highly significant. The finding reflects the fact that the wheat transfer was extra-marginal and binding, and beneficiaries of the in-kind programme are constrained to consume more wheat than they would have chosen to do under an equal-valued cash transfer. This provides clear evidence of a constraint on the composition of household food expenditure.

The rice share is lower for beneficiary households under both programmes when compared to those households that received no benefits. Both coefficients were not statistically significant and the t-test of the difference found that there are no differences between programmes. This either indicates that most in-kind rice transfers were infra-marginal or the fact that households in this round of the survey received most of their benefits in wheat rather than

rice. The coefficient for the FFE programme provides an interesting story. Regression (2) confirms the findings in the survey that most of the in-kind food transfer was given out in the form of wheat and that, in response to the constraint that households faced in higher wheat consumption, they substitute away from rice consumption. This type of substitution is predicted by the theory of rationing (Neary and Roberts, 1980; Deaton, 1981) mentioned earlier is Section 2.2.

This substitution story is confirmed in regression (3), when we look at expenditure on cereal share. Both rice and wheat fall under the cereal category and in fact make up around 92% of households' cereal consumption on average. Though the coefficient of the in-kind programme is positive, suggesting that in-kind beneficiary households did consume more cereal than non-beneficiary households, the increase was not statistically significant. This indicates that beneficiary households were able to substitute away from rice in favour of wheat to leave themselves less constrained overall. This explains why we cannot find the impact of the in-kind programme on calorie consumption despite the fact that the wheat transfer did constrain people to consume more wheat. Households under the cash programme choose to consume less cereal than non-beneficiary households but again the difference between the two programmes is only marginally significant, which strengthens our argument that in-kind beneficiary households were able to adjust their behaviour and substitute effectively. It is important to note that, while the constraint in wheat consumption caused by the FFE programme is clear, wheat on average represents only a small part of households' total expenditure. Consequently, it could easily be offset by household-level adjustments intended to decrease other parts of their food expenditure in order to free themselves from the constraint.

The next two regressions show the result of both programmes' impact on consumption of meat and fish and fruit and vegetables, which are the other major parts of households' food consumption. The findings on both categories are similar in that none of the coefficients of the in-kind and the cash programmes are individually significant. The t-test for the difference between the in-kind and the cash transfer programmes fails to reject the null hypothesis that there are no differences between the two transfers for both categories of food. Given that the categories of food presented in this table made up the majority of households' food consumption, we can safely conclude that, while they may be evidence of some constraints

as a result of the FFE programme, households were able to shift their consumption among food items around so that the overall constraint is minimal.

Regression (5) shows the results for food share. Given the substitution argument earlier, it is not surprising that the coefficient of the in-kind variable is statistically insignificant. The effect of the cash programme on overall food share is also not statistically significant. However, the size of the coefficient suggests that cash beneficiary households have around a 2% lower food share than non-beneficiaries, compared to 4.6% higher food share for the in-kind beneficiary households. Though the t-test of differences between the two programmes cannot reject the null hypothesis that they are the same, there are some indications that households under the cash programme choose to spend less on food than households under the in-kind transfer programme. This provides some evidence that, once free from the constraint of the in-kind transfer, households choose to spend their money differently, allocating more of their funds towards non-food items.

If the cash-receiving households were free from the constraint faced by the in-kind households, and choose to spend less on food, it is important for policymakers to know what they choose to spend the extra money on. Paternalistic fears would be confirmed if these households choose to spend the extra cash on goods deemed to be undesirable by society, such as cigarettes or alcohol. However, if households in fact spend their extra money positively on goods that are deemed beneficial (such as children's education), then paternalistic fears would be unjustified.

Another important issue in the planning of redistribution programmes is the issue concerning the indirect targeting of a specific group of the population, such as young children. Transfers that are aimed at children are often given through parents and as a result it is possible that, if given in cash, such transfers could be diverted by parents to spend on things that do not improve children's welfare. As a result, a paternalistic government would provide the benefits in-kind by making the transfer conditional or by giving out the benefits in in-kind goods that benefit the children directly, such as school uniforms, school lunch or food. Both the FFE and PES programmes are conditional on children's education and, hence, aim to improve the welfare of young children in Bangladesh. It is therefore also important to determine whether the benefits from both programmes are being sufficiently passed on to young children whose welfare, apart from school enrolment, is only indirectly targeted.

Regression (7) looks at the impact of both transfer programmes on households' spending on clothing. Both In-kind and cash beneficiary households are found to have spent more on clothes than non-beneficiaries. However, the t-test of the difference between the two coefficients rejects the null hypothesis and shows that there are significant differences between the two programmes at the 10% level with cash programme households spending larger share of their income on cloth than in-kind households. Regression (8) looks at the impact of both transfer programmes on households' spending on children's education. Beneficiary households from the in-kind programme spend no more on children education than household who received no benefits at all. On the contrary, cash beneficiary households spend 2% more on children education than non-beneficiary households though this is not statistically significant. The t-test of the difference between the two coefficients found no significance different between the two programmes.

It is possible to drawn from these results a conclusion that cash transfer beneficiary households who were not under any constraints were able to spend their extra money directly on useful personal items, such as clothes. Households under in-kind programme however, are already facing constrain with in-kind goods (wheat), and can only manage to substitute between food categories to a certain extent. As a result, they were not able to generate enough leftover cash to spend on non-food items as much as cash beneficiary households can.

Regression (10) shows that cash beneficiary households spend statistically less on cigarettes than households that received no benefits, while in-kind beneficiary households spent no differently on smoking than non-beneficiary households. The t-test found that there is a difference between programmes. It is evident that cash beneficiary households spend less on smoking than in-kind beneficiary households. Regressions (7), (8) and (10) tell a positive story that not only does the cash beneficiary household chose to spend their extra money drawn away from food expenditure positively, they are not spending the extra cash on vices such as cigarette. This should ease the paternalistic fear that poor households make bad decisions once given the option of spending freely.

In addition, regression (9) also provides evidence that counters paternalistic fears concerning transfer programmes in general; the indirect targeting of children. Both in-kind and cash programme beneficiaries were found to have spent more on children than non-beneficiary

households. This finding provides evidence for the "flypaper" effect (Jacoby, 2002), where if a benefit is aimed at a specific member of the household, the benefits seem to stick to that member even though the household could have reallocated the benefits across all members. In this case, as the transfers were only obtainable as a result of children attending school, parents might view the benefits received as a transfer to their children and therefore ended up spending more on their children. The effect is different between the two programmes with cash beneficiary households able to spend more on children, but only marginally significant. This should eliminate the paternalistic fear regarding indirect transfers to children possibly being reallocated away under cash transfers, though we must remember that even though the mode of the transfer is different, both are conditional transfer programmes and are both technically in-kind, so we cannot draw the conclusion here too strongly.

In-kind transfer beneficiaries were able to adjust their behaviour to a certain extent and manage to substitute effectively in order to keep their food consumption at the same level as those who did not receive the benefits, but were still consuming more than they would have liked to do under equal-valued cash transfers. It is a puzzle as to why these households did not substitute enough to allow them to spend the extra money saved from not spending on food in a similar manner as those who did this under the cash transfer programme.

The answer must lie in the behavioural response by households to the different types of transfer. It is possible that when transfers are given in-kind and the transaction cost for reselling is high, households may substitute just enough so that they consume roughly about the same amount of food as before but with a change in the composition of consumption. The extra money saved is then used to spend on what the household wants and, in this case, the shift shows up in expenditure on children and clothes. However, under the cash transfer programme, households are able to spend the transfer freely and are not constrained to consume the food transfer. It is possible that, once money is being spent on non-food items, the extra money from the transfer acts as an extra source of cash liquidity that was combined with current income to purchase various non-food items that households weren't able to afford before.

2.7.3 Effect on Health Status

In the last section, we established that in-kind transfer lead to a constraint in the behaviour of the recipient. Households under the cash transfer were able to spend more on clothes and children's non-food consumption. The constraint caused in-kind households to spend more on food as a total share of expenditure than cash households, although these differences did not cause in-kind households to consume significantly more calories than cash households. In this section, we now look at how these constraints in the behaviour may have an effect on household welfare.

One measure of household welfare is the health status of the members of households. We look specifically at two measures of health status of young children between the ages 0 to 12 years old; proportion of stunted and underweight children within a household. Anthropometric measures were taken only in the 2000 and 2006 rounds, so only data from these two rounds is used in the regressions. In addition, we can only investigate these effects for households who have children age 0 to 12 in both rounds of the survey. The overall sample sizes are therefore reduced from 1365 to 632 observations. Again, the specification in all of the regressions in Table 2.11 is the preferred FE-IV regression, $\hat{\beta}^{FE-IV}_{InKind}$ and $\hat{\beta}^{FE-IV}_{Cash}$ here are consistent estimators of $LATE^K$ and $LATE^C$ on health outcomes.

Regression (1) in Table 2.11 shows the impact of both the FFE and PES programmes on the proportion of stunted children within a household. Neither programme had an impact in reducing the number of stunted children within households, as the coefficients on both programmes' variables are not statistically different from zero, due to the relatively large standard error that indicates that both variables were not precisely estimated. Looking at the magnitude and direction of each coefficient, the coefficients on both in-kind and cash transfers are negative and indicates that children within the beneficiary households are less stunted than non-beneficiaries. Despite the large difference in magnitude, a t-test of the difference between the two programmes' effects fails to reject the null of no difference.

Regression (2), which looks at the programmes' impact on the proportion of children who are underweight, tells a similar story. The effects of both programmes are not statistically significant from non-beneficiary households and are also not statistically different from each other. However, if we look at the magnitude and the direction of the coefficient, the in-kind coefficient is negative though small. Surprisingly, the cash coefficient is relatively large

and negative, indicating that nearly 86% of children in cash households are not underweight. Overall, the regression evidence shows that two programmes led to no difference in the health outcomes of young children within households. As before, we do not report the outcomes of the value regressions, but the results in this section are robust to using the value specification.

It must be noted however, that health outcomes were only measured in 2000 and 2006. Given the panel ageing, it is most likely that the health outcome measures at the two spots in time were of different individual children. In addition, it is almost impossible to immediately observe changes in health variables such as height or weight over a short period of time. Given the short time frame, the impact of the FFE programme on the health status is likely to be underestimated as children have not had enough time to fully benefit from the transfers. On the contrary, given that health is a stock variable that is likely affected by all previous treatment, the impact of the PES programme on health status is likely to be overestimated as some children would have been the original beneficiaries of the in-kind programme making the final health measure a result of an accumulation of both transfer programmes instead. Therefore, given these considerations, it is hard to draw a definite conclusion about the programmes' impact on long-term health outcomes of children. However, the fact that children of PES beneficiary households did not exhibit statistically significantly superior health outcomes over non-beneficiaries despite the possibility that the impact is overestimated does signal the lack of impact of both transfer programmes on children's health.

Table 2.11: Health Regression

	(1) Stunted	(2) Underweight
In-Kind	**-0.187**	**-0.066**
	(0.188)	(0.196)
Cash	**-0.699**	**-0.856**
	(0.728)	(0.730)
Log Household Size	0.188	-0.014
	(0.138)	(0.137)
Log PC Expenditure	0.004	-0.035
	(0.083)	(0.077)
Children 0-4	-0.226	-0.968**
	(0.489)	(0.491)
Children 5-9	-0.099	-0.438
	(0.463)	(0.459)
Children 10-14	-0.076	-0.170
	(0.461)	(0.453)
Adult 15-54	-0.175	-0.099
	(0.480)	(0.469)
year==2006	0.214	0.266
	(0.220)	(0.220)
Daylabour	0.130	0.166
	(0.117)	(0.117)
Low Income Occupation	0.150*	0.047
	(0.091)	(0.175)
Female-headed Household	0.222	0.103
	(0.157)	(0.147)
Child Enrolment	0.288	0.354
	(0.456)	(0.465)
Observations	632	632
Inkind=Cash	**0.426**	**0.209**

Robust-Clustered standard error in parentheses
Main statistics for InKind=Cash are the p-values
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.7.4 Summary

The results in this section show that the in-kind wheat transfer given by the FFE programme is extra-marginal and binding. Consistent with the theory, this in-kind wheat transfer did lead to a higher wheat share among beneficiaries of the in-kind transfer than they would have chosen if they were given equal-valued cash transfer. The difference in the magnitude of total food share of in-kind and cash beneficiary households gives an indication that in-kind transfer households may be forced to consumed more food than households with cash transfers, though the result were only marginally significant.

Cash households who face no constraints divert more spending to non-food consumption than in-kind households. Contrary to the usual paternalistic fears, households spend their extra cash income positively on useful items such as clothes and children-focused non-food items, and spend no more on negative consumption such as smoking. Despite the difference

in the constraint imposed by the two programmes, the substitution behaviours by in-kind households, in the end, lead to only marginal difference in the calorie consumption between households under the two programmes, which illustrates that households were able to adjust effectively. Neither programme had a significant effect on the health status of children in Bangladesh. The lack of effects could be because of the small size of the transfer, which only represents around 5% of total expenditure on average, but also the difficulty of measuring long-term health outcome must be noted. More importantly, there is no significant difference in the measured health outcomes between the two programmes, which further proves that paternalism could lead to unnecessary distortions and constraints that lower household utility without achieving better social or household longer-term welfare.

2.8 Conclusion and Policy Implications

A significant number of redistribution programmes around the world occur in-kind. This is a puzzling fact, as providing benefits in-kind not only lower recipients' utility, but are often subject to higher administration costs than providing benefits in cash. While there are many possible explanations as to why transfers more often occur in-kind rather than in cash, one of the main justifications used to provide the motivation behind most in-kind programmes is that of "Paternalism".

Paternalistic in-kind transfers intentionally cause a distortion to recipients' behaviour in order to promote outcomes that might not be achieved under cash transfers in the hope of generating a positive social welfare gain as a result. This chapter attempts to provide additional empirical evidence contributing to the debate between in-kind versus cash transfers. The empirical exercises seek to address whether such distorting paternalistic acts are really necessary given the cost involved.

The chapter makes use of a natural experiment, where a set of households experiences an exogenous change in the nature of the transfer they received, from in-kind to cash. This is captured by a three-wave longitudinal household survey collected by IFPRI from 2000 to 2006, spanning the period of two conditional transfer programmes in Bangladesh; the FFE and PES programmes.

The FE-IV approach was used in the estimation of the impact of the two programmes. This approach allows us to control for the unobserved time-invariant individual-specific effects, such as household preference, via the fixed effect (within) estimator. Failing to take these households' fixed effect into account would cause bias in the estimation under a normal pooled OLS specification. The IV approach was added to the analysis to control for the endogeneity of programme participation and possible measurement errors in the identification of the actual beneficiary households of the programmes. However, using the IV or fixed effect estimator by itself would not provide desirable consistent estimates. The FE-IV approach, which is a combination of the two, is preferred, although it requires stronger assumptions regarding the validity of the instruments. Differences in the value of the transfer were allowed for by using both the actual and hypothetical value of the transfer received instead of the beneficiary status indicator variable, which, in the end, produces similar results.

The in-kind food grain transfer, which was completely infra-marginal in terms of total

food, as predicted by the theory, has minimal impact on overall calorie consumption when compared to equal-valued cash transfer. However, in the analysis of individual food items and categories, the in-kind wheat transfer was both extra-marginal and binding, forcing beneficiary household to consume more wheat than they would have otherwise chosen to do. Households, however, were able to substitute effectively out of similar foodstuffs such as rice, keeping the overall cereal consumption around the same level in order to ease the constraint and minimise distortions. However, households under cash transfers who faced no constraint were able to substitute even more effectively out of all food categories to finance expenditure on non-food items. These substitutions result in a marginally significant difference in calorie consumption between the beneficiaries of the two programmes. Although the total food share was only marginally statistically different, the large difference in the magnitude of the two programmes gave an indication that in-kind beneficiary households were constrained to spend more on food than the cash beneficiary households.

The findings in this chapter provide minimal evidence in support of the paternalistic motivation behind the in-kind food transfer. In-kind households face a distortion in their consumption at little gain to their calorie consumption and no gain on the health outcomes among children of the households. Moreover, the extra cash drawn out of spending on food by households under the cash programme was spent positively on useful items such as clothes and children's non-food consumption, with no increase in spending on consumption of undesirable items like cigarettes. These results suggest that paternalistic in-kind transfer only caused distortion to individual consumption, lowering individual household utility at almost no extra gain to households' long-term or social welfare.

The findings in this chapter are similar to a finding in a study on a transfer programme in Mexico by Cunha (2010), which cleanly estimates the impact of cash versus in-kind transfers using a randomly assigned experiment. Cunha found little evidence of differences between cash and in-kind households and similarly concludes that the evidence supporting the paternalistic motivations behind in-kind transfers is weak. However, the findings of no programme effects in this chapter contradict earlier findings by Ahmed (2005), who concluded that households under the in-kind FFE programme consumed more calories than non-beneficiary households, while the same is not true for cash PES beneficiary households, hence concluding that there were differences between the two programmes. However, these differences could

have been the result of failing to control for the unobserved household fixed effect, which results in a bias in the estimates.

In addition to the theoretical discussion of paternalistic in-kind transfers versus cash transfers, we must not forget the policy implications of empirical research in this particular subject. In-kind transfer programmes involve higher administrative costs in both the storage and distribution of in-kind goods such as food grains. These costs may be justified if the provision of in-kind transfers leads to a substantial increase in welfare in comparison to its cash counterpart. Such an increase may outweigh these administrative costs on top of the cost to recipients' utility whose consumption was distorted by the programme. The precise comparison was not possible in this chapter due to data limitations. Given the sizeable cost saving that could result from replacing in-kind transfers with cash transfers, more empirical evidence looking more closely at the cost benefit analysis of cash versus in-kind transfer is needed.

Despite the conclusion here that paternalism may not be a sufficient justification for provision of transfer in-kind over cash, there are other valid justifications recently being investigated and these make for interesting and important findings. The pecuniary effect may provide an important justification, as recent findings show that the general equilibrium effect on the prices of in-kind and cash transfers works in an opposite direction, which could lead to different outcomes for beneficiary households (Jayachandran et al., 2010). Other considerations, such as the self-targeting induced by in-kind transfers and the political economy explanation, are also important and will provide the opportunity for a great deal of further research on the issue of in-kind versus cash transfers and the design of transfer programmes in general.

2.A Chapter Appendix

2.A.1 Price Effect

Table 2.A.1: Wheat and Rice Price Regression

	(1) Wheat Price	(2) Rice Price
FFE union	**-0.883*****	**-0.013**
	(0.165)	(0.086)
Log PC Expenditure	0.354***	0.282***
	(0.087)	(0.058)
year==2003	1.778***	2.138***
	(0.150)	(0.082)
year==2006	10.403***	6.480***
	(0.164)	(0.098)
Constant	8.750***	9.048***
	(0.583)	(0.391)
Observations	1365	1365
R^2	0.903	0.857

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.A.2 Enrolment Regression

Table 2.A.2: Enrolment Regression

	(1) Proportion of Children enrolled in primary school
In-Kind	**0.205*****
	(0.050)
Cash	**0.532*****
	(0.108)
Log Household Size	-0.034
	(0.030)
Log PC Expenditure	0.010
	(0.020)
Children 0-4	0.039
	(0.112)
Children 5-9	0.237**
	(0.109)
Children 10-14	0.117
	(0.111)
Adult 15-54	-0.022
	(0.073)
Wheat Price	0.003
	(0.005)
Rice Price	0.012
	(0.011)
year==2003	-0.214***
	(0.048)
year==2006	-0.271***
	(0.091)
Daylabour	-0.047*
	(0.025)
Low Income Occupation	-0.004
	(0.051)
Female-headed Household	-0.058
	(0.040)
Observations	1365
InKind=Cash	**0.0005*****

b coefficients; se in parentheses
Robust-Clustered standard error in parentheses; main statistics for InKind=Cash are the p-values
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.A.3 Attrition

The dependent variable, Attrition, take the value of one for households who had left the sample any time after the first wave (2000) and zero otherwise. Explanatory variables are baseline values in year 2000 for all variables that are believed to have affected the outcome of interest. Upazila 1 dummy variable was dropped as all households from Upazila 1 remain in the sample for all 3 waves.

Table 2.A.3: Attrition Probit Result

	(1) Attrition	
Household head age	-0.043***	(0.002)
Household head age^2	0.003***	(0.001)
Household Head education	-0.000	(0.999)
Household size	-0.126	(0.120)
Children age 0 to 4	0.603	(0.737)
Children age 5 to 9	0.520	(0.765)
Children age 10 to 14	0.224	(0.903)
Adult age 15 to 54	0.923	(0.530)
Asset	0.038	(0.719)
Per capita Expenditure 2000	0.026	(0.939)
Per capita Calorie 2000	0.000	(0.451)
Upazila 1==Nilphamari Sadar	0.000	(.)
Upazila 2==Mohadebpur	-1.283***	(0.001)
Upazila 3==Sherpur Sadar	-0.717*	(0.068)
Upazila 4==Modhupur	-1.075***	(0.008)
Upazila 5==Kalia	-1.137***	(0.002)
Upazila 6==Agailjhara	-0.709*	(0.074)
Upazila 7==Haziganj	-0.437	(0.198)
Constant	-0.348	(0.890)
Observations	420	
Psuedo R$\hat{2}$	0.193	
Wald Test	55.774	

main statistics are coefficients; p-value in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.A.4 First Stage Regression

Table 2.A.4: First Stage Regression of Table 2.9

	(3.1) In-Kind	(3.2) Cash	(4.1) In-Kind	(4.2) Cash	(5.1) In-Kind	(5.2) Cash	(6.1) In-Kind	(6.2) Cash
Inkind Instrument	**0.470*****	**-0.053*****	**0.458*****	**-0.102*****	**18.804*****	**-2.688*****	**21.376*****	**-2.730*****
	(0.041)	(0.013)	(0.043)	(0.032)	(1.819)	(0.729)	(2.051)	(0.764)
Cash Instrument	**-0.031*****	**0.156*****	**-0.021**	**0.134*****	**0.376**	**3.466*****	**0.362**	**3.784*****
	(0.010)	(0.036)	(0.025)	(0.042)	(1.039)	(0.877)	(1.180)	(0.984)
Log Household Size	0.019	0.081***	-0.019	0.019	-1.849	-2.803**	-2.340	-3.652**
	(0.018)	(0.029)	(0.038)	(0.053)	(1.446)	(1.272)	(1.653)	(1.513)
Log PC Expenditure	-0.023*	-0.064***	-0.006	-0.007	0.787	-0.476	0.893	-0.336
	(0.014)	(0.021)	(0.024)	(0.035)	(1.022)	(0.835)	(1.146)	(0.905)
Children 0-4	-0.085	0.054	-0.212	0.091	-5.194	3.968	-6.164	4.980
	(0.074)	(0.098)	(0.147)	(0.187)	(5.629)	(4.241)	(6.445)	(4.519)
Children 5-9	-0.060	0.197**	-0.196	0.147	-3.053	3.354	-4.356	3.910
	(0.075)	(0.094)	(0.136)	(0.170)	(5.498)	(3.946)	(6.212)	(4.464)
Children 10-14	-0.044	0.224**	-0.206*	0.328**	-6.185	5.951*	-7.948	8.723**
	(0.061)	(0.091)	(0.113)	(0.148)	(4.388)	(3.318)	(4.942)	(3.936)
Adult 15-54	-0.030	-0.020	-0.140*	0.168	-4.744	2.843	-4.696	4.181
	(0.040)	(0.056)	(0.083)	(0.119)	(3.319)	(2.686)	(3.694)	(3.032)
Wheat Price	-0.001	-0.004	-0.005	-0.002	0.519*	-0.197	0.592*	-0.161
	(0.004)	(0.006)	(0.006)	(0.008)	(0.291)	(0.174)	(0.338)	(0.177)
Rice Price	0.009*	0.000	-0.012	-0.018	-0.259	-0.350	-0.591	-0.257
	(0.005)	(0.009)	(0.013)	(0.018)	(0.558)	(0.341)	(0.626)	(0.389)
year==2003	-0.142***	0.350***	-0.102**	0.342***	-5.701***	7.567***	-5.969***	7.520***
	(0.027)	(0.039)	(0.040)	(0.055)	(1.544)	(1.125)	(1.745)	(1.203)
year==2006	-0.157**	0.341***	-0.004	0.371***	-8.959**	7.138**	-8.531*	7.381**
	(0.062)	(0.083)	(0.108)	(0.142)	(4.225)	(3.044)	(4.715)	(3.128)
Daylabour	-0.002	0.056**	-0.034	0.098**	-1.195	2.038**	-1.364	1.956**
	(0.017)	(0.026)	(0.028)	(0.040)	(1.173)	(0.836)	(1.329)	(0.952)
Low Income Occupation	-0.044	0.005	-0.022	-0.008	-1.007	0.393	-0.931	0.730
	(0.048)	(0.062)	(0.085)	(0.081)	(3.682)	(1.594)	(4.200)	(1.899)
Female-headed Household	-0.002	0.014	-0.039	0.100	-0.952	2.527*	-1.084	2.316
	(0.022)	(0.029)	(0.039)	(0.064)	(1.527)	(1.373)	(1.758)	(1.594)
Child Enrolment	0.228***	0.554***	0.247***	0.610***	10.469***	12.564***	11.917***	14.152***
	(0.064)	(0.085)	(0.083)	(0.108)	(3.529)	(2.418)	(4.000)	(2.803)
Constant	0.186	0.120						
	(0.133)	(0.164)						
Observations	1365	1365	1365	1365	1365	1365	1365	1365
F	27.34	39.40	61.31	14.91	57.77	21.25	59.25	19.79

b coefficients; Robust-Clustered standard error in parentheses
Column (3.1) and (3.2) are first stage of Column (3) in Table 2.9
Column (4.1) and (4.2) are first stage of Column (4) in Table 2.9
Column (5.1) and (5.2) are first stage of Column (5) in Table 2.9
Column (6.1) and (6.2) are first stage of Column (6) in Table 2.9
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3

Dynamics of Living Standards in Thailand

3.1 Introduction

There are three main aspects of the income distribution that have been studied in the economics literature: economic growth, inequality and dynamics. Changes in the mean income, inequality and dynamics each describe a particular aspect of the joint distribution of income, Y_t, over T periods: $f(Y_1, Y_2, ..., Y_T)$. Economic growth between two periods, t and $t + k$, reflects differences in the means of the marginal distribution of income in these periods. Changes in inequality reflect changes in the variance of the marginal distribution, and changes in income dynamics reflect changes in the covariance of income across years. Economic growth and inequality have been studied in great detail while, in comparison, works on income dynamics are still in their early stage. In this chapter, it is this dynamic aspect, which involves the movement of household earnings within the earnings distribution over time, that we focus on. We aim to find consistent estimates for the degree of earnings and income persistence as well as investigating how the dynamic processes involved vary with household characteristics in Thailand.

Background on Thailand

In the past three decades, the economy of Thailand has expanded considerably. GDP at current prices shows that from 1980 to 2011, the size of the Thai economy expanded nearly

sixteen-fold when measured in Thai Baht. Thailand had enjoyed high rates of economic growth between the mid- 1980s and the mid-1990s. According to the IMF World Economic Outlook database, between 1985 and 1996, Thailand had an average GDP growth rate of 8.8% per year, which peaked at 13.3% in 1988. During this period of high growth, Thailand was tipped to be in the next generation of East Asian Tigers following Hong Kong, Singapore, Taiwan and South Korea.

The Asian financial crisis, which started after the devaluation of the Thai Baht in June 1997, has caused a sharp decline in the Thai economy. The country's average growth rate fell to -5.9 percent per year during 1997–1998. In terms of the national currency, the country's GDP dropped from THB 3.115 trillion at the end of 1996 to THB 2.749 trillion at the end of 1998. In terms of the U.S. Dollar, it took Thailand around 10 years to regain the amount of the GDP it had in 1996. The unemployment rate went up nearly three-fold, from 1.5 percent of the total labour force in 1996 to 4.4 percent in 1998.

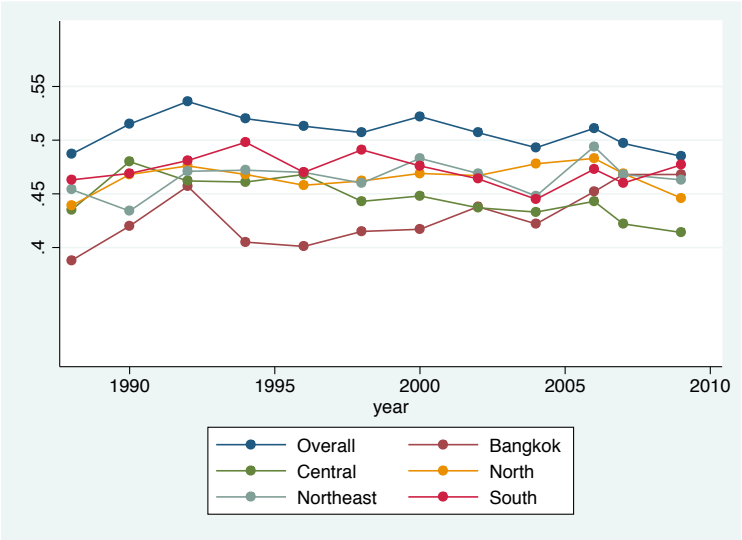
From 1999 to 2006, Thailand recovered from the downturn and had an average growth rate of 5.0 percent per year. Since 2007, however, the Kingdom has confronted a number of challenges - ranging from a military coup in late 2006, political turmoil from 2008 to 2011, and the U.S. financial institution crisis that peaked in 2008 to a historic flood in 2011. As a result, during 2007–2011, the average GDP growth rate of the country slumped once again to 2.6 percent per year.

Despite all the setbacks, Thailand is recognised by the World Bank as "one of the great development success stories". It is now an upper-middle income country. Thailand's GDP per capita at constant prices in national currency in 1985 was THB 22,997; in 2011 it was THB 71,802. This represents around a three-fold increase in real per capita income level within the past 27 years. Within 22 years, the percentage of the population living below the national poverty line decreased dramatically from 42.21 percent in 1988 to 7.75 percent in 2010. The Gini coefficient[1], a measure of income equality, however, has remained persistently high throughout. Figure (3.1) plots Thailand's Gini coefficients between 1988 to 2009. The Gini coefficient was 0.487 in 1988, and was 0.485 in 2009, its lowest level, having fluctuated around the 0.5 level for the majority of this period. Inequalities at the sub-national level also exhibit the pattern of being persistently high in every region. Thailand's income

[1]The Gini coefficient is a measure of income equality. A Gini coefficient of zero indicates perfect equality where everybody has exactly equal income. A Gini coefficient of one indicates maximal inequality where one person has all the income in the economy.

distribution is skewed towards a small percentage of the population as the richest 20% of the Thai population accounted for over half of the total household income in Thailand (54% in 2009).

Figure 3.1: Thailand Gini Coefficients 1988 to 2009



Studies based on the comparisons of cross-sections of the population over time often showed that the distribution of incomes during the period of high growth in Thailand has become more unequal or, at the very least, has remained as unequal as in the past. A study by Krongkaew and Kakwani (2003), which used the Thai Socio-Economics survey, a cross-sectional survey that collects information on incomes and household information, discovered the growth-equity trade-off pattern in the Thai economy, where high growth leads to a more unequal society. However, studies of inequality that are based on comparable cross-section of individuals provide only a snapshot, which reflect the income distribution at a specific point in time and reveal nothing about the dynamics of the income distribution.

Income and Earnings Dynamics

This study examines income and earnings persistency and also income and earnings mobility, a closely related concept, in Thailand between 1985 and 2011. Income persistency and mobility, which are studies of the dynamics of the income distribution and how it is changing over time, is a crucial concept. The degree of persistency in income is often seen as a measure of the equality of opportunity, and of the flexibility and freedom of movement in the labour market (Atkinson et al., 1992). The welfare of individuals in two societies with similar levels

of income inequality but different patterns of income dynamics would be expected to differ. Individuals in a society where incomes are less persistent would enjoy a greater incentive to exert effort and climb up the income distribution than individuals in a society where income persistency is high, resulting in higher productivity in the overall economy.

Mobility and persistence can be studied over multiple time frames. Intra-generational mobility measures the movement of individuals within the income distribution in comparison with their peers over the course of their life. On the other hand, inter-generational mobility measures the movement of individuals within the income distribution across generations. In terms of persistence, if individual income is represented by $Y_{i,t}$, intra-generational income persistence measures the correlation between $Y_{i,t}$ and $Y_{i,s}$, where t and s are the different time periods in which the same individual income is observed. This could be a correlation between current and last year or even further lags of income. For inter-generational persistence, however, t and s represent different generations. In this setting, individual i's income observed in time period t is compared with the income of his/her parents observed in time period s.

There are two different concepts to measure income mobility: absolute and relative. Absolute mobility measures the growth in individual income across times or generations. It is possible for society as a whole to achieve upward or downward absolute income mobility as everyone's income move up or down together. Relative mobility, on the other hand, measure the relative movements and positional changes that occur within the income distribution. Relative mobility is a zero-sum game in a sense that as certain individuals achieve upward movements, another group of individuals will have to move down the income distribution to compensate. Measures of the degree of absolute mobility can be captured by the regression framework, while the degree of relative mobility are usually studied via the construction of income transition matrices.

Within the regression framework, measures of mobility and persistence are further classified in to unconditional and conditional. Unconditional persistence measures correlation between current earnings and past earnings without explicitly holding other things equal. It is a measure of overall persistence in the economy for everyone regardless of their individual characteristic, whether or not ones are similar. Conversely, conditional persistence measures the correlation of earnings among individuals with similar characteristics. Conditioning

variables differ among different studies but mainly include variables that affect income and earnings such as education, household characteristics, sector transitions, etc.

This chapter studies intra-generational income mobility and persistence in Thailand from both absolute and relative aspects. Unconditional persistence and persistence conditional on observables such as education, household size, the sector of employment as well as the time-invariant unobservables, will also be estimated. Given the high level of inequality observed in Thailand throughout the period of its economic expansion, a study that quantifies the degree of persistency and investigates the dynamic process will provide important additional information on how individuals have fared and assist policy makers in dealing with the high level of income inequality.

Evidence from Panel Data

As empirical studies of income dynamics involve measuring changes in income over time, it is necessary to observe household or individual income in more than one period. Longitudinal data are thus crucial for this type of analysis. Given the availability of long-running panel data in developed countries, many empirical studies have long been conducted in countries such as the US and the UK. Jarvis and Jenkins (1998) used the 1991–1994 British Household Panel Survey (BHPS) data to study both income mobility and persistency in the UK, using the transition matrix and the Galton-Markov regression-based approach respectively. They find that while there was a high level of mobility (low persistence) in income among the population studied, even over a one-year interval, the degree of mobility observed was not very large. A more recent estimation by Jenkins (2011), which employed more recent rounds of the BHPS (1991–2006), had similar findings to the earlier study. Importantly, the study finds no clear trend for income persistency over time in Britain.

Mobility estimation was also documented for the USA in the 1970s, 1980s, and 1990s by Hungerford (1993, 2011), who used decile transition matrices for family income calculated from Panel Study of Income Dynamics (PSID) data. The results are similar to those for Britain in the sense that "for the most part there was considerable movement within the distribution during the two observation periods (1969–76 and 1979–86), but the movement generally is not very great in either direction". Hungerford also found that aggregate mobility in the USA did not change much between the 1970s and 1980s, but fell slightly between the

1980s and 1990s.

The recent availability of panel data in developing countries has made estimation of income and earnings persistence based on the Galton-Markov regression-based approach in these countries possible. Fields et al. (2003) studied income and earnings persistence in Indonesia, South Africa and Venezuela and finds "unconditional convergence" in income in these countries. Convergence, in this case, refers to the concept that the largest changes in income or earnings are experienced by those who have the lowest incomes and earnings to begin with. Hence, there is a process whereby the initially poor are catching up with the initially rich. "Unconditional" here refers to the method of performing a regression of changes in income only on past income to determine degree of persistency without conditioning on any other variables. Similar findings of unconditional convergence in Argentina and Mexico (Fields et al., 2007) and in Vietnam (Glewwe and Nguyen, 2002) were also reported. In China, having used the transition matrix method, Khor and Pencavel (2010) found that income mobility was high in the 1990s. However, by the early 2000s, while China's income inequality had increased further, income mobility had decreased, implying that the probability of being stuck at a lower level of income has increased for households in China.

These studies of income dynamics, however, suffer from two major problems. First is the measurement error problem in reported income or earnings in survey data, which leads to the classic attenuation bias problem, resulting in an under-estimation of the degree of income and earnings persistence in the economy. Researchers in the field have attempted several methods to solve this major problem. Dragoset and Fields (2006), used both survey earnings data and the less error-prone employer-reported total annual labour earnings from the Social Security Administration's Detailed Earnings Record (DER) to estimate the earnings regression, and found greater persistence with the more accurate administrative data. However, the lack of access to reliable administrative data to conduct similar analysis, especially in developing countries, has prevented this method from gaining more widespread adoption.

Another approach is to use instrumental variable methods. Many instruments which are believed to be correlated with lagged income or earnings but uncorrelated with measurement errors, such as lagged expenditure (e.g. McCulloch and Baulch, 2000), measures of predicted permanent income based on asset and land holding (e.g. Fields et al., 2007), and weight (Glewwe and Nguyen, 2002) have been used in the literature. However, Antman and

McKenzie (2007b) have shown that under a certain set of general assumptions the key conditions for the validity of the instruments are that the instruments must be uncorrelated with the measurement error and must also have the same degree of persistence as the earnings or income being instrumented. This condition is extremely unlikely to be satisfied in practice.

The second major problem with studies of income dynamics is attrition bias. Non-random attrition, where households leave the sample in a systematic manner that is correlated with the outcome being investigated, will cause bias in the estimates. In the case of mobility, it may be that households that leave the sample by moving away are the more mobile ones, which can move to take advantage of other opportunities in comparison to those that remain in the sample. This may cause the calculation based on the remaining biased sample to under-estimate the true level of mobility or to over-estimate the level of persistency. Thomas et al. (2001) discussed the experience of the Indonesia Family Life Survey, which explicitly tracked movers, and found that those who moved are different in terms of their initial characteristics than those who stayed. However, the literature on income and earnings dynamics has devoted less attention to assessing the impact of attrition on estimates of mobility and persistency and correcting attrition bias.

Moreover, one of the biggest obstacles to the study of income and earnings dynamics in developing countries is the lack of availability of representative, long-running longitudinal data sets. In most of the studies in developing countries mentioned above, the panel data sets used are only available in small numbers of waves and therefore, the researchers were only able to study dynamics of earnings over a short time period. This gave a biased picture of the degree of persistency observed as it is likely that individuals or households would experience little change in income over a short time span. Alternatively, most of the observed movement within the income distribution could be a result of transitory shocks, which happened to occur during these short periods of observation, yielding an under-estimation of true persistency. The lack of panel data has been one of the major factors that has prevented studies of the dynamic aspect from taking centre stage in discussion regarding the distribution of income in developing countries.

Evidence from Pseudo-Panel Data

Deaton (1985) suggested a method to circumvent the unavailability of genuine panel data

by following cohorts of individuals over time instead. These cohorts are constructed from repeated cross-sectional household surveys to create pseudo-panel data. However, only recently has a method for estimating income dynamics using these pseudo-panel data become available. There are several benefits to using pseudo-panel over standard panel data in investigating income and earnings dynamics. First, while most panel data tend to focus on certain geographic regions or specific groups, cross-sectional household surveys tend to be more representative of the population as a whole. Second, repeated cross-sectional household surveys are available in most countries, especially the developing ones, and are available over longer periods and a greater number of observations than genuine panel data. Third, given that pseudo-panel data track cohorts instead of individuals over time, they do not suffer from the problem of attrition bias as it is not necessary for the same households to be resampled in the next period to analyse cohort dynamics. Last, and most important, is the ability of pseudo-panel data to deal with bias from measurement error. Construction of pseudo-panel data involves taking cohort means within each time period, and this averaging process helps eliminate individual-level measurement errors in the cross-section. One major drawback with this approach, however, is the fact that within-cohort dynamics of income and earnings cannot be investigated.

There are several recent works that used pseudo-panel data to analyse income and earnings dynamics in many countries, mainly those in Latin America. Countries in this region draw the attention of researchers as they have suffered from persistently high levels of income inequality, making the issue of income dynamics an important one. In addition, many countries in Latin America do not currently have long-running genuine panel data but have decent collections of repeated cross-sectional household surveys available for use. Antman and McKenzie (2007b) showed that dynamic pseudo-panel methods can be used to consistently estimate the earnings persistence parameter when genuine panel data are not available and also in the presence of non-classical measurement errors, provided that the dynamic process of earnings is truly generated by the Galton-Markov process. They applied this method to data on earnings from Mexico and found that unconditional persistency in Mexico is very high, suggesting that the high level of inequality found in the cross-section is persistent. However, conditional persistency in Mexico is low so households are able to recover quickly from earnings shocks.

Calonico (2006) used the dynamic pseudo-panel method to study earnings persistence in eight Latin American countries and found the degree of persistence to be very high in all countries. Navarro (2010) also employed this approach to study earnings dynamics in Argentina, using the same data set as Calonico, but found a lower degree of persistency. The difference stemmed mainly from the fact that Navarro analysed the data over a different time span, made use of a different measure of earnings and only used a sub-sample of the survey in order to create a larger pseudo-panel. Most recently, Cuesta et al. (2011) estimated earnings persistence from 1992–2003 in 14 Latin America countries. The authors also found a high degree of unconditional persistence but a lower degree of conditional persistence, in agreement with the findings in Calonico as well as Antman and McKenzie.

Outside of the Latin America region, dynamic pseudo-panel methods have been used to analyse earnings dynamics in Korea (Bourguignon et al., 2004). The authors show that under the assumption that individual earnings dynamics obey some basic properties and follow a simple stochastic process, the main parameters of this process can be recovered from repeated cross-sectional data. With these parameters they simulated the earning dynamics of an individual and estimated other measures of interest, such as the individual's vulnerability to poverty. They found a high degree of earnings persistence in Korea. Their results showed that model parameters recovered from pseudo-panels approximate reasonably well those estimated directly from genuine panel data that were available in Korea.

Evidence from Quantile Regression

All these studies of income and earnings dynamics using both the standard and pseudo-panel data have concentrated only on one aspect of the inter-temporal linkage of the income distribution, the persistence of the conditional mean. This ignores persistence at other parts of the conditional income distribution although it is likely that individuals at different parts of the income distribution will experience different degrees of earnings persistence. By only focussing on the inter-temporal linkage of the mean through the use of the Ordinary Least Squares (OLS) method, estimates obtained could only give a partial picture of the actual degree of earnings persistence experienced by individuals. These shortcomings necessitate the move beyond OLS to other techniques that will allow the investigation of persistence at different parts of the income distribution. Quantile regression (Koenker and Bassett, 1978)

is one of the methods that could achieve this goal.

There are several studies that have estimated inter-generational income persistence using the quantile regression model. Eide and Showalter (1999), using the PSID data, estimated the degree of inter-generational persistence for the US. They found that the inter-generational linkage between fathers' and sons' earnings is greater at the bottom of the sons' conditional earnings distribution than at the top. They also found that controlling for the sons' education reduces the inter-generational earnings correlation, and that education is more valuable at the bottom of the conditional earnings distribution than at the top. Several other studies using quantile regression also find differing degrees of inter-generational persistence at different parts of the distribution (Fertig (2001); Bratberg et al. (2005)).

In an intra-generational framework, as studied in this chapter, there are very few studies that make use of the quantile regression technique to study income dynamics. An exception is a recent study of intra-generational earnings persistence in Argentina using quantile regression by Navarro (2007). The author found that persistence is greater at the bottom and the middle of the conditional earnings distribution than at the top, thus rejecting the hypothesis that inter-temporal linkages in income are identical across the conditional income distribution. Nevertheless, there has not been enough effort in the intra-generational studies of income dynamics to move beyond mean regression.

Permanent Income Model

The aforementioned studies all seek to estimate the persistence parameter mainly from the Galton-Markov model of earnings dynamics. While this simple model can provide a lot of information on the dynamics of the earnings process, it is purely a statistical model of earnings. To gain greater understanding of the various components that determine the earnings process, it is necessary to move beyond this statistical model to models with greater economic structure such as the permanent earnings model.

The permanent earnings model attempts to model the covariance structure of earnings and decompose earnings into permanent and transitory components. The relative importance of the two components is crucial in determining the dynamic process. If the permanent components dominate the earnings process, an individual's initial position in the earnings distribution will matter more and there will be relatively little change in an individual's

relative position over time. Earnings differences are very persistent in this case. However, if the transitory component dominates, there will be more movement within the earnings distribution and the initial difference in the starting position of individuals will matter less. As earnings differences at a point in time reflect a combination of long-term permanent and short-term transitory effects it is important to investigate this type of process in order to shed light on the appropriate policy response as the policy implications of these two effects are drastically different.

There are a great number of studies that strive to model the covariance structure of earnings. An early contribution was made by Lillard and Willis (1978) who, using the PSID sample, modelled the dynamics of earnings by allowing the transitory component to follow the simple AR(1) process while holding the permanent component fixed over the life cycle. Several studies have extended this basic model to allow for a more general time series structure covering the transitory element such as the autoregressive moving average (ARMA) process (MaCurdy, 1982)). Others have sought to model the permanent component by allowing it to vary over the life cycle following a random walk in age (Moffitt and Gottschalk, 1995) or using a heterogeneous growth model (Lillard and Weiss, 1979; Guvenen, 2009) or both (Baker and Solon, 2003). The maximum likelihood estimation technique was first used to identify parameters in these earnings models, but the method of Minimum Distance Estimation developed by Chamberlain (1984), which was first used to study covariance structure of earnings by Abowd and Card (1989) is now the common procedure.

Even with the vast literature on the modelling of the covariance structure of earnings, almost all have concentrated on developed countries. This is due mainly to the availability of reliable data in these countries. One of the few studies on a developing country was conducted in Brazil (Santos and Souza, 2007) and has reported a high degree of time-invariant individual permanent components that cannot be accounted for by education or age. It will be interesting to investigate whether this finding will also be observed in Thailand.

Contributions and Findings

This chapter contributes to the existing literature by providing the most complete and comprehensive empirical study of income and earnings dynamics in Thailand to date. The study spans a period of 27 years, covering periods of high growth, the financial crisis, the recovery

from the crisis and the slowdown as a result of political turmoil. The dynamic pseudo-panel method is used on pseudo-panel data constructed from the 1985 to 2011 annual Labour Force Survey (LFS) to obtain consistent estimates of earnings persistence in Thailand. In addition, genuine panel data from the Townsend Thai Project, which are available from 1997 to 2010, are used to study income dynamics, providing comparisons with the result obtained from the pseudo-panel data. This chapter, to the best of our knowledge, is the first time that the Townsend Thai panel data, the only long-running panel data set in Thailand, are being used to study income dynamics.

This study makes use of the recently proposed method to control for the age, period and cohort (APC) effects all at once, using the maximum entropy principle of Browning, Crawford and Knoef (2012). This method overcomes the classic APC perfect multicollinearity problem to arrive at a more accurate estimate of conditional earnings and income persistence. The quantile regression technique is also employed to study earnings dynamics at different parts of the earnings distribution in Thailand. In addition, this study will present evidence beyond the Galton-Markov regression framework and seek to model the covariance structure of earnings and income in Thailand. This is the first time that such variance decompositions are performed for Thailand, thus providing greater understanding of the dynamic process of earnings and income in the country.

We find that estimates obtained from the Townsend panel data suggest a low degree of unconditional and conditional income persistence in Thailand. However, these estimates are biased and likely to understate the degree of persistence as a result of measurement error in income, non-random attrition and non-representativeness of the sample. Dynamic pseudo-panel methods are applied to pseudo-panel data constructed from nationally representative LFS to obtain consistent estimates of unconditional and conditional income and earnings persistence in the presence of measurement error and sample attrition. Pseudo-panel estimates indicate a high level of unconditional earnings persistence in Thailand. However, controlling for the APC effects lowers the estimated persistency significantly. These findings of low conditional earnings persistence suggest that households are able to recover quickly from shocks to their labour earnings. However, the fact that unconditional earnings persistence remains high suggests that household fixed effects and characteristics are important and that the high level of income inequality found in the cross-section will likely persist over

time.

Quantile regression estimates of both sets of data find that the degree of income and earnings persistence differ at various parts of the conditional distribution. While the degree of persistence is low at the bottom end of the income distribution, income exhibits a great level of persistence over time at the top end. In addition, we find that the income distribution is diverging, which suggests that those at the top end of the distribution are moving further away from those at the bottom over time.

Analysis of the covariance structure of income and earnings confirms the findings obtained under the Galton-Markov model. In both samples, estimated parameters from the permanent income model indicate a relatively low degree of persistence in the transitory element over time. In addition, the magnitude of the transitory component was found to be much greater than the permanent component of the earnings process. These findings suggest a low degree of conditional earnings and income persistency in Thailand, consistent with earlier findings. Moreover, estimation of various specifications of the permanent earnings model confirms that the basic AR(1) model with fixed individual effects provides the best fit for data in both samples, thus providing support for the use of the Galton-Markov process to model income and earnings dynamics in the first part of our analysis.

The structure of this chapter is as follows. Section 3.2 presents various measures of living standards and earnings dynamics to arrive at those chosen for the analysis in this chapter. Section 3.3 discusses the identification of the parameters of interest and explains the econometric specification for estimating these parameters. Section 3.4 describes the survey data and presents the descriptive statistics from both the LFS pseudo-panel data and the Townsend Thai panel data. Section 3.5 presents the empirical results. Section 3.6 offers a discussion of the policy implications of the findings and concluding remarks.

3.2 Measures of Living Standards and Income Dynamics

In this section we present the various measures of living standards and earnings dynamics that are studied in the current literature. While we only give brief introductions to most of them, we will discuss at length the measures used in this chapter: the origin independence and persistence of household real equivalised earnings and income.

3.2.1 Measures of Living Standards

Households in a cross-section differ in size, structure and many other characteristics. In addition, the same household observed in different periods also experiences different price levels. In this chapter we will investigate the dynamics of household living standards over time and also determine the degree of persistence in these living standards. It is important, therefore, to have a measure of economic wellbeing that can be compared over time across both the same and different households in the survey.

Income and expenditure data in surveys are normally collected at the household level. For survey data used in this chapter, the Townsend Thai panel data set only provides detailed household-level income data. While the LFS does provide earnings for each individual in the household who works during the survey, it is more meaningful to first aggregate the earnings/income data at the household level in order to reflect the pooling of income within a household. As a result welfare comparisons between households in these surveys, both within the cross-section and over time, require two main adjustments to be made to the nominal household-level income data, namely price and demographic adjustment.

In order to compare household welfare we make use of the concept of money metric utility, where the indifference curves of individual household preference are labelled by the amount of money needed to reach the utility levels corresponding to those indifference curves for some fixed set of prices. Money metric utility can be approximated by real income or real expenditure, the two most commonly used welfare measures. This allows us to avoid having to specify a parametric utility function.

In a standard consumer choice theory, a consumer maximises a utility function $\nu(q)$ by choosing a bundle of goods q to satisfy the budget constraint $p \cdot q = y$. If the utility-maximising choice of household i results in a utility value of u^i, income y must be the minimum cost of attaining u^i at prices p and characteristics z. This minimum cost can be

represented by the cost function $C(u^i, p, z)$, and we can write $C(u^i, p^i, z^i) = y_i$. Although utility values are not observable, the cost function is a monotonic and increasing transformation of the utility function; hence, it can be used as a measure of utility and the values compared at fixed p and z. Therefore, the main issue we want to resolve here is how to fix these p and z in the cross-sections and over time to arrive at $C(u^i, p^0, z^0)$, a measure of individual household utility at a reference price (p^0) and with reference characteristics (z^0) that would allow for comparisons of living standards between households.

The cost-of-living indices are comparisons of the costs of obtaining the same utility at two different sets of prices. Formally, this is

$$P_i(p^i, p^0; u^i, z^i) = \frac{C(u^i, p^i, z^i)}{C(u^i, p^0, z^i)} \tag{3.1}$$

Where u^i is household i utility level and z^i is a vector representing household i characteristics such as demographic structures, which more specifically are the size and age structure of the household. The scalar value P_i is the relative cost of maintaining the standard of living u^i for a household with the characteristic z^i at the two sets of price vectors, the current price p^i and the reference price p^0. We can use this cost-of-living index to adjust for differences in the price levels that households face and convert nominal to real income. The validity of the cost-of-living indices is based on the assumption of unchanging tastes. In this chapter, the Thai Consumer Price Index (CPI) was used for these price adjustments.

In addition, we must make further adjustment concerning household demographic structures. An important assumption made is that the differences in tastes between households are only a result of variations in observable demographic characteristics z. Households that face the same price will only have different welfare levels as a result of differences in z.

One of the crudest and most popular adjustment methods is to deflate total expenditure by a simple headcount of the number of members within a household. However, this simple approach is problematic for two main reasons.

1. Per capita adjustment ignores the variation of need with age. Young children need less than older children, who also need less than adult. Ignoring this age variation would result in an under-estimation of the economic wellbeing of a household with lots of children in comparison to a household which consists of only adults.

2. Per capita adjustment overlooks the opportunities for economies of scale in consumption. Four people do not need double the amount of two people as they can usually share the fixed cost. Buying and cooking food for more people is also cheaper on a per capita basis. Again, ignoring this leads to the under-estimation of wellbeing for households with more members.

For these reasons we would prefer a measure involving equivalence scales, which, at the very least, take into account the fact that children's needs are only a fraction of those of adults. These equivalence scales are to welfare comparisons between households with different characteristics what cost-of-living indices are to welfare comparisons for a given household facing different prices.

If u^i and p^0 are the household utility level and the reference price vector respectively, the equivalence scale is

$$E_i(z^i, z^0; u^i, p^0) = \frac{C(u^i, p^0, z^i)}{C(u^i, p^0, z^0)} \tag{3.2}$$

The reference characteristics z^0 could represent any characteristics of choice. As an example, we could think of z^0 as representing a household consisting of two adults and z^i as a household with two adults and a newborn child. $C(u^i, p^0, z^0)$ here would therefore measures the minimum cost of reaching utility level u^i at the reference price p^0 for a household with two adults, while $C(u^i, p^0, z^i)$ would measures the minimum cost of also reaching utility level u^i at the reference price p^0 but for a household with two adults and a child. The excess of $E(z^i, z^0; u^i, p^0)$ over unity would be the cost of the child as a ratio of total household income $C(u^i, p^0, z^0)$ for the reference group of two adults. $E(z^i, z^0; u^i, p^0)$ itself would be the required adult equivalence scale.

Unlike the cost-of-living indices, unique estimation of equivalence scales cannot be calculated from knowledge of the demand function alone (Pollak and Wales, 1979) as the reallocation of resources within the household, which does not affect the demand of goods that the household purchases but affects how these resources are shared among its members, is not directly observable. Therefore, different cost functions, which represent different preferences and different costs of children, could result in the same observable behaviour (demand and total expenditure). This inability to determine a unique cost function that generates the data from knowledge of the demand function alone results in the under-identification of

equivalence scales. A solution to this problem is to make an extra identifying assumption that would allow a unique measurement.

The oldest method for constructing equivalence scales, and one that is still widely used in practice, was proposed by Engel (1857). It is based on the identifying assumption that the share of food expenditure in total outlay correctly indicates welfare between households of differing demographic composition. A large household and a small household are equally well-off if, and only if, they devote the same fraction of their budget to food. Child costs are then defined as the difference in total spending for two households of different sizes (with number of children being the difference in size) when the families spend the same share of their budget on food. However, a major problem with Engel's method lies in its implicit assumption that adults (parents) and children consume the same shares of food relative to shares of non-food items out of the total household budget. Children are generally believed to be food-intensive, in that food is a bigger share of what is spent on children. If children consume food as a greater proportion of total spending, then the Engel methodology will compensate with even greater levels of income to return the family's overall spending on non-food items to its pre-child share. This leads to the over-estimation of child costs.

Another widely used method is one proposed by Rothbarth (1943), which measures a household's economic wellbeing based on the level of spending on selected goods consumed only by the adults in the household. For a given level of income, additional children reduce the amount of household spending on adult goods as resources within the household are redirected towards the children. Rothbarth's method works by calculating how much additional income is needed to restore the pre-child level of spending on those adult goods for the household. Child costs are the difference in total spending for two households of different size (with number of children being the difference in size) when the households spend the same amount on adult goods. Alcohol, tobacco and adult clothing have been used as adult goods in the literature.

In this chapter, we will not attempt to estimate the equivalence scales for Thailand. Instead, we make use of the adult male equivalence scale for India estimated from a dietary survey which measured caloric intake at the individual level for distinct age-sex categories in Townsend (1994), which more closely resemble the sample of the Thai population than the OECD equivalence scales, and was used for an analysis of inequality in Thailand by Jeong

(2008). The equivalence scales are listed in Table 3.1.

Table 3.1: Equivalence Scale

Individual Characteristic	Townsend (1994) Equivalence Scale
Male: over 18	1.0
Female: over 18	0.9
Male: 13 to 18	0.94
Female: 13 to 18	0.83
Both: 7 to 12	0.67
Both: 4 to 6	0.52
Both:1 to 3	0.32
Both: Less than 1	0.05

With appropriate adjustment for price and household demographics (y_i/P_iE_i), we arrived at the final real equivalised income $C(u^i, p^0, z^0)$, a standard of living measure that could be compared across households.

$$\frac{y_i}{P_iE_i} = C(u^i, p^i, z^i)\,\frac{C(u^i, p^0, z^i)}{C(u^i, p^i, z^i)}\,\frac{C(u^i, p^0, z^0)}{C(u^i, p^0, z^i)} = C(u^i, p^0, z^0)$$

The cost-of-living indices (Thailand CPI) and the Townsend equivalence scales discussed above are used to adjust all earnings and incomes, our measures of living standards, in both the LFS and Townsend data sets.

3.2.2 Measures of Income Dynamics

Measures of income dynamics are further divided into those which measure economic mobility and those which measure income persistence. While these two concepts are closely related, they are distinctly different from one another. Economic mobility examines the movement of individuals within the income distribution as time passes. On the other hand, income persistence measures the correlation of income over time and does not necessarily convey any information about the movement of individuals within the income distribution.

Despite the main focus on measures of income persistence in this chapter, we will also investigate economic mobility in Thailand through the construction of the transition matrix from the Townsend Thai panel data.

3.2.2.1 Measures of Economic Mobility

While income mobility is a very important concept, there is relatively less agreement on how one can measure mobility, especially in comparison to the measurement of inequality. These difficulties often arise as a result of the lack of clarity as to which aspect of economic mobility we should be measuring. As put by Fields and Ok (1996a), "The mobility literature does not provide a unified discourse of analysis. This might be because the very notion of income mobility is not well-defined; different studies concentrate on different aspects of this multi-faceted concept". Different studies make use of different indices that measure different aspects or concepts of economic mobility. It is important, therefore, to specify clearly which concepts and aspects of mobility we are measuring.

Jenkins (2011), in his review of the literature on income mobility concepts and measures, distinguishes among three main concepts of economic mobility: mobility as reduction of longer-term inequality, mobility as income growth, and mobility as positional change. While it is this latter concept of mobility as positional change, more specifically the origin independence of income, that is the main focus of study in this chapter, we provide a brief discussion on all of the concepts above.

Mobility as Reduction of Longer-Term Inequality

The first mobility concept links income mobility with its impact on inequality. Income mobility means that there are changes in incomes for most people between one year and the next. Therefore, in any given year, an individual's income in that year will differ from the average of his/her income taken over several years. Averaging across time smooths the longitudinal variability in each person's income. In addition, differences across individuals in these averaged incomes will be less than the dispersion across individuals in their incomes for any single year. Mobility can therefore be characterised in terms of the speed at which inequality is reduced as the accounting period for income is lengthened from one year to a longer period (Shorrocks, 1978b). More mobility corresponds to a greater reduction in inequality over an accounting period of a given length.

One measure of mobility as an equaliser of longer-term income proposed by Fields (2005a) is $\varepsilon = 1 - \frac{I(Y_a)}{I(Y_1)}$, where Y_a is a vector of individual average income, Y_1 is the vector of base-year income and $I(.)$ is an inequality measure such as the Gini coefficient or the Theil's

index. When income over the longer period is distributed as unequally as the income in the base period, $\varepsilon = 0$. $\varepsilon > 0$ ($\varepsilon < 0$) only when longer period income is distributed more (less) equally than the base-year income. The greater the magnitude of (positive) ε the greater the mobility.

Mobility as Income Growth

Mobility as income growth refers to measures summarising the change in income experienced by each individual within the society between different points in time. The change might be a gain or a loss, and can be in absolute terms (expressed as an increase or decrease in a monetary unit) or proportionate terms (expressed as a percentage of original income). Income growth is defined for each individual separately and overall income mobility for society is obtained through the aggregation of the experience of each individual. This concept of mobility can be thought of as measuring "absolute" mobility, where positive income growth for every individual is regarded as an increase in mobility even though relative positions are maintained.

One of the measures of mobility based on the income growth concept is the mean absolute income change, per capita, of Fields and Ok (1996a). This is $m_n(x, y) = \frac{1}{n}\sum_{i=1}^{n} |x_i - y_i|$ where x and y are individual i incomes in two different periods. This measure looks at the size of the fluctuations in individuals' incomes and weighs the gains and losses similarly without taking the direction of change into account. Alternative measures that look at individuals' income changes and also their direction are average income change ($m_n(x, y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)$), the average amount gained by the winners and lost by the losers.

In addition to the directional aspect of income gain and loss, the value of those gains and losses relative to an individual's original level of income may be important. For example, two individuals A and B both received an increase in income of 1,000 pounds per month. It is plausible to believe that the effect of this increase on each individual's mobility would be drastically different depending on A's and B's original levels of income prior to the increase. If A and B were earning 1,000 and 100,000 pounds per month respectively prior to the increase, while the 1,000 pound increase lead to a doubling of A's income, it represents only 1 percent of B's income. Therefore we would realistically expect A to have experienced more income mobility than B despite an equal absolute gain in their income. Fields and Ok (1996a)

also devised a measure of the mean change in log income $m_n(x,y) = \frac{1}{n}\sum_{j=1}^{n}[log(x) - log(y_i)]$ which combines directional income gains and losses while also taking in to account the income levels of the gainers and the losers as previously discussed. For all measures stated above, higher values of $m_n(x,y)$ indicate higher level of economic mobility.

Mobility as Positional Change

In contrast to the concept of mobility as income growth, mobility as positional change is based on the idea that mobility for a specific individual also depends on others' positions in the income distribution. This concept of mobility can be thought of as measuring "relative" mobility, where mobility for a person depends not on whether his or her income has increased or decreased but on how the change alters his or her position relative to others. With positional change mobility, not everyone can be upwardly mobile.

Measures of positional change encapsulate the situation of maximal or perfect mobility in two different ways. One view is that perfect mobility occurs when one's current income is completely unrelated to one's income origin; this is often termed "origin independence" in the literature. Another view is that perfect mobility occurs when final positions are a complete reversal of positions at the origin; this is often termed "rank reversal".

In this chapter, the measures of income mobility that we employed are based on the origin independence concept of mobility. Origin independence is a measure of the degree to which current income is determined by past incomes. Origin independence is at its minimum when current income is completely determined by past incomes and this would suggest that there is no mobility at all in the economy. Mobility and origin independence are greatest when current and past incomes are unrelated.

The most common method of measuring origin independence is through the construction of a transition matrix. A transition matrix is a matrix that determines the portion of people who stay or move between income classes, in this case an income quintile, between the base year and the final year. The base year is normally the row while the final year is the column of the transition matrix. The entries of the transition matrix are conditional probabilities. For example, p_{12} (entry in the first row, second column) is the conditional probability of a household moving to the second quintile in the final year given that it has started in the first quintile ($p_{12} = P[Q_2\, in\, final\, year |\, Q_1\, in\, base\, year]$). These probabilities sum up to one

in each row and column. If income were perfectly positively origin-dependent, individuals would not move between the quintiles at all and the quintile transition matrix would equal the identity matrix, with all entries on the diagonal equal to one as shown below.

$$
\begin{array}{rc}
 & \textit{Income quintile in final year} \\
 & \begin{array}{ccccc} 1st & 2nd & 3rd & 4th & 5th \end{array} \\
\textit{Income Quintile in base year} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}
\end{array}
$$

In the case of perfect origin independence, 20% of those in each base year income quintile would be found in each final year income quintile, which would produce a matrix like the one below.

$$
\begin{array}{rc}
 & \textit{Income quintile in final year} \\
 & \begin{array}{ccccc} 1st & 2nd & 3rd & 4th & 5th \end{array} \\
\textit{Income Quintile in base year} & \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}
\end{array}
$$

The actual transition matrix that we observed is likely to be between the two extremes shown above. Some measures of the degree of positional change between incomes in a pair of years based on the transition matrix include the mean number of quintiles moved or the mobility ratio ($1 - Trace(M)$, where M is the income transition matrix).

Despite the widespread use of the transition matrices in income mobility studies there are some major drawbacks to this approach. First, the degree of mobility observed from the transition matrix depends on the division of income into bins such as quintiles or deciles, in which the selection is entirely at the discretion of the researcher. With smaller bins, even minor changes in real income may result in a shift between ranks, suggesting a higher degree of mobility. Second, a transition matrix neglects the within-quintile changes and the different absolute income changes that underlie a change in income quintiles, which may result in a misleading picture of the mobility of income. For example, if there is no mobility between quintiles, the transition matrix approach will create the false appearance of complete

immobility even though there might be substantial mobility within quintiles. Finally, the individuals or households in the bottom quintile cannot move downward even though their incomes might have decreased greatly, while households in the top quintile cannot move upward even if they have experienced tremendous rises in income – the floor and ceiling effects. Despite these shortcomings, transition matrix still presents a very informative and intuitive way to observe movement of individuals within the income distribution.

In this study, transition matrices will be calculated for the Townsend panel data only. It is not meaningful to construct transition matrices for the pseudo-panel data as the repeated units we observed are not made up of the same individuals. In addition, too few cohorts were observed to make the quintile transition matrix informative.

3.2.2.2 Measures of Persistence

Income persistence measures the degree of correlation of income over time. While this is very closely related to the concept of origin independence mobility discussed above, they are not the same. Correlation between incomes over time does not necessarily provide any information on individuals' movement within the income distribution, unlike the transition matrix, which clearly indicates where individuals have moved, but it does provide information on the degree of dependence of current income on past income. In addition, measures of persistence capture the concept of absolute mobility, as with lots of changes in incomes, the degree of persistence observed will be lower. Therefore, a society with a high degree of income and earnings persistence can, to a certain extent, be viewed as a society with less economic mobility (if mobility is viewed from an income growth perspective).

One measure of income persistence is the coefficient of the regression of individuals' income on their lagged income. This measure based on the Galton-Markov linear regression framework, also referred to as the "statistical model of earnings dynamics" (Atkinson et al., 1992), is commonly used in the empirical literature on income dynamics (Jarvis and Jenkins, 1998; Fields et al., 2003). Unlike the transition matrix, which breaks income into quintiles, the regression-based method uses all available information inherent in the actual distributions. Moreover, being based on a regression framework, this method of investigating income dynamics allows for the use of various techniques such as instrumental variables and the pseudo-panel method as possible fixes for problems of measurement errors and attrition

that often trouble efforts to measure income persistence accurately. The degree of income and earnings persistence will be estimated for both the Townsend and the LFS data. We discuss this method and its extensions in more detail in Section 3.3 on identification and econometric specifications.

3.2.3 Summary

The main objective of this study is to investigate the dynamics of living standards in Thailand. In this section we have established that real equivalised household income and earnings are the living standards of interest. Analysis of both mobility and persistence in these living standards will be performed in an examination of income dynamics in Thailand.

Mobility analysis will be based on the construction of transition matrices. Given the data limitation, the transition matrix and, hence, income mobility can only be directly investigated from the Townsend panel data. Measures of persistence based on the regression framework of the statistical model of earnings dynamics will be calculated for both the Townsend and LFS data.

The next section provides detailed discussion of the statistical model of earnings dynamics investigated in this chapter as well as the identification and the econometric specifications employed to obtain consistent estimates of income and earnings persistence in this process. In addition, an alternative "permanent earnings model" which involves the decomposition of earnings into permanent and transitory components will be examined to provide a more complete picture of the dynamics of earnings in Thailand.

3.3 Identification, Econometric Specifications and Estimation

In this section we discuss the identification, econometric specifications and methodology used to obtain consistent estimates of income and earnings persistence in Thailand. We first present the identification and estimation of persistence based on the Galton-Markov framework for standard and pseudo-panel data, dealing with the problems of measurement error, attrition and the non-representativeness of the Townsend Thai panel sample for the Thai labour force. We then proceed to the permanent income model, where we attempt to identify and estimate persistence as well as parameters of the variance decomposition of income and earnings into permanent and transitory components.

3.3.1 Statistical Models of Earnings Dynamics

In this subsection, we focus on one of the simplest but most intuitive measures of earnings persistence based on a regression framework: the slope coefficient in a regression of earnings on its lagged value.

3.3.1.1 The Galton-Markov Earnings Model

A simple statistical model of earnings dynamics is the first-order autoregressive model (also sometimes referred to as the Galton model). The data generating process for earnings dynamics in this model is assumed to be given by Equation (3.3).

$$Y_{i,t} = \alpha + \beta Y_{i,t-1} + \epsilon_{i,t} \tag{3.3}$$

Where $Y_{i,t}$ represents earnings (either in a monetary unit or log scale) of household i at time t. The model embodies a number of assumptions:

1. The stochastic error term $\epsilon_{i,t}$ is assumed to be independent of $Y_{i,t-1}$.

2. The stochastic error term $\epsilon_{i,t}$ is distributed independently and identically across individuals i and periods t.

3. There is population homogeneity and time homogeneity. This means that the parameters α and β of the earnings process are constant across individuals and across time.

4. Earnings at date t are independent of earnings in any period before $t-1$.

These assumptions allow the parameter β to be estimated consistently by OLS. Assumption (4) in particular is the first-order Markov assumption where the value of earning at time period $t-1$ contains all the relevant information from the individual's earnings history. Given that earnings satisfy this Markov property, the statistical model above is often referred to as the Galton-Markov model (Stewart, 2010).

The main parameter of interest in this model is the slope coefficient β, a measure of the degree of persistency in earnings; this demonstrates how change in lagged earnings relates with the change in current earnings.

$$\triangle E(Y_{i,t}|Y_{i,t-1}) = \beta \cdot \triangle Y_{i,t-1} \tag{3.4}$$

From Equation (3.4), a unit change in lag earnings will lead to a β unit change in expected current earnings. A value of $\beta = 1$ would indicate that an expected change in current period earnings is highly persistent and perfectly dependent on the previous period's change in earnings. In this case individuals tend to hold their relative earnings positions except for purely random shifts, the size of which depends on $\epsilon_{i,t}$; this is a random walk process. This point is where earnings persistence is greatest. At the other extreme, $\beta = 0$ indicates that the expected current period earnings change is completely unrelated to the previous period's earnings change. This point is where earnings persistence is at minimum. A typical value of β is likely to be between these two extremes ($0 < \beta < 1$), where a value of β closer to zero would suggest a lower degree of persistence in earnings.

Two special cases arises when $\beta > 1$ and $\beta < 0$. $\beta > 1$ suggests that the expected current earnings change will occur at a greater magnitude than the previous period's change. This can be thought of as indicating divergence in the earnings distribution. Alternatively, $\beta < 0$ suggests a certain degree of reversal as those whose changes in earnings are greatest in the last period are now experiencing the most negative changes in earnings in the current period.

It should be noted that the degree of persistence estimated here tells us nothing about where the individual has moved. The only context where the first-order Galton-Markov regression has been used to indicate movement within the earnings distribution is when it is interpreted from the mean reversion perspective. However, as will be shown later in Subsection 3.3.1.3, this interpretation is incorrect as mean reversion is purely a statistical

artefact. Estimates of the coefficient β must only be taken to indicate persistence, or the correlation of earnings with its lag over time, as discussed above.

Lastly, the socially optimal value of β will involves a trade-off between the degree of aversion to inequality and the degree of aversion to inter-temporal fluctuation in earnings (Gottschalk and Spolaore, 2002). The former places positive value on reversal within the earnings distribution that promotes equality in lifetime earnings and favours a lower value of β. The latter places positive value on earnings stability as the uncertainties that accompany frequent random fluctuations in earnings can lead to a loss of welfare, and favours a value of β closer to one.

3.3.1.2 Unconditional and Conditional Persistence

There are two main parameters of interest that we will attempt to estimate in this chapter, unconditional and conditional persistence.

Unconditional persistence measures correlation between current earnings and past earnings without explicitly holding other things equal. It is a measure of overall persistence in the economy for everyone regardless of their individual characteristic, whether or not ones are similar. The purpose of this measure is simply to see who is doing better between the initially rich and the initially poor. Here unconditional persistence is not a biased estimator of conditional persistence, but is another parameter that we are interested in estimating in its own right.

$$Y_{i,t} = \alpha + \beta^U Y_{i,t-1} + \epsilon_{i,t} \tag{3.5}$$

Estimation of Equation (3.5) gives the coefficient $\hat{\beta}^U$ which determine the estimated degree of "unconditional persistence". Conceptually, there are no issues concerning omitted variable bias with unconditional persistence estimation as these variables are left in the error term on purpose.

Conversely, conditional persistence measures the correlation of the earnings over time among individuals with similar characteristics.

$$Y_{i,t} = \alpha + \beta^C Y_{i,t-1} + \gamma X_{i,t} + \varphi_{i,t} \tag{3.6}$$

Equation (3.6), where $X_{i,t}$ is a set of control variables such as gender, education level or geographic location, extends the basic Galton-Markov model to estimate conditional persistence. Here the interest is in whether the initially rich have more or less persistence in earnings than the initially poor with similar characteristics such as education, gender etc. The estimate of β^C in Equation (3.6) gives $\hat{\beta}^C$, a measure of "conditional persistence". The error term $\varphi_{i,t}$ in Equation (3.6) is not the same as $\epsilon_{i,t}$ unless $\gamma = 0$. Omitting relevent and serially correlated explanatory variables will cause the error term $\epsilon_{i,t}$ to be serially correlated which will usually lead to inconsistent estimation of β^U in Equation (3.5). However, as noted earlier that omitted variables are left in the error term on purpose in the estimation of the unconditional persistence parameter, this serial correlation will not be a concern in this case.

A society with high level of unconditional but low conditional persistence offers some hope as there is a scope for individuals to obtain the specific characteristic, such as education, and improve the prospect of moving up the social ladder unless the relevant conditional factors are individual-specific and cannot be replicate by others. One of the main objectives of this study to obtain consistent estimations of the parameter β^U and β^C.

3.3.1.3 Galton's Fallacy of Regression towards the Mean

The parameter β of the Galton-Markov model has been claimed by various researchers[2] to also indicates the degree of convergence or divergence of earnings between the initially rich and the initially poor. In this subsection we illustrate that this interpretation of β suffers greatly from Galton's classical fallacy of regression towards the mean (Friedman, 1992; Quah, 1993). We show that observing $\beta < 1$ does not convey any information on the convergence of the earnings distribution, and suggest an alternative method to study the change in the shape of the conditional earnings distribution.

The β parameter of the Galton-Markov model is frequently interpreted to be a measure of the extent to which convergence in earnings takes place. A value of $0 < \beta < 1$ was claimed to suggest that there is convergence while $\beta > 0$ suggests a divergence in earnings between the rich and the poor. While the term "convergence" can mean a number of different things, these interpretations of β would imply that the term "convergence" used here describe a process whereby poorer individuals are catching up with richer individuals – that the cross-section dispersion of the earning distribution diminishes over time. Below we formally illustrate the

[2] e.g. Antman and McKenzie (2008), Fields et al (2003).

flaws with this type of argument.

Case I: Stationary Distribution

Let us first assume that the earnings distribution is stationary. This implies that

- $E[Y_{i,t}] = E[Y_{i,t-1}] = \mu$
- $Var[Y_{i,t}] = Var[Y_{i,t-1}]$

Equations (3.7) and (3.8) give the OLS estimators of α and β from Equation (3.3) respectively.

$$\hat{\alpha} = E[Y_{i,t}] - \hat{\beta}E[Y_{i,t-1}] \tag{3.7}$$

$$\hat{\beta} = \frac{COV(Y_{i,t}, Y_{i,t-1})}{\sqrt{Var[Y_{i,t-1}]}\sqrt{Var[Y_{i,t-1}]}} \tag{3.8}$$

Given the stationarity assumption above ($Var[Y_{i,t}] = Var[Y_{i,t-1}]$), we can rewrite $\hat{\beta}$ in Equation (3.4) as

$$\hat{\beta} = \frac{COV(Y_{i,t}, Y_{i,t-1})}{\sqrt{Var[Y_{i,t}]}\sqrt{Var[Y_{i,t-1}]}} = \rho \tag{3.9}$$

Where ρ is the correlation coefficient between $Y_{i,t}$ and $Y_{i,t-1}$, whose value lies between $-1 \leq \rho \leq 1$.

Substituting the values of α and β into Equation (3.3) gives

$$E[Y_{i,t}|Y_{i,t-1}] = E[Y_{i,t}] - \hat{\beta}[Y_{i,t-1}] + \hat{\beta}E[Y_{i,t-1}|Y_{i,t-1}] \tag{3.10}$$

$$E[Y_{i,t}|Y_{i,t-1}] = \mu - \rho\mu + \rho Y_{i,t-1} = (1-\rho)\mu + \rho Y_{i,t-1} \tag{3.11}$$

Equation (3.11) shows that expected current earnings are a weighted average between the mean μ and the lagged earnings $Y_{i,t-1}$. A value of $\beta = 1$ ($\rho = 1$) indicates that expected earnings in the current period are completely determined by, and hence perfectly dependent on, previous period earnings ($E[Y_{i,t}|Y_{i,t-1}] = Y_{i,t-1}$). On the other hand, $\beta < 1$ ($\rho < 1$) indicates some degree of mean reversion in expected earnings over time, where the earnings

gap between individuals with earnings below and above the average earnings will become narrower over time as expected earnings for both groups are moving closer to the mean. In the extreme, the value of $\beta = 0$ ($\rho = 0$) indicates that expected current period earnings will be equal to the mean regardless of the level of previous earnings. Lower persistence interpreted in this manner would, therefore, suggest that expected lifetime earnings will be more equal over time as individuals move closer to the population mean earnings.

However, we have made a restrictive assumption that the earnings distribution is stationary over time ($Var[Y_{i,t}] = Var[Y_{i,t-1}]$), so that there can be no convergence in the sense that the dispersion of the earnings distribution is being reduced over time. The fact that we still obtain an estimate of $\hat{\beta} < 1$ is a direct consequence of the stationarity assumption made earlier, which ensures that the value of $\beta = \rho$ will always be less than 1 except in the case of perfect correlation in earnings. The estimate of β, interpreted as indicating the degree of mean reversion as used above, is purely a statistical artefact of the stationarity assumption. Therefore, in this case, obtaining a regression estimate of $\hat{\beta} < 1$ would tell us nothing about the convergence or divergence of the earnings distribution. To wrongly infer that the population earnings distribution is narrowing and that individuals' earnings are converging towards each other as a result of obtaining an estimate of $\hat{\beta} < 1$ is to commit Galton's fallacy.

Case II: Non-stationary Distribution

Relaxing the stationarity assumption to allow both the mean and the variance of the earnings distribution to change over time does not affect the Galton's fallacy result obtained earlier. In this case, we can rewrite $\hat{\beta}$ as

$$\hat{\beta} = \rho \frac{\sigma_{Y_t}}{\sigma_{Y_{t-1}}}$$

Where σ_{Y_t} and $\sigma_{Y_{t-1}}$ are the standard deviations of the current and previous period earnings distributions respectively. Here, it is entirely possible to obtain an estimate of $\hat{\beta} < 1$, even when $\sigma_{Y_t} > \sigma_{Y_{t-1}}$, which would actually indicate that the earnings distribution is diverging over time (this is possible if the value of ρ is small enough). As before, a value of $\hat{\beta} < 1$ provides no information about convergence in the earnings distribution.

However, the same is not true of the process of divergence. As $\rho \leq 1$, it is only possible

to observe $\beta > 1$ as a result of an increase in the spread of the earnings distribution ($\sigma_{Y_t} > \sigma_{y_{t-1}}$). A value of $\beta > 1$ indicates divergence causing the widening of the earnings gap, as t increases. The "reversal" result discussed earlier ($\beta < 0$) also holds as it only depends on observing a negative relationship $\rho < 0$, since $\frac{\sigma_{Y_t}}{\sigma_{Y_{t-1}}} > 0$ always. This reversal interpretation makes no claim on indicating the change in shape of the earnings distribution.

We have shown here that earlier interpretations of the parameter β used in intra-generational dynamics studies are plagued by Galton' fallacy. The value of β must only be interpreted as indicating the degree of persistence in earnings as before, and no inference on the shape of the distribution can be drawn from observing the value of $\hat{\beta} < 1$ alone.

To investigate the change in the shape of the earnings distribution it is necessary to move beyond the least squares estimation of the Galton-Markov earnings model that only focuses on the conditional mean, which represents just one aspect of the earnings distribution. In this chapter we propose the use of quantile regression as a method to determine the degree of persistence at different parts of the earnings distribution and also to investigate the change in the shape of the earnings distribution, to establish whether it has converged or diverged.

3.3.1.4 Quantile Regression of the Galton-Markov Earnings Model

In this subsection, we discuss in detail the method of quantile regression (Koenker and Bassett, 1978), which is used to extend the analysis of income and earnings dynamics to comprehensively study the earnings distribution.

Our earlier estimates of earnings persistence from the Galton-Markov model are estimates of "average" inter-temporal earnings persistence. By estimating the conditional average using the least squares technique, we made a number of assumptions, most notably assumption (3) in Subsection 3.3.1.1, that β is constant across all individuals. This implies that the degree of earnings persistency is identical across the entire conditional earnings distribution and the parameter β is a single measure of the dynamics of the earnings process. However, the inter-temporal linkage of individual household earnings may differ depending on the position each individual household has in the conditional earnings distribution. If this non-uniformity exists across the earnings distribution, studies that only concentrate on the persistence of the means of the distribution will not show a complete picture of earnings persistence. Therefore, a relaxation of the population homogeneity assumption made earlier is necessary.

Quantile regression provides a framework to study income persistence at different parts of the income distribution. As we have argued in Subsection 3.3.1.3, measures of income and earning persistence based on the OLS framework, which only concentrates on the mean, can never truly indicate the degree of convergence or divergence of the income distribution. Quantile regression makes it possible to measure the effect of the explanatory variables at any point in the conditional distribution of the dependent variable. Therefore, by estimating the degree of persistence at various points of the distribution we will be able to observe and make an accurate statement about the changing shape of the conditional distribution.

The starting point for quantile regression is the condition quantile function (CQF). Let Y be a continuously distributed random variable with a well-behaved density, then the CQF at quantile τ given a vector of regressors, $X_{i,t}$, is defined as

$$Q_Y[\tau|X_{i,t}] = F_Y^{-1}(\tau|X_{i,t}) \tag{3.12}$$

In this case the dependent variable, Y, is either current household income or earnings ($Y_{i,t}$). $X_{i,t}$ consists of any relevant independent variables but the one that will be of main interest is one-period-lagged earnings or income, $Y_{i,t-1}$. $F_Y(y_{i,t}|X_{i,t})$ is the distribution function for Y at $y_{i,t}$ conditioned on $X_{i,t}$. When $\tau = 0.25$, for example, $Q_Y[\tau|X_{i,t}]$ gives us the conditional bottom quartile of Y given $X_{i,t}$, while $\tau = 0.5$ gives the conditional median.

The CQF is the conditional quantile version of the conditional expectation function (CEF). While the CEF can be derived as the solution to the familiar mean-squared error problem as shown in Equation (3.13), Koenker and Bassett (1978) showed that the CQF is the solution to the minimisation problem in Equation (3.14).

$$E[Y_{i,t}|X_{i,t}] = \underset{m(X_{i,t})}{arg\,min}\, E[\{Y_{i,t} - m(X_{i,t})\}^2] \tag{3.13}$$

$$Q_{Y_{i.t}}[\tau|X_{i,t}] = \underset{q(X_{i,t})}{arg\,min}\, E[\rho_\tau(Y_{i,t} - q(X_{i,t})] \tag{3.14}$$

Where $\rho_\tau(u)$ is called the "check function" and is given by Equation (3.15) below.

$$\rho_\tau(u) = 1(u > 0) \cdot \tau|u| \; + \; 1(u \leq 0) \cdot (1 - \tau)|u| \tag{3.15}$$

The key idea in their methodology implies replacing the process of ordering and sorting

sample observations with the process of optimisation. If $\tau = 0.5$, Equation (3.14) is the least absolute deviations as $\rho_{0.5}(u) = \frac{1}{2}|u|$ and $Q_{Y_{i.t}}[\tau = 0.5|X_{i,t}]$ is the conditional median as this minimises absolute deviations. For other values of $\tau \, \epsilon \, [0, 1]$, the check function weights positive and negative terms asymmetrically and it is this asymmetric weighting that generates a minimand that gives the τth conditional quantile as the solution.

With continuous or high-dimensional $X_{i,t}$, the CQF may be hard to estimate and summarise. It is therefore a common practice, like in the case of the linear CEF, to substitute a linear model for $q(X_{i,t})$ in Equation (3.14) in order to reduce the function to a smaller set of numbers, just one for each $X_{i,t}$. This gives

$$\beta(\tau) \equiv \underset{b}{arg\,min}\, E[\rho_\tau(Y_{i,t} - X_{i,t}^{'} b)] \tag{3.16}$$

$\beta(\tau)$ is an estimator of the linear population conditional quantile regression. $\hat{\beta}(\tau)$ is the corresponding sample analogue ($\hat{\beta}(\tau) \equiv \underset{b}{arg\,min}\, \Sigma_{i=1}^{n}[\rho_\tau(Y_{i,t} - X_{i,t}^{'} b)]$). The above minimisation problem can be reformulated and solved efficiently by linear programming methods.

Here we are interested in estimating the degree of persistence which was represented as a linear first-order autoregressive Galton-Markov model of income or earnings as shown in Equation (3.3) and restated in Equation (3.17) below.

$$Y_{i,t} = \beta Y_{i,t-1} + \epsilon_{i,t} \tag{3.17}$$

The least squares method employed earlier to a model of this type only models the conditional mean, which represents only one aspect of the income and earnings distribution. The quantile equivalence of Equation (3.17) above is given by

$$Q_{Y_{i,t}}[\tau|Y_{i,t-1}] = \beta_0(\tau) \, + \, \beta_1(\tau) Y_{i,t-1} \tag{3.18}$$

$$Q_{Y_{i,t}}[\tau|Y_{i,t-1}, X_{i,t}] = \beta_0(\tau) \, + \, \beta_1(\tau) Y_{i,t-1} + \beta_2(\tau) X_{i,t} \tag{3.19}$$

Where Equation (3.18) represents an unconditional quantile persistence regression, while Equation (3.19) represents the conditional quantile persistence regression. These quantile regressions enable the investigation of whether households at different parts of the conditional distribution experience a different degree of earnings persistence. $\beta(\tau)$, the slope coefficient

of the τth quantile regression line, shows how the value of current income, $Y_{i,t}$, changes with lagged income, $Y_{i,t-1}$, along the τth percentile of the distribution of values $Y_{i,t}$ at each value of $Y_{i,t-1}$. It is important to note that while the quantile estimation of the dynamics of earnings involves the relaxation of the population homogeneity assumption made earlier in Subsection 3.3.1.1, the first-order Markov property of the Galton-Markov model still holds at each quantile.

In addition, by estimating sufficiently many quantile regression coefficients, we can determine the shape of the conditional earnings distribution and investigate the change between the periods t and $t-1$. In this sense, we will be able to say with a degree of certainty whether there has been a convergence or divergence of the earnings distribution. Estimates of $\beta(t)$ that exhibit a spread in their values as we move from the bottom to the top quantile indicate divergence in the conditional earnings distribution. Alternatively, estimates of $\beta(t)$ that exhibit a collapse in their values as we move from the bottom to the top quantile would indicate convergence in the conditional earnings distribution. Therefore, it is much more comprehensive to estimate a range of quantile regressions instead of only concentrating on the mean, which only gives a partial picture of the shape of the conditional earnings distribution.

3.3.2 Estimation of the Galton-Markov Earnings Model

In this subsection we present the econometric specifications and the identifying assumptions made to ensure that the parameter β, a measure of earnings persistence, can be consistently estimated for both the conditional mean and the conditional quantile regression of the Galton-Markov earnings model.

3.3.2.1 Estimating the Unconditional and Conditional Earnings Persistence Regression

The unconditional Galton-Markov regression Equation (3.3) and its quantile counterpart are restated below.

$$E[Y_{i,t}|Y_{i,t-1}] = \alpha + \beta^U Y_{i,t-1} \tag{3.20}$$

$$Q_{Y_{i,t}}[\tau|Y_{i,t-1}] = \beta_0(\tau) + \beta_1^U(\tau)Y_{i,t-1} \tag{3.21}$$

The unconditional persistence parameter β^U in Equation (3.20) can be consistently estimated using OLS methods on the assumption that the error term $\epsilon_{i,t}$ is independent, identically distributed (iid) with a mean of zero and uncorrelated with lagged earnings $E[\epsilon_{i,t}|Y_{i,t-1}] = 0$. Similarly, $\beta_1^U(\tau)$, the unconditional quantile persistence parameter in Equation (3.21), can be consistently estimated if the error term is uncorrelated with lagged earnings at each quantile. As stated earlier, omitted variable bias will not be a concern when estimating the unconditional persistence regressions above as omitted variables are left in the error term on purpose. $\hat{\beta}^U$ and $\hat{\beta}_1^U(\tau)$ determine the degree of persistence based solely on last period earnings without holding any characteristics constant.

$$E[Y_{i,t}|Y_{i,t-1}, X_{i,t}] = \alpha + \beta^C Y_{i,t-1} + \gamma X_{i,t} \tag{3.22}$$

$$Q_{Y_{i,t}}[\tau|Y_{i,t-1}, X_{i,t}] = \beta_0(\tau) + \beta_1^C(\tau)Y_{i,t-1} + \beta_2(\tau)X_{i,t} \tag{3.23}$$

With the conditional persistence regression and its quantile counterpart restated above in Equation (3.22) and (3.23) respectively, we examine persistence conditional on a set of co-variates $X_{i,t}$. $X_{i,t}$ here consists of control variables for household demographic structure, education level, sector of employment and year effect. Equation (3.22) can be consistently estimated by OLS where the identifying assumption of conditional independence, $E[\varphi_{i,t}|Y_{i,t-1},X_{i,t}] = 0$, is needed. This assumption is that having controlled for a set of co-variates, the error term is now exogenous and uncorrelated with lagged earnings ($Y_{i,t-1}$). A similar conditional independence assumption at each quantile is required to obtain consistent estimates of the conditional quantile persistence parameter $\beta_1^C(\tau)$ (Koenker and Bassett, 1978).

The conditional independence assumption is unlikely to hold in practice as there are other factors, observed and unobserved, which are omitted from the regression that will cause omitted variable bias. There are also several other problems that will cause violations of the identifying assumption above and compromise the consistency of the final persistence estimates of both the mean and the quantile regressions. We discuss these in detail in the

next subsection.

3.3.2.2 Problems with the Basic Identifying Assumptions

Autocorrelation in the Error Terms

The exogeneity assumption for both the unconditional and the conditional regression will not hold when the unobserved shocks to earnings ($\epsilon_{i,t}$) are autocorrelated. For example, unobserved negative shocks to the labour market that affect wages could be persistent and affect earnings in many periods ($E[\epsilon_{i,t}, \epsilon_{i,s}] \neq 0$, $E[\varphi_{i,t}, \varphi_{i,s}] \neq 0$ for $s \neq t$), causing $E[\epsilon_{i,t}|Y_{i,t-1}] \neq 0$ and $E[\varphi_{i,t}|Y_{i,t-1,}X_{i,t}] \neq 0$, which will result in an inconsistent estimation of β^U and β^C respectively.

The issue of correlated errors in the Galton-Markov model will not be dealt with here and we maintain our assumption (2) from Subsection 3.3.1.1 that the error term is iid for the remainder of this chapter.

Omitted Variable Bias

Despite ruling out persistence in the error term there are still other problems that prevent β^U and β^C from being consistently estimated. As stated earlier, omitted variable bias presents a problem for consistent estimation of conditional persistence β^C. While observable omitted variables can normally be dealt with easily by including such variables directly in the regression, some relevant observable variables cannot be included all at once without causing an identification problem. In this case, age, period and cohort (APC) are all independently relevant control variables of the earnings dynamics process. However, these three variables cannot be included in the regression at the same time because of the perfect multicollinearity problem presented by the fact that all three variables are a perfect linear combination of each other. Therefore, the bias remains until all relevant variables can be effectively controlled for. We propose to deal with this problem by utilising the maximum entropy methods that will allow for all three effects to be accounted for in the same regression.

An additional problem is the existence of unobserved omitted variables in the error term that are correlated with past earnings, the explanatory variable in this case, such as individual ability and motivation. Not being able to explicitly control for these effects will result in the estimate of conditional persistence $\hat{\beta}^C$ being biased. In the presence of the

unobserved effect, the direction of the bias is ambiguous and depends on the nature of the correlation between lagged earnings and the unobserved effect. However, if we are willing to make a plausible assumption that these omitted unobservables are fixed over time, we can easily deal with this issue in a standard panel data framework.

Measurement Error

Perhaps the biggest problem faced by empirical works on income and earnings dynamics is the issue of measurement error. Earnings, especially those recorded in surveys in developing countries, are often susceptible to errors in measurement. Errors in the earnings variable will result in an inconsistent estimation of both the unconditional β^U and conditional β^C persistence parameter.

We present a simple illustration of why measurement errors in the regressor will cause a problem:

$$Y_{i,t} = Y^*_{i,t} + \nu_{i,t} \tag{3.24}$$

$$\begin{aligned} Y^*_{i,t} &= \beta Y^*_{i,t-1} + \epsilon_{i,t} \\ Y_{i,t} - \nu_{i,t} &= \beta Y_{i,t-1} - \beta\nu_{i,t-1} + \epsilon_{i,t} \\ Y_{i,t} &= \beta Y_{i,t-1} + \eta_{i,t}; \qquad where\, \eta_{i,t} = \epsilon_{i,t} + \nu_{i,t} - \beta\nu_{i,t-1} \end{aligned} \tag{3.25}$$

Where $Y_{i,t}$ in Equation (3.24) are earnings which are measured with errors and $Y^*_{i,t}$ are the true earnings. Equation (3.24) assumes classical measurement error, where the error occurs in an additive fashion, and where $E[\epsilon_{i,t}|\nu_{i,t}] = 0$, $E[Y^*_{i,t}|\nu_{i,t}] = 0$, $Var(\nu_{i,t}, \nu_{i,t-1}) = 0$, $Var(\nu_{i,t}) = \sigma^2_\nu$ and $Var(\nu_{i,t-1}) = \sigma^2_{\nu,-1}$. Recall Equation (3.3) which represents the data generating process with true earnings $Y^*_{i,t}$ and $Y^*_{i,t-1}$, that are measured without errors in variables, as regressors. We cannot perform regressions on Equation (3.3) as we never observed $Y^*_{i,t}$ or $Y^*_{i,t-1}$ in practice. Using Equation (3.24) to substitute for $Y^*_{i,t}$ and $Y^*_{i,t-1}$ in Equation (3.3) gives Equation (3.25) above. Running a pooled OLS regression on Equation (3.25) would yield an estimator $\hat{\beta}$ in the following form:

$$
\begin{aligned}
plim\hat{\beta}^{POLS} &= \beta + \frac{COV(Y_{i,t-1}, \epsilon_{i,t} + \nu_{i,t} - \beta\nu_{i,t-1})}{VAR(Y_{i,t-1})} \\
&= \beta + \frac{COV(Y^*_{i,t-1} + \nu_{it-1}, \epsilon_{i,t} + \nu_{i,t} - \beta\nu_{i,t-1})}{VAR(Y_{i,t-1})} \\
&= \beta - \frac{\beta VAR(\nu_{i,t-1})}{VAR(Y_{i,t-1})} = \beta[1 - \frac{\sigma^2_{\nu,-1}}{VAR(Y_{it-1})}] \\
&= \beta[\frac{VAR(Y_{it-1}) - \sigma^2_{\nu-1}}{VAR(Y_{it-1})}] = \beta[\frac{VAR(Y^*_{it-1} + \nu_{i,t-1}) - \sigma^2_{\nu,-1}}{VAR(Y^*_{it-1} + \nu_{i,t-1})}] \\
&= \beta[\frac{\sigma^2_{Y*} + \sigma^2_{v,-1} - \sigma^2_{\nu,-1}}{\sigma^2_{Y*} + \sigma^2_{v,-1}}] = \beta[\frac{\sigma^2_{Y*}}{\sigma^2_{Y*} + \sigma^2_{v,-1}}]
\end{aligned}
\tag{3.26}
$$

This illustrates that classical measurement error will produce attenuation bias towards zero in an estimate of β, which in this context means that degree of earnings persistence is being understated.[3] It is important to note that the attenuation bias result obtained above rests on a key assumption that $COV(Y^*_{i,t-1}, \epsilon_{i,t}) = 0$. This assumption will be violated if the error term $\epsilon_{i,t}$ consists of unobserved effects that are correlated with true lagged earnings as discussed earlier.

The problem of measurement error affects the estimates of quantile regression as well as those from the OLS framework. Montes-Rojas (2011) derives the bias of the quantile regression estimator in the presence of classical additive measurement error, and shows its connection to least squares models. Classical errors in variables will also lead to bias in the quantile regression framework that also depends on the noise to signal ratio, thus resulting in attenuation bias as in the least squares case. However, in the quantile regression case, this bias is weighted by the conditional density function of the dependent and independent variables. In this chapter, we will not attempt to fix the problem of measurement errors for the quantile regression estimates of standard panel data.

Attrition

Attrition is also a concern for estimates obtained from standard panel data. Non-random attrition, where households with certain characteristics that are correlated with the outcome of interest systematically leave the survey, will cause bias in the estimates. For example, households with high geographic mobility are those more likely to leave the survey and it is possible that these households are the ones with lower earnings persistence as they move to seek better employment opportunity elsewhere. Not being able to track these households will leave a biased sample in the final panel, which leads to biased estimates of earnings

[3] This attenuation bias towards zero also occurs similarly in conditional persistence regression even though the derivation is not shown here.

persistence. In this chapter, we will not attempt to fix the problem of attrition bias in standard panel data, as we do not possess the necessary exogenous instrument for doing so.

In the subsequent subsections we discuss various methods that provide possible solutions to the problems presented in this subsection in greater detail.

3.3.2.3 Age, Period and Cohort: Maximum Entropy Estimation

In this subsection, we discuss in detail the principle behind the maximum entropy approach proposed by Browning, Crawford and Knoef (2012), which is used in the decomposition of real equivalised earnings and income in to APC effects, and also in the estimation of the degree of persistence while controlling for APC effects all at once thus overcoming the perfect multicollinearity problem discussed earlier.

Age, period and cohort are three distinct concepts that affect the evolution of earnings and income. Different age groups within the population are at different stages of life in relation to their work and family formation. Young individuals who only recently entered the labour market will have different levels of earnings to those who are in the middle of their working life despite having similar characteristics. At various time periods, individuals are also exposed to different events, which have widespread effects such as macroeconomic shocks. Individuals who were born in different birth cohorts experience different histories and institutions that may also affect their lifetime earnings and income.

While these three effects are important and conceptually independent in their own rights, the standard linear model that attempts to identify the APC effects all at once suffers from a well-known perfect multicollinearity problem that prevents unique point identification of each separate effect. This occurs because there is a perfect linear relationship among these three effects: age plus year of birth equals the year in which the individual is observed in the survey, i.e. the period. Following Browning et al. (2012) we formally express this problem and the solution, the maximum entropy approach, employed.

In a panel data setting, we observe individual i in period t, whose age is denoted $a_{i,t}$. Given these information, a year of birth cohort variable can be calculated, $c_{i,t} = t - a_{i,t}$. We also observe individual earnings or income $Y_{i,t}$, our variable of interest in this case. Later in Section 3.4, we will be interested in decomposing these individuals' earnings into APC component. These components are specified as additively decomposable and enter the

regression in a less restrictive non-parametric fashion as in Equation (3.27) below.

$$Y_{i,t} = \sum_{a} \alpha_a d_{i,t}^{a} + \sum_{t} \lambda_t d_t + \sum_{c} \gamma_c d_{i,t}^{c} + \epsilon_{i,t} \tag{3.27}$$

Here, $d_{i,t}^{a}$ is a dummy variable that takes the value one if individual i is of age a at the end of survey year t. Similarly, $d_{i,t}^{c}$ is a dummy variable that takes the value one if individual i was born in year c, and d_t is a dummy variable that takes the value one if individual i earnings, $Y_{i,t}$, are observed in year t. α_a, λ_t and γ_c are the corresponding age, period and cohort coefficients and the summations in Equation (3.27) are taken over all possible values of the three variables.

Considering a simple example to illustrate the problem without loss of generality we focussed on just two cohorts ($c = 1, 2$), which are surveyed for three periods ($t = 3, 4, 5$) so that ages range from 1 to 4 ($a = 1, 2, 3, 4$). We have

$$Y_{i,t} = \alpha_1 d_{i,t}^{a=1} + \alpha_2 d_{i,t}^{a=2} + \alpha_3 d_{i,t}^{a=3} + \alpha_4 d_{i,t}^{a=4} + \lambda_3 d_3 + \lambda_4 d_4 + \lambda_5 d_5 + \gamma_1 d_{i,t}^{c=1} + \gamma_2 d_{i,t}^{c=2} + \epsilon_{i,t} \tag{3.28}$$

Equation (3.28) cannot be estimated as it faced two identification problems. First, each set of age period and cohort dummy variables sums up to one, which leads to the well-known perfect multicollinearity problem. This can be dealt with by setting the coefficients of any two dummy variables from a different set of dummies to zero and interpreting the remaining coefficients relative to this reference. For example, we could set the coefficient for the youngest age and the most recent cohort (α_1, γ_2) to zero, which would make the reference group an individual who was born in period 2 (second cohort) and observed in year 3, at the age of 1.

$$Y_{i,t} = \alpha_2 d_{i,t}^{a=2} + \alpha_3 d_{i,t}^{a=3} + \alpha_4 d_{i,t}^{a=4} + \lambda_3 d_3 + \lambda_4 d_4 + \lambda_5 d_5 + \gamma_1 d_{i,t}^{c=1} + \epsilon_{i,t} \tag{3.29}$$

Equation (3.29) sets (α_1, γ_2) to zero as discussed above. However, even with this normalisation, Equation (3.29) is not identified as we still have another perfect multicollinearity problem arising from the linear relationship among age, period and cohort: age plus year of birth equal period as discussed earlier. This imposes a linear relationship on the dummy

variables, which in this example can be formally represented as

$$-d_{i,t}^{a=2} - 2d_{i,t}^{a=3} - 3d_{i,t}^{a=4} + d_4 + 2d_5 + d_{i,t}^{c=1} = 0 \tag{3.30}$$

Given the relationship in (3.30), the matrix of the regressors in Equation (3.30) does not have full column rank and is not invertible and therefore, cannot be estimated using OLS method. Solving for $d_{i,t}^{a=2}$ in (3.30) ($d_{i,t}^{a=2} = -2d_{i,t}^{a=3} - 3d_{i,t}^{a=4} + d_4 + 2d_5 + d_{i,t}^{c=1}$) and substituting this into Equation (3.29) gives the reduced form Equation (3.31) below.

$$Y_{i,t} = b_1 d_{i,t}^{a=3} + b_2 d_{i,t}^{a=4} + b_3 d_3 + b_4 d_4 + b_5 d_5 + b_6 d_{i,t}^{c=1} + \epsilon_{i,t} \tag{3.31}$$

The matrix of the regressors in Equation (3.31) now has full column rank and can be estimated using OLS. The relationship between the reduced form parameter and the original parameter is given by

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \gamma_1 \end{bmatrix} \quad or \;\; b = \mathbf{A}\delta \tag{3.32}$$

This represents an under-determined linear simultaneous equation system with seven unknowns and only six equations. While we can calculate the six reduced form estimates, it is not possible to recover a unique solution for the original seven parameters of interest as there is an infinite number of possible solutions to this system of equations.

Even though the original parameters δ are not point identified, Browning et al. (2012) showed that under the condition that the support of $Y_{i,t}$ is bounded ($Y \in [Y_{min}, Y_{max}]$), the APC coefficient vector δ is set identified and lies in a closed, convex parameter space given by the hypercube $\mathcal{B}$, centred at zero with each side of length $2\,[Y_{max} - Y_{min}]$. While several methods of obtaining the APC point estimates have been proposed in the literature, mostly through making additional assumptions or increasing restrictions on certain parameters to

transform the ill-posed inverse problem in to a well-posed one[4], here the authors utilise an information-based approach using the maximum entropy principle. The identified set given by the hypercube $\mathcal{B}$ can be represented as

$$\mathcal{B} = \{\delta | \delta = \mathbf{S}p\} \tag{3.33}$$

Where $\mathbf{S} = [s_1, s_2, ..., s_J]$, the s_j are vectors representing the vertices of the cube and p is a vector of non-negative weights that sum to one and which are used to form all of the convex combinations of the vertices. Here, δ can be expressed as a convex combination of the extreme points of $\mathcal{B}$; this is possible as a result of the crucial assumption made that the support of $Y_{i,t}$ is bounded. Substituting (3.33) into (3.32) gives

$$b = \mathbf{AS}p \tag{3.34}$$

The problem of solving (3.32) for the original parameter of interest, δ, is now reformulated as a problem of solving (3.34) for p. While the system of equation in (3.34) is still under-determined (as $J > 6$), the problem is now expressed in a way that allows for the use of the maximum entropy principle. Here, p_j have all the characteristics of probabilities, which, in the current context, is interpreted as a measure of the current state of knowledge as opposed to being a limiting frequency. We can treat the p vector as a discrete probability distribution over the J multivariate outcomes represented by the column of matrix $\mathbf{S}$. The solution to (3.34) requires a constructive principle for choosing one probability distribution over another rather than choosing one parameter vector over another.

The principle adopted for choosing a probability distribution is based on Jaynes (1957a; 1957b) and is an extension of LaPlace's principle of insufficient reason in situations in which some information about the problem is available from the data. The main idea is that a distribution that does not unduly favour one outcome over another (a probability distribution that is as flat as possible), subject to some necessary constraints, should be chosen. These constraints are that the probabilities are non-negative, sum to one and satisfy any data-based restriction, which in this case is the reparameterised reduced form estimates in Equation 3.34. An objective function that will achieve this is the entropy function suggested by Shannon (1948) as a measure of uncertainty.

[4]See Browning et al. (2012) for a review.

$$H(p) = -p'ln(p) \tag{3.35}$$

The entropy function (3.35) is maximised when the probabilities are uniformly distributed (all outcomes are equally likely, which is interpreted as being maximally uncertain), and it is minimised if the probability distribution is degenerate on a particular outcome (interpreted as perfect certainty about that outcome). When elements of p have zero probability, $H(p) = 0ln(0) \equiv 0$. The constrained optimisation problem is therefore

$$\underset{p \geq 0,\, 1'p=1}{max} \; -p'ln(p) \; subject\, to\, b = \mathbf{AS}p \tag{3.36}$$

Equation (3.36) is a non-linear optimisation problem with a unique solution p^*, which is a vector of probabilities over support points. The maximum entropy probability distribution p^* is then used to recover the APC coefficients ($\delta^* = \mathbf{S}p^*$), which satisfy the data constraints in (3.32) exactly, and can be interpreted as the expected value of a discrete multidimensional random variable consistent with the entropy maximising choice of the underlying probability distribution.

In addition to the decomposition of earnings into APC components, another interest in this chapter is the estimation of earnings and income persistence controlling for APC effects all at once. The maximum entropy approach discussed above made this estimation possible. Due to a constraint with the APC programming command (O'Dea, 2011), which only permits the inclusion of a dependent variable without allowing for any additional independent variable apart from the APC dummy variables, estimates of the earnings persistence parameter β were obtained via the Frisch-Waugh formula. We first estimate Equation (3.37) below.

$$Y_{i,t} = \kappa + \sum_{a,a\neq 20} \alpha_a d^a_{i,t} + \sum_{t,t\neq 1985} \lambda_t d_t + \sum_{c,c\neq 1941} \gamma_c d^c_{i,t} + \epsilon_{i,t} \tag{3.37}$$

$Y_{i,t}$ is household income (in the case of the Townsend panel data) or cohort average earnings (in the case of the LFS pseudo-panel). Equation (3.37) is a regression of current period earnings or income on age period and cohort dummy variables. As the regression equation also includes an intercept κ, one dummy variable from each set of APC dummies must be set to zero to avoid the perfect multicollinearity problem. Here we set the coefficients of the youngest age (20 years old), the first year of the survey (1985) and the oldest cohort

(1941) to zero. From Equation (3.37), we obtain the predicted residuals $\hat{\epsilon}_{i,t}$, which capture any variations in $Y_{i,t}$ not explained by the APC effect. Next we estimate Equation (3.38)

$$Y_{i,t-1} = \kappa + \sum_{a,a\neq 20,21} \alpha_a d^a_{i,t} + \sum_{t,t\neq 1985,1986} \lambda_t d_t + \sum_{c,c\neq 1941} \gamma_c d^c_{i,t} + \xi_{i,t} \tag{3.38}$$

Equation (3.38) is a regression of lagged earnings or income on age period and cohort dummy variables. As lagged earnings are used, we need to set two further dummy variables from any two of the three sets to zero to avoid the dummy variable trap as before. Here in addition to earlier, the coefficients of the next youngest age (21 years old) and the next year of the survey (1986) are set to zero. Again we obtain the predicted residuals $\hat{\xi}_{i,t}$, which capture any variations in lagged earnings $Y_{i,t-1}$ not explained by the APC effect.

According to the Frisch-Waugh regression anatomy formula, an estimate of β from a regression such as Equation (3.39), which cannot be directly estimated due to the programming constraint mentioned above, can also be obtained by regressing the predicted residuals obtained in (3.37) on those obtained in (3.38) as in Equation (3.40) below.

$$Y_{i,t} = \kappa + \beta Y_{i,t-1} + \sum_{a,a\neq 20} \alpha_a d^a_{i,t} + \sum_{t,t\neq 1985} \lambda_t d_t + \sum_{c,c\neq 1941} \gamma_c d^c_{i,t} + \epsilon_{i,t} \tag{3.39}$$

$$\hat{\varepsilon}_{i,t} = \beta \hat{\xi}_{i,t} + \zeta_{i,t} \tag{3.40}$$

As the earnings persistence estimate $\hat{\beta}$ is obtained using generated regressors ($\hat{\epsilon}_{i,t}$ and $\hat{\xi}_{i,t}$), the standard error obtained for $\hat{\beta}$ from the regression in Equation (3.40) will be incorrect. This will be corrected for by bootstrapping for the standard error[5].

The final estimate of β obtained using the above method is unlikely to be significantly different from that obtained from a simpler treatment of including age, period and cohort as dummy variables and dropping more categories from each set to allow identification. However, employing the APC method has made the simultaneous point identification of the

[5] An alternative Frisch-Waugh regression anatomy formula for obtaining APC $\hat{\beta}$ is

$$\hat{\beta} = \frac{Cov(Y_{i,t}, \hat{\xi}_{i,t})}{Var(\hat{\xi}_{i,t})}$$

Where $\hat{\xi}_{i,t}$ is the predicted residual from a regression of $Y_{i,t-1}$ on all APC covariates as in Equation (3.38). This is equivalent to $\hat{\beta} = \frac{Cov(\hat{\varepsilon}_{i,t}, \hat{\xi}_{i,t})}{Var(\hat{\xi}_{i,t})}$ in Equation (3.39) as the fitted values removed from $Y_{i,t}$ to obtain $\hat{\varepsilon}_{i,t}$, the residual from a regression of $Y_{i,t}$ on APC covariates as in Equation (3.37), are uncorrelated with $\hat{\xi}_{i,t}$.

impact of each age, period and cohort possible. Studying the decomposition of earnings in to these three independent components will help increase the understanding of the the underlying factors behind earnings and income dynamics in Thailand. The APC decomposition of earnings and income are performed later in Section 3.4.

In this chapter we will not be performing the maximum entropy APC regression directly on the quantile regression as there are currently no theoretical findings on how to perform such regression. Alternatively, we will use the result drawn from the APC decomposition of earnings and income to guide our quantile APC estimation strategy.

3.3.2.4 Dealing with Measurement Errors without Individual Fixed Effects

As illustrated in Subsection 3.3.2.2, measurement errors will cause bias in OLS estimates of earnings persistence. The common solution in dealing with measurement errors utilises the instrumental variable (IV) approach. Within the earnings dynamics literature, instruments such as education, lagged expenditure (McCulloch and Baulch, 2000), and anthropometric measures such as body weight (Glewwe and Nguyen, 2002) have been used.

For an instrument to be valid it must satisfy the following condition:

1. Instrument Relevance: $E(Y^{*}_{i,t-1}\,, Z_{i,t}) \neq 0$

   The instrument must be correlated with real lagged earnings that are not measured with error.

2. Instrument Exogeneity: $E(Z_{i,t}\,, \eta_{i,t}) = 0$

   The instrument must not be correlated with the composite error term $\eta_{i,t}$ in Equation (3.25). This means that it must be exogenous with respect to $\epsilon_{i,t}$, the original error term, and it must be uncorrelated with current and lagged measurement errors ($\nu_{i,t}$ and $\nu_{i,t-1}$).

However, in a setting with dynamic model and measurement errors Antman and McKenzie (2007b) have shown that the conditions required for instruments to be valid are more stringent than usual and are unlikely to be met by any of the previous instruments employed in the literature.

Following Antman and McKenzie (2007b), we have an instrument $Z_{i,t-1}$ which is related to the real lagged earnings by the linear relationship $Y^{*}_{i,t-1} = \phi + \gamma Z_{i,t-1} + \xi_{i,t-1}$. This can

be rewritten in terms of observed lagged earnings as $Y_{i,t-1} = \phi + \gamma Z_{i,t-1} + \xi_{i,t-1} + \nu_{i,t-1}$. In addition, the authors assumed that the instrument exhibits the time series properties of the form $Z_{i,t} = \theta + \rho Z_{i,t-1} + \omega_{i,t}$. The IV estimator $\hat{\beta}_{IV}$ from instrumenting Equation (3.25) using $Z_{i,t-1}$ is given by

$$\hat{\beta}_{IV} \xrightarrow{p} \beta + \frac{\gamma(\rho-\beta)Var(Z_{i,t-1}) + E(Z_{i,t-1}\nu_{i,t}) - \beta E(Z_{i,t-1}\nu_{i,t-1}) + \gamma E(Z_{i,t-1}\omega_{i,t}) + E(Z_{i,t-1}\xi_{i,t})}{\gamma Var(Z_{i,t-1}) + E(Z_{i,t-1}\nu_{i,t-1})} \tag{3.41}$$

The conditions laid out in Antman and McKenzie (2007b) for which the instrument must satisfy are restated below.[6]

1. Instrument Relevance: $E(Y^*_{i,t-1}, Z_{i,t-1}) \neq 0$ or $\gamma \neq 0$

2. $E(Z_{i,t-1}, \nu_{i,t}) = 0$ and $E(Z_{i,t-1}, \nu_{i,t-1}) = 0$

   The instrument must be uncorrelated with current and lagged measurement error.

3. $E(Z_{i,t-1}, \omega_{i,t}) = 0$ and $E(Z_{i,t-1}, \xi_{i,t}) = 0$

   The instrument must be uncorrelated with the error terms in the true lagged earnings ($\xi_{i,t-1}$) and instrument dynamic ($\omega_{i,t}$) equations. This condition will be violated if the true lagged earnings, $Y^*_{i,t-1}$, contain individual fixed effects which are correlated with the instrument, or if the dynamic process governing the evolution of the instrument itself contains an individual fixed effect.

4. $\rho = \beta$

   The degree of autocorrelation in the instrument must perfectly match the degree of autocorrelation in earnings.

Conditions [3] and [4] together ensure that the instrument is exogenous from the original error term without the measurement error, $E(Z_{i,t-1}, \epsilon_{i,t}) = 0$. Conditions [2], [3] and [4] combine to form the instrument exogeneity assumption, $E(Z_{i,t-1}, \eta_{i,t}) = 0$. These conditions, especially condition [4], are unlikely to be met by most of the instruments used in the literature thus far and it is not clear if an instrument exists that will satisfy all these conditions simultaneously.

[6] For the derivation of $\hat{\beta}_{IV}$ and these conditions, see Antman and McKenzie (2008)

In addition, if all the conditions above hold except condition [4], the IV estimator will converge in probability to the degree of autocorrelation in the instrument ($\hat{\beta}^{IV} \xrightarrow{p} \rho$). Therefore, using an instrument that exhibits zero correlation over time will cause $\hat{\beta}^{IV}$ to converge to zero while using an instrument that does not vary over time such as education, as we will do later to illustrate this point, will cause $\hat{\beta}^{IV}$ to converge to one. Therefore, measurement error is a major problem for consistent estimation of earnings persistence and it is unlikely that the standard IV framework will help solve this problem.

3.3.2.5 Dealing with Measurement Errors with Individual Fixed Effects

In the last subsection we saw that the presence of an individual fixed effect could exacerbate the measurement error problem by rendering the normal instrumental variable approach invalid. This is due to the violation of condition [3] set out above. In addition, the presence of an individual fixed effect in the error term itself would cause the OLS estimation of conditional persistence to be inconsistent.

Equation (3.3) is rewritten as

$$Y^*_{i,t} = \alpha + \beta Y^*_{i,t-1} + \kappa_i + \mu_{i,t} \tag{3.42}$$

Where the original error term $\epsilon_{i,t}$ in Equation (3.3) is now decomposed into κ_i, a time-invariant unobservable component such as ability or motivation, and $\mu_{i,t}$ the time-varying unobservables. When the individual fixed effect is correlated with true lagged earnings ($E[Y^*_{i,t-1}, \kappa_i] \neq 0$), OLS estimation of Equation (3.42) will yield an inconsistent estimate of β.

Also likely to be problematic is the use of instruments such as expenditures or lagged earnings themselves, which are closely related to earnings, to correct for the measurement error problem as is common in the literature. This is because individual fixed effect that correlates with earnings is likely to be correlated with expenditure as well (and will definitely be correlated with lagged earnings), resulting in the violation of condition [3] set out in the previous section. Given that expenditure and lagged earnings are the few instruments that would most likely satisfy the very stringent condition [4], which requires that the degree of autocorrelation in the instrument match the degree of autocorrelation in earnings, being unable to use this as a result of the presence of an individual fixed effect is a concern.

In the presence of the fixed effect with or without measurement error, the solution that is usually implemented is the Arellano-Bond Generalised Method of Moments (AB) estimator (Arellano and Bond, 1991). This approach begins by taking the first differences (FD) of the estimating equation to remove the individual fixed effect that is correlated with the explanatory variable. Taking the first-differences of Equation (3.42) yields the FD equation below:

$$\Delta Y^*_{i,t} = \beta \Delta Y^*_{i,t-1} + \Delta \mu_{i,t} \tag{3.43}$$

The parameter β still cannot be estimated consistently in Equation (3.43) as $E[\Delta Y^*_{i,t-1}, \Delta \mu_{i,t}] \neq 0$ since $E[Y^*_{i,t-1}, \mu_{i,t-1}] \neq 0$. Without measurement error, consistent estimate of β is obtained by using the lagged level of the explanatory variable such as $Y^*_{i,t-2}$ as an instrument for $\Delta Y^*_{i,t-1}$. This instrument is valid as it satisfies both instrument relevance ($E[\Delta Y^*_{i,t-1}, Y^*_{i,t-2}] \neq 0$) and instrument exogeneity ($E[Y^*_{i,t-2}, \Delta \mu_{i,t}] = 0$) requirements. In this case any lagged earnings of two periods or more are valid instruments ($Y^*_{i,t-s}$; $s = 2, ..., T$). Here we need to have panel data with at least three periods ($T = 3$). Note that by taking the first differences we lost the first period equation and that any observable time-invariant characteristics (e.g. education attainment, sex) are also removed from the estimated equation.

However, with classical measurement error as in Equation (3.24), the method employed above is incorrect. With classical measurement error the FD equation contains a moving average error term as shown below in Equation (3.44).

$$\Delta Y_{i,t} = \beta \Delta Y_{i,t-1} + (\Delta \mu_{i,t} + \Delta \nu_{i,t} - \beta \Delta \nu_{i,t-1}) \tag{3.44}$$

As before the explanatory variable is correlated with the error term. However, instruments that consist of lagged earnings of two periods are no longer valid instruments in this case as $E[Y_{i,t-2}, (\Delta \mu_{i,t} + \Delta \nu_{i,t} - \beta \Delta \nu_{i,t-1})] \neq 0$ since $E[Y_{i,t-2}, \nu_{i,t-2}] \neq 0$. Therefore, in order to obtain a consistent estimate of β, further lagged earnings are needed as instruments. In this case lagged earnings of at least three periods or more ($Y_{i,t-s}$; $s = 3, ..., T$) would be valid instruments. Therefore, we now need to have panel data with at least four periods ($T = 4$) in order to perform the Arellano-Bond method.[7]

[7]However, this assumes classical measurement error where the errors are serially uncorrelated. If this

3.3.2.6 Representativeness

In the last subsection we showed that consistent estimates of earnings persistence from genuine panel data can be obtained even under the presence of unobserved individual fixed effect and classical errors in variables through the use of the AB method. Nevertheless, one major problem remains with AB $\hat{\beta}$, our most preferred estimate of earnings persistence, which is the non-representativeness of the data.

The Townsend Thai data, while being the only long-running panel data set in Thailand with detailed information collected on household income and characteristics that are desirable for estimation of income dynamics, is only collected in 4 out of 76 provinces in Thailand.[8] Furthermore, these provinces are only located in two regions of Thailand, the Central and Northeastern regions (2 provinces per region), which leaves the Northern, Eastern and Southern regions out of the analysis. While this sampling problem is more acceptable in some forms of research than others, it is certainly problematic in this situation, where one seeks to determine the country's degree of earnings persistence.

Therefore, while AB $\hat{\beta}$ is a consistent estimate of earnings persistence, it is only a consistent estimate among a sub-population of Thai people and its external validity remains in question when inference is made from this sub-population to the overall Thai population. This non-representativeness of the Townsend Thai data therefore makes AB $\hat{\beta}$ likely to be an inconsistent estimate of earnings persistence of overall Thai workers.

We therefore propose the use of the dynamic pseudo-panel method to fix this problem. In the next subsection we proceed to show that pseudo-panel estimate not only overcomes the non-representativeness issue as the data is constructed from a long-running and nationally representative cross-sectional data set, but also successfully deals with both measurement errors and attrition, which are often the source of inconsistencies in the estimates of earnings persistence in standard panel data settings.

3.3.3 Pseudo-Panel Data

In many developing countries there is a lack of genuine representative panel data where the same households or individuals are followed over time. However, repeated cross-sectional

assumption does not hold, we will face more difficulties in finding valid instruments.

[8] Thailand now has 77 provinces with the addition of Bungkhan, a new province that was divided from another province only very recently in 2011.

surveys, where different random samples are taken from the population at different points in time, are often more readily available, contain more observations and span a longer period of time. Nevertheless, in this setting, where samples are newly drawn in every time period, it is impossible to track individuals to model their dynamics and control for individual fixed effects.

Deaton (1985) suggests the use of pseudo-panel data constructed from these repeated cross-sectional surveys where cohorts instead of individuals are track over time. A cohort is defined as "a group with fixed membership, individuals of which can be identified as they show up in the survey", such as year of birth cohort, education or sex cohort. In this section we show how pseudo-panel data can be used to provide consistent estimates of earnings persistence while also effectively dealing with problems commonly confronted when using standard panel data such as measurement error and attrition.

3.3.3.1 Identification of Earnings Persistence in Pseudo-Panel Settings

Restating the assumed Galton-Markov data generating process in Equation (3.6) below with a decomposed error term to include the individual-specific fixed effect ($\varphi_{i,t} = \kappa_i + \mu_{i,t}$) for repeated cross-sections gives

$$Y_{i(t),t} = \alpha + \beta Y_{i(t),t-1} + \gamma X_{i(t),t} + \kappa_{i(t)} + \mu_{i(t),t} \tag{3.45}$$

Where the use of the index $i(t)$ instead of i makes explicit the fact that different individuals are sampled in each period. With repeated cross-sectional data $Y_{i(t),t-1}$ is the value of actual earnings Y at time $t-1$ for an individual i, who is only observed at time t; hence, $Y_{i(t),t-1}$ is unobserved. The solution to this is to construct an estimate for $Y_{i(t),t-1}$ using information on earnings of other individuals observed at $t-1$ instead.

Following Deaton (1985), we divide the population into C mutually exclusive cohorts ($\sum_{c=1}^{C} n_c = N$, where N is the total number of observations at time t), with fixed membership over time such as year of birth cohort. Taking cohort averages of Equation (3.45) over the n_c individuals observed in cohort c at time t yields Equation (3.46) below:

$$\bar{Y}_{c(t),t} = \alpha + \beta \bar{Y}_{c(t),t-1} + \gamma \bar{X}_{c(t),t} + \kappa_c + \bar{\mu}_{c(t),t} \tag{3.46}$$

Where $\bar{Y}_{c(t),t} = (1/n_c)\sum_{i=1}^{n_c} Y_{i,t}$ is the average of earnings of all individuals in cohort c at time t. The other sample means are defined similarly while $\bar{\kappa}_c = (1/n_c)\sum_{i=1}^{n_c} \kappa_c = \kappa_c$ is the cohort fixed effect. The resulting data set is now a pseudo-panel with repeated observations for C cohorts over T time periods. As before, $\bar{Y}_{c(t),t-1}$, the sample average of earnings in period $t-1$ of individuals in cohort c observed at time t, is not observable as individuals are only observed once in repeated cross-sectional data. To solve this problem, we replace $\bar{Y}_{c(t),t-1}$ with $\bar{Y}_{c(t-1),t-1}$, the sample mean of earnings of different individuals who are in the same cohort c but were observed at time $t-1$. The resulting regression equation is

$$\bar{Y}_{c(t),t} = \alpha + \beta\bar{Y}_{c(t-1),t-1} + \gamma\bar{X}_{c(t),t} + \kappa_c + \bar{\mu}_{c(t),t} + \pi_{c(t),t} \tag{3.47}$$

$$for\, c = 1, .., C\,;\;\; t = 2, ..., T$$

Where $\pi_{c(t),t} = \beta(\bar{Y}_{c(t),t-1} - \bar{Y}_{c(t-1),t-1})$ is the prediction error. The parameter β cannot be consistently estimated by an OLS regression of Equation (3.47) as the regression composite error term will be correlated with the lagged earnings variable $Y_{c(t-1),t-1}$ in finite samples. This happens as a result of measurement error from observing different individuals in time t to those observed at time $t-1$ (McKenzie, 2004). In addition, there is a correlation between the unobserved cohort fixed effect and the explanatory variable that would cause a bias in the estimate of earnings persistence.

McKenzie (2004) shows that, as the number of individual in each cohort becomes large ($n_c \to \infty$), prediction error would diminish as both $\bar{Y}_{c(t),t-1}$ and $\bar{Y}_{c(t-1),t-1}$ get closer to the population mean for cohort c at time $t-1$. Hence, $\pi_{c(t),t}$ converges to zero and we can ignore this term. However, this condition rests upon one very important implicit assumption that individuals sampled at different point in time are from populations with the same mean properties. If there are significant deaths and inward and outward migration that drastically change the composition of the population in the defined cohort from one period to another in a non-random manner, then this condition will not be satisfied. Therefore, in this chapter we assume that there is no inward and outward migration that significantly changes our sample population from period to period.

As the measure of the cohort used in this chapter is the year-of-birth cohort, we only need that there is no significant migration in and out of Thailand for each year-of-birth cohort to keep the sample representative of the population. Rural-urban migration will not cause a

bias in the result as it does not alter the underlying population tracked, which in this case, are individuals who were born in the same year. In fact, migration within the country to seek better living standards is one of the things we want to measure in this study. We will limit our focus to cohorts that are not too old to prevent the problem of deaths significantly altering the sample.

In addition, the close population assumption above allows us to control for the cohort fixed effect. The unobserved cohort fixed effect κ_c may actually not be "fixed", having originated from the individual fixed effect of different individuals in different time periods. However, if the cohort population is close and large enough as $n_c \to \infty$, individual errors will cancel each other out and variation over time can be ignored (Moffitt, 1993). Therefore, by including cohort dummies as regressors as in Equation (3.48) below, we can also capture the cohort fixed effect and solve the endogeneity problem.

$$\bar{Y}_{c(t),t} = \alpha + \beta \bar{Y}_{c(t-1),t-1} + \gamma \bar{X}_{c(t),t} + \lambda Z_{c(t)} + \bar{\mu}_{c(t),t} \tag{3.48}$$

Where $Z_{c(t)}$ is the cohort dummies variable and we set $\pi_{c(t),t} = 0$ as in our assumption above.

Alternatively, it is also possible to consider the pseudo-panel approach as a type of IV regression that fixes the endogeneity problem of having unobserved $Y_{i(t),t-1}$ in the data generating process in Equation (3.45). Verbeek and Vella (2005) show that OLS regression on Equation (3.47) is identical to using a set of individuals' cohort dummies as the time-invariant instruments ($Z_{i(t)}$) for both the unobserved lagged dependent variable $Y_{i(t),t-1}$ and the exogenous variable $X_{i(t),t}$ in Equation (3.45) (they termed this the "augmented IV estimator"). The following conditions must be satisfied.

1. Instrument Relevance and Rank Condition

   $Z_{i(t)}$ must exhibit sufficient correlation with the unobserved lagged dependent variable $Y_{i(t),t-1}$ and the exogenous variable $X_{i(t),t}$. The correlations must exhibit time variation and not be collinear.

2. Instrument Exogeneity

   $Z_{i(t)}$ should not be correlated with the error term $\mu_{i(t),t}$.[9] This rules out the possibility

[9] We are only concerned with $\mu_{i,t}$ as we assumed that the fixed effect could be taken care of by the within estimator.

that there is a time-varying unobservable cohort effect.

The usual downward bias problem with estimating dynamic panel data models (Nickell, 1981) does not apply here because the error term $\bar{\mu}_{c(t),t}$, which is the within cohort average of individual error terms, is asymptotically zero as $n_c \to \infty$ provided that $\mu_{i,t}$ contains no time-varying cohort-level component as assumed above (Verbeek and Vella, 2005). If these assumptions are all met, the OLS regression of the pseudo-panel in Equation (3.48), which is equivalent to the IV scheme, will give a consistent estimate of earnings persistence β.

3.3.3.2 Dealing with Measurement Error, Attrition and Representativeness

The last subsection dealt with the issue of identification of earnings persistence using pseudo-panel data rather than standard panel data. In this subsection, we show that even when earnings variables are measured with errors, pseudo-panel regression still provides a consistent estimate of earning persistence.

The pseudo-panel Equation (3.48) is rewritten as

$$\bar{Y}^*_{c(t),t} = \alpha + \beta \bar{Y}^*_{c(t-1),t-1} + \gamma \bar{X}_{c(t),t} + \lambda Z_{c(t)} + \bar{\mu}_{c(t),t}$$

Where $Y^*_{i,t}$ denotes the true earnings without measurement error. With classic measurement error of the form $Y_{i,t} = Y^*_{i,t} + \nu_{i,t}$, we can rewrite the above pseudo-panel equation as

$$\bar{Y}_{c(t),t} = \alpha + \beta \bar{Y}_{c(t-1),t-1} + \gamma \bar{X}_{c(t),t} + \lambda Z_{c(t)} + \bar{\mu}_{c(t),t} + \bar{\nu}_{c(t),t} - \beta \bar{\nu}_{c(t-1),t-1} \tag{3.49}$$

Where $\bar{\nu}_{c(t),t}$ is the mean measurement error in earnings at time t for individuals in cohort c. Following Antman and McKenzie (2007b), our identifying assumption is that the law of large numbers applies within a cohort, so that as the number of individuals within a cohort becomes large, $n_c \to \infty$, we have

$$\frac{1}{n_c} \sum_{i=1}^{n_c} \nu_{i,t} \to 0 \tag{3.50}$$

This implies that both cohort averages of measurement error $\bar{\nu}_{c(t),t}$ and $\bar{\nu}_{c(t-1),t-1}$ converge to zero as $n_c \to \infty$ and we can ignore the measurement error terms as these measure-

ment errors average out in the process of constructing the pseudo-panel. Assumption (3.50) is quite weak and covers a large number of cases which violate the assumption of classical measurement error. In addition, dynamics in individual measurement errors, where measurement errors on earnings of one individual are likely to be correlated over time, will not be a problem in this setting as we observe different individuals in each time period. As long as we have a sufficiently large number of observations in each cohort (this is commonly $n_c \geq 100$ in the literature), the OLS estimate of the pseudo-panel equation will give a consistent estimate of the Galton-Markov earnings persistence parameter β.

However, in the case where the assumption (3.50) does not hold, such as when the same measurement errors are common to every individual within the cohort, so that averaging does not make the errors disappear, then the above analysis will not give a consistent estimate of β. Nevertheless, this is very unlikely to happen.

In addition to dealing with measurement error, pseudo-panel data does not suffer from the issue of non-random attrition unlike that faced by some long-running standard panel data where individuals systematically left the survey, leaving a biased sample. This is not a problem in a pseudo-panel setting as new representative and random samples are drawn in each period.

The pseudo-panel data used in this chapter is also constructed from random and nationally representative samples from the LFS. Therefore, the earnings persistence estimate obtained from this pseudo-panel regression will be more externally valid than the estimate from Townsend Thai data. The only major drawback of the pseudo-panel data is that intra-cohort individual dynamics cannot be investigated.

The pseudo-panel regression discussed above extended straightforwardly to the quantile regression framework. Equation (3.51) shows the conditional quantile pseudo-panel specification.

$$Q_{Y_{i,t}}(\tau|\bar{Y}_{c(t-1),t-1}) = \beta_0(\tau) + \beta_1(\tau)\bar{Y}_{c(t-1),t-1} + \beta_2(\tau)\bar{X}_{c(t),t} \tag{3.51}$$

Pseudo-panel data will be used to estimate the unconditional and conditional Galton-Markov persistence parameter for both the OLS and quantile regressions. We will compare the results of these specifications with those of the pooled OLS, quantile regression, IV regression without the fixed effect and Arellano-Bond regression obtained using standard

panel data in Section 3.5. It is important to note that as the earnings and income persistence estimates $\hat{\beta}$ are obtained using generated regressors ($\bar{Y}_{c(t),t}$ and $\bar{Y}_{c(t-1),t-1}$), the standard error of $\hat{\beta}$ will be incorrect. This will be corrected for by bootstrapping for the standard error in all pseudo-panel regression.

3.3.4 Permanent Earnings Model

In the last three subsections, we have been focussing on the identification and estimation of the persistence parameter β of the Statistical Model of Earnings Dynamics, specifically the Galton-Markov model. While this simple statistical model has provided a powerful tool for the quantification of the overall degree of persistence, the model has been based on a very restrictive set of assumptions that is only representative of a very specific process of earnings dynamics. To obtain a more complete understanding of the dynamic process of earnings in Thailand, it is necessary to move beyond the purely statistical model to a more structural model. In this respect, the permanent earnings model that is based on the permanent earnings hypothesis provides a natural point of departure.

3.3.4.1 The Canonical Variance Components Model of Earnings Dynamics

In the setting of what is often referred to as the canonical variance components model of earnings dynamics, earnings are decomposed in to permanent and transitory components in a similar spirit to that of Friedman's (1957) work on consumption.

The earnings or the logarithm of earnings of individual i at time period t, denoted by $Y_{i,t}$, are assumed to be determined by two components: a permanent component, δ_i, that is assumed to be constant over time, and a transitory component, $v_{i,t}$.

$$Y_{i,t} = \delta_i + v_{i,t} \tag{3.52}$$

The permanent component reflects the personal qualities of the individual that in this model are fixed over time such as an educational degree. The transitory component captures the chance factors influencing earnings that cause random deviation from the permanent component in a particular period. The expected values of both earnings components are assumed to be zero and there is no correlation between the two. These assumptions are stated formally as

$$\delta_i \sim (0, \sigma_\delta^2), \quad v_{i,t} \overset{iid}{\sim} (0, \sigma_v^2), \quad E(\delta_i v_{i,t}) = 0 \tag{3.53}$$

In this simple permanent earnings model, the degree of persistence depends on the proportion of variation in earnings ($Y_{i,t}$) accounted for by the transitory component. Taking the correlation coefficient between earnings in period t and period $t+1$, denoted by $r(t, t+1)$, as a measure of persistence, and using the first and second moment assumptions in Equation (3.53), the correlation coefficient can be calculated as

$$r(t, t+1) = \frac{Cov(Y_{i,t} Y_{i,t+1})}{\sigma_y(t) \cdot \sigma_y(t+1)} = \frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_v^2} \tag{3.54}$$

This suggests that the correlation of earnings is equal to the proportion of total variance accounted for by the permanent component. The lower the proportion of the transitory variance component in earnings variation, the higher the degree of persistence.

However, this model consists of a few shortcomings as a result of the restrictive assumptions. The iid assumption on the transitory components precludes serial correlation in the chance factors, and transitory fluctuation such as labour market shocks will completely disappear after each period. As a result, there is no specific relation between earnings in one period and those in the next and, in this sense, there are no real dynamics to the model. Another unrealistic implication is the fixed nature of the permanent component. This simple model does not allow for individual development over the life cycle in accordance with the human capital story.

In the subsequent subsections we deal with these weaknesses and extend the model accordingly.

3.3.4.2 Serial Correlation of the Transitory Component

As discussed above, the iid assumption on the transitory component where transitory shocks to earnings last no more than one period has led to the lack of real dynamics in the model. A common practice in the literature is to introduce serial correlation in the form of a low order AR or ARMA process.

In a pioneering paper by Lillard and Willis (1978), the authors extend the above simple canonical model to allow for an AR(1) structure in the transitory component. The model is

given by

$$Y_{i,t} = X_{i,t}\beta + \mu_{i,t} \tag{3.55}$$

$$\mu_{i,t} = \delta_i + v_{i,t} \tag{3.56}$$

$$v_{i,t} = \gamma v_{i,t-1} + \eta_{i,t} \tag{3.57}$$

Where $X_{i,t}$ are some conditioning variables and $\mu_{i,t}$ is the error term, while δ_i and $v_{i,t}$ are the permanent and transitory components as before. The transitory component $v_{i,t}$ are now assumed to follow an AR(1) process in Equation (3.57), where γ is the AR(1) coefficient that measures the degree of persistence in the transitory element and is closely related to the β coefficient in the Galton-Markov model studied earlier.[10]

The assumptions of the model are:

- δ_i is a random individual component ($\sim (0, \sigma_d^2)$); σ_d^2 is assumed to be homogeneous throughout the population
- $\eta_{i,t}$ is a purely random component ($\overset{iid}{\sim} (0, \sigma_\eta^2)$)
- δ_i and $\eta_{i,t}$ are assumed to be independent of each other and of $X_{i,t}$

Assume that the earnings at the beginning of the work history are given by an error of the form

$$v_{i,0} = \frac{\eta_{i,0}}{\sqrt{1-\gamma^2}} \tag{3.58}$$

and $\eta_{i,t}$ thereafter ($t > 1$). Therefore, $v_{i,t} \sim (0, \sigma_v^2)$, where $\sigma_v^2 = \frac{\sigma_\eta^2}{1-\gamma^2}$.

Residual covariance structure is of the form

$$E[\mu_{i,t}\mu_{j,\tau}] = \begin{cases} \sigma_\mu^2 = \sigma_\delta^2 + \sigma_v^2 & ;\ i = j,\ t = \tau \\ \sigma_d^2 + \gamma^s \sigma_v^2 & ;\ i = j,\ |t - \tau| = s > 0 \\ 0 & ;\ i \neq j \end{cases} \tag{3.59}$$

[10] See Atkinson et al. (1992) for the discussion and derivation.

A potential deficiency with this model concerns the assumption made about the initial condition in Equation (3.58) above. MaCurdy (1982) suggested that the relatively short time dimensions of the data used in this type of study necessitate the inclusion of an initial condition term, $v_{i0} \overset{iid}{\sim} (0, \sigma^2_{v0})$, which captures the accumulation of the stochastic process prior to the first year of the panel.

In this case the covariances are

$$E[\mu_{i,t}\mu_{j,\tau}] = \sigma^2_d + \gamma^{t+\tau}\sigma^2_{v0} + [\frac{\gamma^{t-\tau}(1-\gamma^{2\tau})}{(1-\gamma^2)}]\sigma^2_\eta \quad for\, t \geq \tau \tag{3.60}$$

This can be rewritten as

$$\begin{aligned} E[\mu_{i,t}\mu_{j,\tau}] &= \sigma^2_d + \gamma^{t+\tau}\sigma^2_{v0} + \gamma^{t-\tau}(1-\gamma^{2\tau})\sigma^2_{v*} \\ &= \sigma^2_d + \gamma^{t-\tau}\sigma^2_{v*} + \gamma^{t+\tau}(\sigma^2_{v0} - \sigma^2_{v*}) \qquad for\, t \geq \tau \end{aligned} \tag{3.61}$$

Where $\sigma^2_{v*} = \sigma^2_\eta/(1-\gamma^2) = lim_{t\to\infty} var(v_t)$, the long-run limit variance of v. This moment specification collapses back to that of the earlier specification without an initial condition if $\sigma^2_{v0} = \sigma^2_{v*}$ as we have assumed in (3.58) above.

The model can also be extended to allow the iid element of the transitory component $\eta_{i,t}$ to display some degree of autocorrelation. In this case, we will model $\eta_{i,t}$ to follow an MA(1) process. Equation (3.57) can be rewritten as

$$v_{i,t} = \gamma v_{i,t-1} + \eta_{i,t} + \theta\eta_{i,t-1} \tag{3.62}$$

Where $\eta_{i,t} \overset{iid}{\sim} (0, \sigma^2_\eta)$ as before. The transitory component of the earnings equation now follows an ARMA(1,1) process. If an infinite history of random shocks is assumed the covariances are given by

$$E[\mu_{i,t}\mu_{j,\tau}] = \sigma^2_d + \gamma^{t-\tau}\sigma^2_v + \begin{cases} \gamma^{t-\tau-1}\theta\sigma^2_\eta & if\, t > \tau \\ 0 & if\, t = \tau \end{cases} \tag{3.63}$$

Where $\sigma^2_v = \sigma^2_\eta(1+\theta^2+2\gamma\theta)/(1-\gamma^2)$.

In this chapter, we will estimate Lillard and Willis's AR(1) model, the AR(1) model with the initial variance, and the ARMA(1,1) model and provide comparison of their results.

3.3.4.3 Heterogeneous Growth of the Permanent Component

As discussed in Subsection 3.3.4.1, the assumption that the permanent component of earnings remain fixed over the life cycle can be too restrictive and does not allow for the individual development consistent with the human capital accumulation literature. It is more realistic to allow permanent component to evolve with age. A number of studies have modelled this evolution with either heterogeneous growth (e.g. Lillard and Weiss, 1979; Guvenen, 2009) or a random walk process (Kalwij and Alessie, 2007) or both (Baker and Solon, 2003).

In this chapter we follow in the approach of the literature that modelled the permanent component according to the heterogeneous growth process. This process involves supplementing the permanent earnings component δ_i with an additional parameter that accounts for earnings growth over the life cycle, an individual-specific growth rate. The structure of the heterogeneous growth model is of the form

$$Y^P_{i,a,t} = \delta_i + \lambda_i a_{i,t} \tag{3.64}$$

Where the P-superscript denotes the permanent component of the earnings equation, λ_i is the individual-specific growth rate and $a_{i,t}$ represents age where $a_{i,t} = 0, ...A$.[11] The heterogeneous slope has also been used with respect to years of labour market experience. In this chapter we combine the above heterogeneous growth model of the permanent component with the AR(1) model of the transitory component to give the following model:

$$Y_{i,t} = X_{i,t}\beta + \mu_{i,t}$$

$$\mu_{i,t} = \delta_i + \lambda_i a_{i,t} + v_{i,t}$$

$$v_{i,t} = \gamma v_{i,t-1} + \eta_{i,t}$$

Assumptions on the first and second moments of the individual-specific growth coefficient are given by

[11] Here, age is measured as the deviation from the minimum observed age within the sample for each cohort. The minimum age deviation is therefore equal to 0, while the maximum deviation is T-1, where T is the number of total time periods in the sample. More information is given in the next subsection on how to estimate the heterogeneous growth model.

$$(\delta_i, \lambda_i) \sim [(0,0);\, (\sigma^2_\delta, \sigma_{\delta\lambda}, \sigma^2_\lambda)] \tag{3.65}$$

Here there is a potential correlation between the random growth parameter and the fixed individual-specific parameter. Assumptions on the remaining terms $v_{i,t}$ and $\eta_{i,t}$ are as before. The covariance structure is of the form

$$E[\mu_{i,t}\mu_{j,t+k}] = \begin{cases} \sigma^2_\mu = \sigma^2_\delta + 2\sigma_{\delta\lambda}a_{i,t} + \sigma^2_\lambda a^2_{i,t} + \sigma^2_v & ;\, i = j,\, k = 0 \\ \sigma^2_d + \sigma_{\delta\lambda}(a_{i,t} + a_{i,t+k}) + \sigma^2_\lambda a_{i,t}a_{i,t+k} + \gamma^k\sigma^2_v & ;\, i = j,\, k > 0 \\ 0 & ;\, i \neq j \end{cases} \tag{3.66}$$

In this model σ^2_d represents the time-invariant heterogeneity such as a difference in earnings that results from a difference in the education degrees obtained and remains constant over time. Alternatively, σ^2_λ represents the time-varying heterogeneity that evolves over the life cycle such as that due to the effect of on-the-job training and skill acquisition. Both σ^2_d and σ^2_λ are variances and must take positive values.

The important term in this heterogeneous growth model is the covariance between time-invariant and time-varying individual-specific effect term, $\sigma_{\delta\lambda}$, which could take either a positive or a negative value. A positive value of $\sigma_{\delta\lambda}$ would indicate that there is a positive correlation between δ_i and λ_i, and that these two components reinforce each other. This is a situation where people who possess a high value of a time-invariant individual-specific effect are also those who develop the most during the life cycle, hence possessing a high value of a time-varying individual-specific effect. In such a case, permanent earnings dispersion will increase over time and initially advantaged individuals will move even further away from the rest over the life cycle. A negative value of $\sigma_{d\lambda}$ on the other hand provides an interpretation that is the complete opposite. In this case, an individual with a lower initial endowment (lower degree of education) will experience a faster growth in earnings over the life cycle than the initially better off individual, resulting in a catching up effect. There are greater dynamics within the earnings distribution in this latter case.

3.3.4.4 Estimation of the Covariance Structure

Minimum Distance Estimation (MDE) is employed as an application of GMM to estimate the parameters of the covariance structure of the permanent earnings model specified above. MDE is used to fit the parameters of the theoretical models to the data and allow estimation to be conducted without imposing assumptions regarding the earnings distribution.

The empirical covariance matrix of a balanced panel of observations is calculated as $C = \frac{1}{N}\sum_1^N y_i y_i'$, where y_i is a vector of the residuals of the first stage OLS regression in Equation (3.55) for an individual i of length T.

$$y_i = \begin{bmatrix} y_{i,1} \\ y_{i,2} \\ \vdots \\ y_{i,T} \end{bmatrix}_{Tx1}$$

In the first stage we will regress log individual/cohort earnings ($Y_{i,t}$) on education and the time period dummy variables to obtain the required residuals. In the second stage we use these residuals to calculate the empirical covariance matrix.

The empirical covariance matrix is given by

$$\hat{C} = \begin{bmatrix} \frac{1}{N}\sum_1^N y_{i,1}y_{i,1} & \frac{1}{N}\sum_1^N y_{i,1}y_{i,2} & \cdots & \frac{1}{N}\sum_1^N y_{i,1}y_{i,T} \\ \frac{1}{N}\sum_1^N y_{i,2}y_{i,1} & \frac{1}{N}\sum_1^N y_{i,2}y_{i,2} & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{1}{N}\sum_1^N y_{i,T}y_{i,1} & \cdots & \cdots & \frac{1}{N}\sum_1^N y_{i,T}y_{i,T} \end{bmatrix}_{TxT}$$

Then a vector m_C, which collected all $(T(T+1))/2$ distinct elements of the TxT empirical covariance matrix $\hat{C}$, can be defined as

$$m_c = Vech(\hat{C})$$

The $Vech()$ operator turns a symmetric matrix into a column matrix by stacking all columns of the lower triangular matrix on top of each other.

Let $f(\theta)$ define a non-linear function of the theoretical covariance structure, which embodies the parameter of interest θ. Using the Lilliard and Willis AR(1) model as an illustrative example gives $\theta = \left[\sigma_\delta^2,\, \sigma_v^2,\, \gamma\right]$.

$$m_c = f(\theta)$$

$$\begin{bmatrix} C_{11} = \frac{1}{N}\sum_1^N y_{i,1}y_{i,1} \\ C_{2,1} = \frac{1}{N}\sum_1^N y_{i,2}y_{i,1} \\ C_{3,1} = \frac{1}{N}\sum_1^N y_{i,3}y_{i,1} \\ \vdots \\ C_{T,1} = \frac{1}{N}\sum_1^N y_{i,T}y_{i,1} \\ C_{2,2} = \frac{1}{N}\sum_1^N y_{i,2}y_{i,2} \\ C_{3,2} = \frac{1}{N}\sum_1^N y_{i,3}y_{i,2} \\ C_{4,2} = \frac{1}{N}\sum_1^N y_{i,4}y_{i,2} \\ \vdots \\ C_{T,T} = \frac{1}{N}\sum_1^N y_{i,T}y_{i,T} \end{bmatrix}_{[(T(T+1)/2]\,x\,1} = \begin{bmatrix} \sigma_\delta^2 + \sigma_v^2 \\ \sigma_d^2 + \gamma^1\sigma_v^2 \\ \sigma_d^2 + \gamma^2\sigma_v^2 \\ \vdots \\ \sigma_d^2 + \gamma^{T-1}\sigma_v^2 \\ \sigma_\delta^2 + \sigma_v^2 \\ \sigma_d^2 + \gamma^1\sigma_v^2 \\ \sigma_d^2 + \gamma^2\sigma_v^2 \\ \vdots \\ \sigma_d^2 + \sigma_v^2 \end{bmatrix}_{[(T(T+1)/2]\,x\,1}$$

MDE obtains consistent estimates of θ by minimising the squared distance between the empirical covariance structure m_C and the theoretical covariance structure $f(\theta)$ in the form

$$\hat{\theta} = \underset{\theta}{argmin}[m_c - f(\theta)]'W[m - f(\theta)] \tag{3.67}$$

This is subject to the following restrictions:

- Stationarity: $|\gamma| < 1$
- All variance parameters σ_δ^2, σ_v^2 and σ_η^2 are positives.

The matrix W in Equation (3.67) is a positive definite weighting matrix. The Optimal Minimum Distance (OMD) estimator uses $W = V^{-1}$ as the weighting matrix, where V is a consistent estimator of the asymptotic variance-covariance matrix of m_c. The OMD estimator is asymptotically efficient; however, Monte Carlo evidence given by Altonji and Segal (1996) indicates that the estimator potentially suffers from a serious small sample bias. To avoid this problem, the authors propose the use of the identity matrix as the weighting matrix ($W = I$). The resulting estimation procedure is commonly referred to as Equally Weighted Minimum Distance (EWMD) estimation, and is therefore equivalent to a non-linear least squares procedure. In this chapter, we use the EWMD estimation to obtain the parameters of the covariance structure of earnings for all specifications.

Estimation of the Heterogeneous Growth Model

The estimation of the heterogeneous growth specification of the permanent earnings model is more complicated and involves minor differences from other specifications, which are discussed below.

In order to gain life cycle related information, it necessary to track individuals born in the same cohort over time as they age. In order to separately identify age and time effects it is necessary to observe earnings autocovariances at different phases of the life cycle in each year, which is achieved by estimating the covariance structure separately by birth cohorts (Cappellari, 2004). Ideally, it would be best to track each single year birth cohort separately over the entire range of the sample availability. However, given the small sample size of both the Townsend and the pseudo-panel data, it is necessary to construct different birth cohorts that join several consecutive birth cohorts together to help preserve cell size. Investigating each single cohort separately could result in having only a very small number of observations for the younger and the older cohorts in the sample, which prevent the empirical covariance from being estimated accurately.

The overall sample is divided into four birth cohorts spanning eleven years each for the Townsend sample and five birth cohorts spanning five years each for the LFS pseudo-panel. All birth cohorts can be observed over the full length of the relevant period from 1997–2010 (14 years) in the Townsend sample and 1985–2011 (27 years) in the LFS sample. The central age is then assigned to each birth cohort in the initial year. However, in our analysis of the age structure as noted in Subsection 3.3.4.3, age is measured as a deviation from the minimum in the sample (the central age of each cohort in the initial year of the survey) so that the age would range from 0 in the first year to 13 in the final year in the Townsend sample, and 0 to 26 in the LFS pseudo-panel. This measure of age as deviations for each cohort allows for the identification of the pure age effect excluding the cohort effect. The resulting birth cohort structures are summarised in Tables 3.2 and 3.3.

Table 3.2: Birth Cohort Structure of Townsend Sample

Cohort	Year of Birth	Central Age (1997)	Central Age (2010)
1	1930-1940	62	75
2	1941-1951	51	64
3	1952-1962	40	54
4	1963-1973	29	42

Table 3.3: Birth Cohort Structure of LFS Sample

Cohort	Year of Birth	Central Age (1985)	Central Age (2011)
1	1941-1945	42	69
2	1946-1950	37	64
3	1951-1955	32	59
4	1956-1960	27	54
5	1961-1965	22	49

The empirical covariance matrix will be calculated separately for each birth cohort to control for the cohort effect. In order to take into account the cohort heterogeneity, cohort-specific factor loading is introduce into the earnings model where the residual term of the heterogeneous growth model can be modified as

$$\mu_{i,t} = \pi_C^P(\delta_i + \lambda_i a_{i,t}) + \pi_C^T v_{i,t} \tag{3.68}$$

Where π_C^P is the birth cohort dummy for the permanent component and π_C^T for the transitory component respectively. The subscript c denotes the cohort number ranging from 1 to 4 for the Townsend panel and 1 to 5 for the LFS pseudo-panel. π_C^P and π_C^T must be set to one for the reference cohort to achieve identification. Equation 3.69 shows the complete variance decomposition of the heterogeneous growth model that also accounted for birth cohort factor loading:

$$E[\mu_{i,t}\mu_{j,t-k}] = \begin{cases} \sigma_\mu^2 = \pi_C^P[\sigma_\delta^2 + 2\sigma_{\delta\lambda}a_{i,t} + \sigma_\lambda^2 a_{i,t}^2] + \pi_C^T\sigma_v^2 & ;\ i = j,\ k = 0 \\ \pi_C^P[\sigma_d^2 + \sigma_{\delta\lambda}(a_{i,t} + a_{i,t-k}) + \sigma_\lambda^2 a_{i,t}a_{i,t-k}] + \pi_C^T[\gamma^k\sigma_v^2] & ;\ i = j,\ k > 0 \\ 0 & ;\ i \neq j \end{cases} \tag{3.69}$$

3.3.5 Summary

In this section we have discussed in detail the parameter of interest, the degree of unconditional and conditional persistence in earnings and income. We illustrated that this

persistence can be captured by β, the slope coefficient of an OLS regression of current on lagged earnings in the Galton-Markov model of earnings.

Moreover, we provide a detail discussion on Galton's fallacy of regression towards the mean, relating to the fact that the OLS estimate of β obtained does not provide any information on the process of convergence of the earnings distribution. The quantile regression method is employed to study persistence at a different part of the earnings distribution and will allow an investigation into the changing shape of the earnings distribution.

We attempt to consistently estimate the persistence parameters for the both the OLS and the quantile regression in a standard panel setting by correcting for omitted variable bias and measurement error in the earnings variable that could cause bias in the persistence estimates using various techniques. However, the estimates obtained using standard panel data are still affected by attrition bias and non-representativeness of the Townsend Thai panel sample.

Dynamic pseudo-panel methods, which successfully dealt with the problems of measurement error, attrition bias and non-representativeness, are utilised to overcome these problems. The drawback of this approach, however, is that earnings dynamics within cohorts cannot be studied. In addition, we employ the maximum entropy method for both the standard panel and the pseudo-panel, which overcome the perfect multicollinearity problem and allow for estimation of earnings persistence while controlling for APC effects all at the same time, reducing the extent of omitted variable bias caused by these variables in the Galton-Markov setting.

Lastly, we move beyond the statistical model of earnings dynamics, the Galton-Markov model, towards models that introduce more economic structure in the analysis of the earnings process in order to gain a more complete understanding of the nature of the dynamics of living standards in Thailand. This involves the decomposition of earnings into the permanent and transitory element to determine their relative importance in the earnings process.

The next section provides discussions and descriptive statistics of the data used for income and earnings dynamics analysis in this chapter.

3.4 Data and Descriptive Statistics

In this section, we present descriptions of the two data sets used for income and earnings dynamics analysis in this chapter, the LFS of the Kingdom of Thailand and the Townsend Thai panel data.

3.4.1 Labour Force Survey of the Kingdom of Thailand

The LFS is a national cross-sectional survey, collected by the National Statistical Office of Thailand (NSO) since 1963. From 1984–1997, the NSO carried out three rounds of the survey each year. The first round (January to March) coincided with the non-agricultural season. The second round (April to June) corresponded to the period when the Thai academic year ended and students who had finished school or university joined the labour force. The third round (July to September) matched the agricultural season. The fourth round (October to December) has been additionally conducted since 1998.

The samples for each LFS are drawn from different households throughout the country, using clustered random sampling stratified by geographic regions and provinces. Each quarterly survey contains data for 70,000 – 200,000 people representing around 0.1 – 0.5 percent of the total Thai population. Survey questions include detailed information on employment and unemployment – for instance, labour force status, occupation, industry, hours of work, and earnings – along with characteristics of the labour force and economically inactive persons – such as age, gender, relationship to the household head, marital status, and educational attainment.

3.4.1.1 Data Construction: Pseudo-Panel

This chapter uses the LFS data from 1985 to 2011 to construct pseudo-panel data. Thai agricultural workers tend to migrate to work in the cities during the non-agricultural season and return home during the rainy season in the third quarter to resume farming (Warunsiri and McNown, 2010). Therefore, only the data collected in the third quarter of each survey year are used in the analysis to control for the effect of seasonal labour movement.

The construction of a pseudo-panel, where cohorts are followed through time instead of individual households, started with the determination of household birth cohorts. Household birth cohort information is obtained through the identification of the survey year that the

observation belongs to and the age of the household head. The differences between the two give the household birth cohort. For every survey year, variables of interest such as income and education are averaged across individuals in the same birth cohort to form cohort averages. At the end of the process, we formed a pseudo-panel, which has cohort-year as a unit of observation.

Cohorts are defined for birth years from 1941 to 1965 using the surveys from 1985 through to 2011. This restricts the age of households to 20–70 years in order to focus on working age population and also allow for the investigation of income persistence in retirement years. In total, the data set pooled data from 25 single year-of-birth cohorts over a period of 27 years to form 675 cohort-year observations. The total number of individual household heads used in the construction of the pseudo-panel is 332,875. The cell size of each unit of observation in the pseudo-panel mostly exceeds 100 observations per cell (in 665 out of 675 cells), with only 4 cells having a cell size below 50. (The smallest is 35 observations per cell.)

It is important to note that all individual-level characteristic variables used to form cohort averages, such as education, age and sector of employment, were obtained from individual characteristics of the household head. This is a conventional practice in the construction of a pseudo-panel and is preferred to taking the average of each individual characteristic within the household to arrive at household average age for instance. On the other hand, household-level variables such as household size and the number of children were obtained from aggregation at the household level.

3.4.1.2 Key Variables Construction

The key variables used in the analysis of dynamics of living standards are income and earnings. Earnings are wages received by the household from participating in full-time formal sector employment while income incorporates additional earnings from self-employment activities such as business and farm incomes. The LFS collected data on individual earnings throughout the period from 1985 to 2011. However, the survey stopped collecting data on business and farm incomes from 2001 onwards. As a result, there are earnings data for the whole 27-year period but there are income data for only 16 periods (1985 to 2000).

Another key point that is related to the problem above is the representativeness of the labour force in the survey. Once business and farm incomes are no longer collected,

households whose major incomes are from self-employment are excluded from the pseudo-panel construction in the later years as their incomes are not observed. This leaves only those who are full-time employees in the formal sector in the sample.

The construction of income and earning variables is similar. All earnings and income are converted to monthly earnings and income. Monthly earnings and income are used instead of hourly earnings and income because we are interested in how household living standards change from period to period within the income distribution. These movements in living standards must also include the labour supply effect, in that if individuals work harder they will have more income and earnings, which help increase their overall living standards. In this chapter, we want to incorporate the labour supply effect in the dynamic analysis as supposed to controlling for it, hence the use of monthly instead of hourly income and earnings.

Earnings can be obtained directly from the survey question for waged employees. To construct the income variables, data on net profit from businesses and farms are used in addition to earnings in order to obtain the incomes for self-employed workers. In cases where more than one household member work in the same household farm or business, the profits are then divided equally among them first. All earnings and incomes are adjusted for inflation using the Thai CPI. Individual earnings and incomes are then aggregated to the household level before being converted to real equivalised earnings and income using the Townsend Indian equivalence scale mentioned earlier in Subsection 3.2.1. These household equivalised earnings and income are later weighted[12] and averaged across birth cohort to form cohort-year observations of earnings and incomes.

Figure 3.2 shows the histogram of log real equivalised earnings and income used to construct the cohort data. Both incomes and earnings have a distribution that is approximately log-normal as one would expect.

Figures 3.3 and 3.4 plot weighted average real monthly equivalised earnings and incomes with real GDP per capita. All three series were normalised to one in the year 1985. These plots allow for the verification of earnings and incomes computed from the LFS data. We would expect earnings and incomes growth to track GDP per capita growth closely if both variables were computed correctly. While the levels of neither accumulated earnings nor incomes growth perfectly match those of real GDP per capita, the two variables show similar

[12]There is a survey weight variable in the LFS data set to correct for the random stratification sampling nature of the survey that caused some households to be over- or under-represented. This survey weight provides a correction for the sample size to match the population level.

trend with real GDP per capita accumulated growth across time. Computed earnings and incomes also picked up the effect of the Asian financial crisis as is evident in the dip observed in both curves in 1997. As average incomes are only observed up until 2000, their movement after the crisis is not clear. Average earnings show a sign of growth in a similar fashion to the growth in real GDP after the crisis but the rate of accumulated growth in earnings is slower than that of real GDP per capita. A reason for this may lie in the fact that the sample for earnings from the year 2000 onwards only consists of individuals in waged employment, whose wages are unlikely to be frequently adjusted to be in line with the overall growth of the economy which are mostly driven by export and tourism growth during this period.

Figure 3.2: Histogram of Log Equivalised Earnings and Income

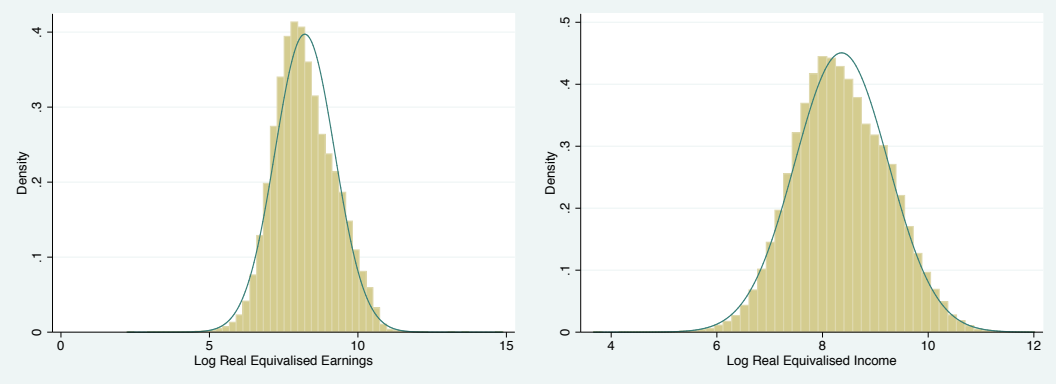


Figure 3.3: Real Equivalised Earnings with Real GDP

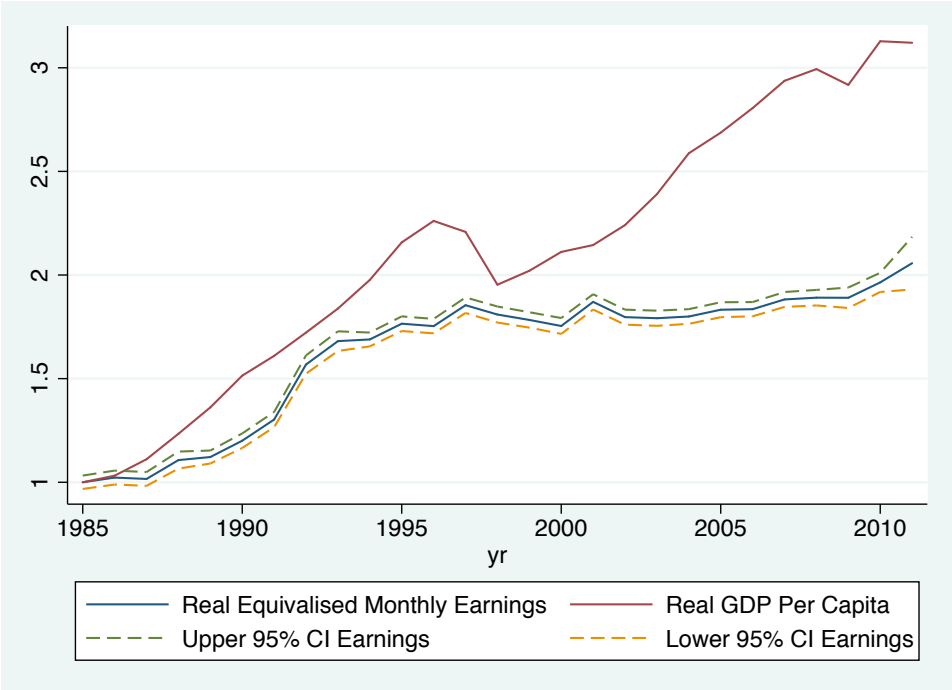
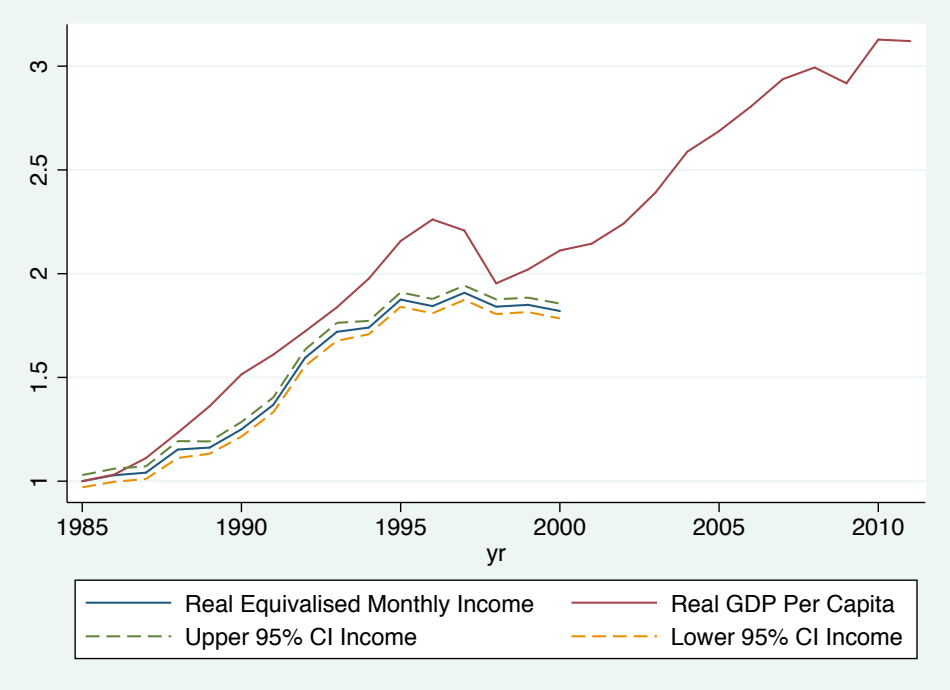


Figure 3.4: Real Equivalised Incomes with Real GDP



3.4.1.3 Descriptive Statistics

Table 3.4 shows the basic summary statistics of the weighted pseudo-panel data. The average age of the LFS pseudo-panel sample is 45 years and is a direct result of the construction of a pseudo-panel, which restricted the age to between 20 and 70 years old. Average monthly real equivalised earnings and income are similar and are in the range of 4,000 Bahts per month. Average years of schooling are around 6–7 years, which is above the primary education level. Consistent with the discussion above about the sample of formally employed workers possibly being over-represented as a result of a change in the survey question, on average around 63% of the sample works in the service sector, which commands more wages than the agricultural and manufacturing sectors, and around 67% live in an urban area.

Comparing the results of Table 3.4 and Table 3.5 in the next subsection, which present similar statistics for the Townsend Thai panel data, illustrates the fundamental differences between the two samples. The LFS sample has higher incomes, higher educational attainment and a higher proportion of people working in the more formal service sector. It represents a wealthier sample of waged employees as opposed to the self-employed farmers that formed the majority of the Townsend Thai panel data sample.

Figure 3.5 plots the quantile of earnings in the LFS data from 1985 to 2011. These plots

provide the first indication that the growth in earnings has not been achieved uniformly over time. Earnings growth at the top end of the distribution has increased more rapidly than those at the bottom end. This lends support to the use of quantile regression to investigate beyond the dynamics of the mean earnings.

Figure 3.6, shows the APC decomposition of monthly real equivalised earnings obtained using the maximum entropy principle (Browning et al., 2012). As discussed earlier, age, period and cohort are three important and distinct concepts that affect earnings and income. Age effects capture the evolution of income and earnings over an individual's life-cycle. Time periods effects capture the widespread effects such as macroeconomic shocks that individuals are exposed to at different point in time. Finally, cohort effects capture the impact that different histories and institutions may have on individual lifetime earnings and income. This maximum entropy method, which was discussed in detail in Subsection (3.3.2.3), allows APC effects to be simultaneously determined overcoming the perfect multicollinearity problem.

The top left panel is the cohort earnings age profile. It does not follow the usual hump shape where individual earnings gradually rise as one gets older and peak in the middle before declining in retirement. The two unusual things about the age profile in Thailand is the relatively constant age effect throughout the entire range except for a slight drop in the late thirties to the mid-forties and the sudden drop in earnings once households reached retirement age at 60. An explanation for the first observation may be that households in their late thirties are at the family formation stage; female workers within the household may have to exit the labour force to take care of children, which lowers the household real equivalised earnings at this stage. These drops in households' earnings prevent the observation of the hump shape that would have been expected if the family formation effect could be separated out. An explanation for the second observation may be found with the problem discussed earlier. As the LFS stops collecting income data on individuals who are self-employed, the earnings data consist entirely of those who are waged employees. As a result, we see a sharp decline in formal labour force participation and earnings once individuals have to retire at the retirement age of 60. A slight increase in real equivalised income at around 70 years old may be a result of inaccurate estimation as the cell sizes of the cohort data get smaller around the extremes. An alternative explanation may be that children in Thailand tend to move back in with their parents to take care of them in old age, especially after retirement.

This may serve to increase the household real equivalised earnings.

The top right panel is the period or the year effect, which picks up the macroeconomic change that households faced. It is no surprise therefore that the shape of the curve resembles the real GDP plotted earlier. The curve also picks up the effect of the Asian financial crisis that lowered earnings in 1997. Lastly, the bottom left panel plotted the cohort effect, or the effect of being born in a certain period. We can see from the curve that the younger cohort whose members were born in the later half of the 1950s were better off than those who were born slightly earlier in the early 1950s. This is because the 1955–60 cohorts of households were entering the labour market when the Thai economy was beginning to achieve spectacular growth in the mid-1980s. Cohorts whose members were born slightly earlier would have been entering the labour market during the mid-1970s, when the Thai economy suffered as a result of unstable military rules and political turmoil. Cohorts whose members were born in the early and mid-1960s, however, faced a different fortune as they would have been entering the labour market during the height of the financial crisis in the late 1990s. These APC effects provide lots of information on the evolution of household earnings and they will certainly affect earnings persistence over time. Later in the result section we will attempt to estimate the level of earnings persistence within the Thai economy while partialling out all of the APC effects discussed here.

Table 3.4: Summary Statistics of Thai Labour Force Survey

	count	mean	sd
Real Equivalised Earnings	675	4157.961	959.4573
Real Equivalised Income	400	4512.832	1137.175
Age	675	45	10.62233
Household size	675	4.588906	.5299335
Years of Education	675	6.563222	.9658072
Proportion in Agrictural Sector	675	.1968176	.1781953
Proportion in Manufacturing Sector	675	.176373	.0916989
Proportion in Service Sector	675	.6268094	.1444354
Proportion in Urban Area	675	.666037	.1199338
Proportion in Rural Area	675	.333963	.1199338

Figure 3.5: Quantile Plot: LFS

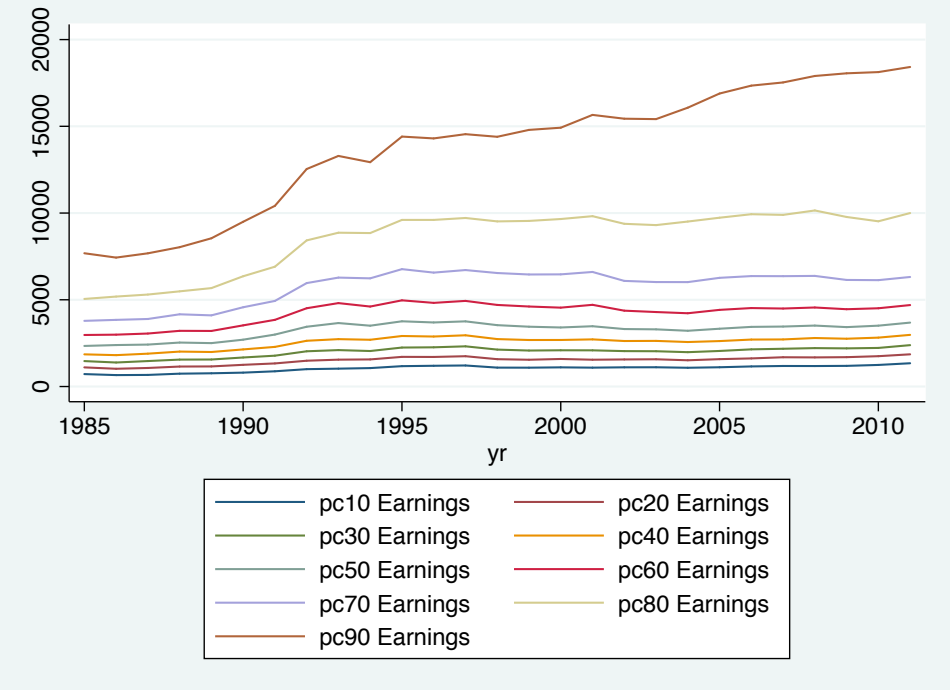
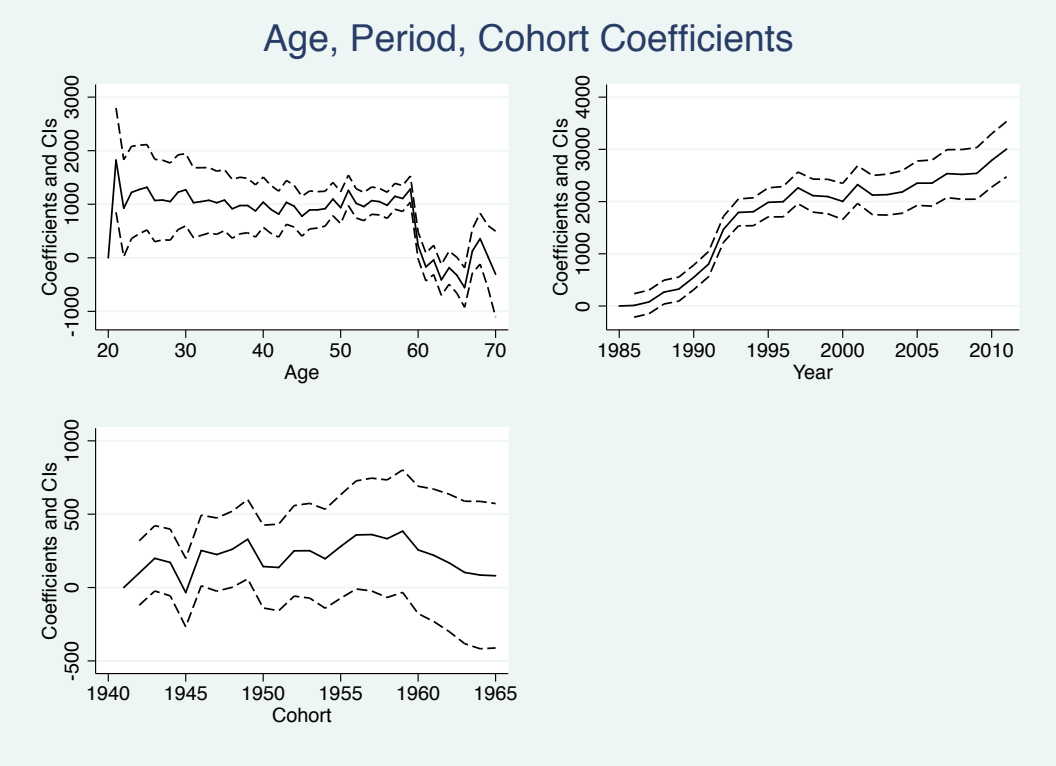


Figure 3.6: Age Period Cohort Decomposition for Thai LFS



3.4.2 Townsend Thai Data

The Townsend Thai household annual survey is a longitudinal survey collected as part of the Townsend Thai Project. The Townsend Thai Project began in 1997 with a cross-sectional survey. It was fielded in May, prior to the financial crisis, which began with the devaluation of the Thai Baht in July 1997.

Two separate regions were deliberately chosen, the more highly developed Central region located near Bangkok and the poorer semi-arid Northeast. Four provinces (changwats) were chosen in total, two in the Central region – Lop Buri and Chachoengsao – and two in the Northeast – Sisaket and Buriram. Within each of the four provinces, 12 sub-counties (tambons) were selected at random using stratification. Within each sub-county, four villages were selected at random for inclusion in the sample. Thus there are in total l92 villages in the 48 sub-counties of the four changwats. In May 1998 one-third of the original sample of villages was resurveyed. In each of the four survey provinces, four sub-counties were selected at random from the 12 original sub-counties in the cross-sectional 1997 survey. In 2003 two more provinces, Satun and Yala, were added to the survey, and another two provinces, Phrae and Petchabun, were added shortly after in 2004. The project is still on-going and at present the latest round of the annual household resurvey data released is the 2010 round.

3.4.2.1 Data Construction

This chapter uses the Townsend annual household panel data from 1997 to 2010 in the analysis of dynamics of living standards. The Townsend Thai data consists of 73,833 individuals and 3,584 households in total. However, many of these households were only observed once in the 1997 cross-sectional survey, and many more were added as a result of the expansion of the Townsend Thai Project to include four more provinces and also additional urban surveys in later years.

In this analysis, we will only focus on households that were observed for the whole 14-year period between 1997 and 2010. 688 households were observed for the full 14 years since the project inception out of a possible 956 households that were resurveyed in 1998. As before, individual-level characteristics such as education and sex were obtained from the household heads who were assumed to represent the households. Household-level variables such as incomes are aggregated at the household level first before being converted into real

equivalised incomes.

The need to use a representative household head to obtain individual-level characteristic variables further reduced the sample size as although some households were observed for 14 years, the household heads were missing, had been replaced or had died in some years. Though an effort was made to reconcile the change in household head a further 38 households were dropped, which left a final 650 households for the final analysis. The following reconciliation of the change in the household heads were employed:

- If the household heads had died and were replaced by a member of the household such as the spouse or one of the children, who had also been present in the household from the beginning, then we adjusted the data so that the person who later took over became the head of the household from the start in 1997.

- If the household head had retired in an earlier period (within the first 7 years of the 14-year period) and was taken over by a member who had been present in the household from the beginning, then we adjusted the data so that the person who later took over became the head of the household from the start in 1997.

- If the household head had retired in a later period (after the first 7 years of the 14-year period) and was still living in the household in the last period (2010), then we adjusted the data so that the retired household head remained the head of the household until the last period in 2010.

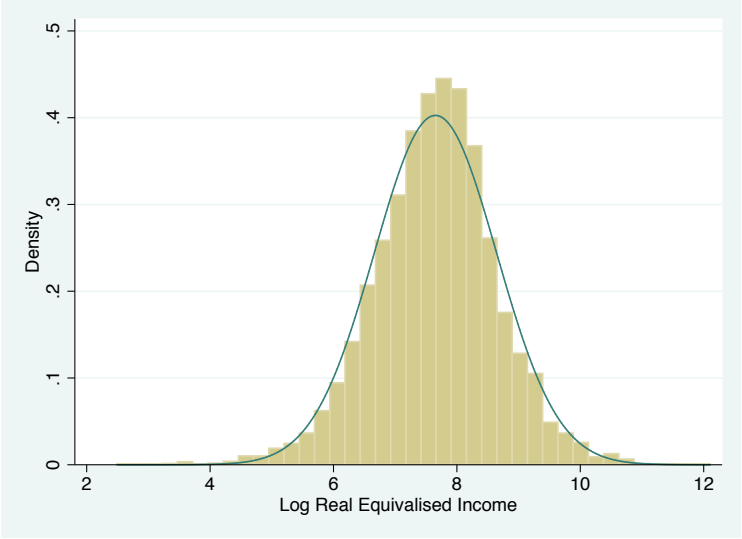
3.4.2.2 Key Variables Construction

The key variable used for analysis in the Townsend Thai panel data is household income. The Townsend Thai panel study only collected data on household-level incomes as opposed to individual earnings and incomes collected by the LFS. Household incomes collected in the survey include wage earnings, as well as net profit from farms and businesses. Similar to before, household incomes are converted to real equivalised monthly incomes using Thailand CPI and the Townsend Indian equivalence scale.

Figure 3.7 shows the histogram of the logarithm of real equivalised household income, which is approximately log-normal. Figure 3.8 plots average real monthly equivalised incomes with real GDP per capita. Both series have been normalised to one in the year 1997. The

Townsend Thai data average real equivalised income accumulated growth tracks the real GDP per capita accumulated growth figure relatively closely both in terms of level and trend. The average cumulative incomes growth of the Townsend Thai panel data track the real GDP per capita series better than the LFS incomes and earnings data, but it is worth noting the wider confidence interval, which suggests that the average incomes in the Townsend Thai data may be less precisely estimated. This observation could possibly be explained by the nature of farm income, the main source of income for the Townsend sample, which is mainly driven by macroeconomic factors such as the global prices of agricultural products. Thailand GDP is driven mostly by export, so it is not surprising that the two series track each other much better than the LFS earnings series.

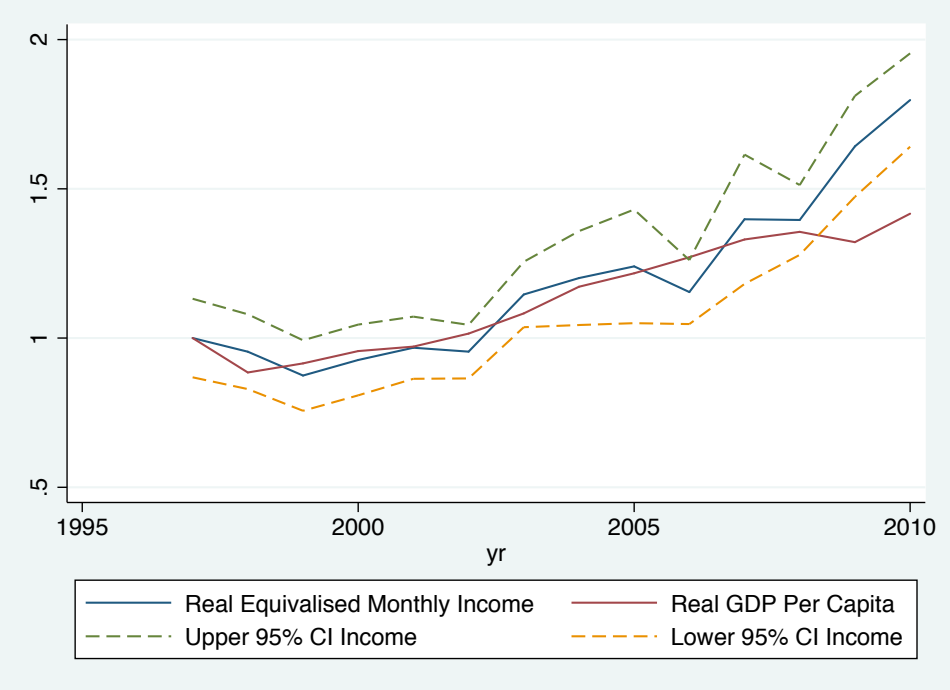
Figure 3.7: Histogram of Log Real Equivalised Income



3.4.2.3 Descriptive Statistics

Table 3.5 shows the descriptive statistics of key variables that are used in the analysis. As we have mentioned earlier in the discussion of the LFS summary statistics, the Townsend Thai data summary statistics exhibit the characteristics of a less developed region with lower income, lower level of educational attainment and higher proportion of people working in the informal agricultural sector. The average real equivalised income is around 1,000 Bahts less than the average income and 500 Bahts less than real equivalised earnings observed in the LFS sample. It is worth noting the large standard deviations of the income variable, which suggest that the estimate may be imprecise. Education attainment of 4.3 years is around 2

Figure 3.8: Real Equivalised Incomes with Real GDP



years less than in the LFS sample. In addition, the average age of the sample is much older at 53 years of age. 75% of the sample works in the agricultural sector as opposed to only 20% in the LFS sample. All of this is again consistent with the fact that the majority of the Townsend Thai sample is made up of self-employed agricultural workers as the survey are conducted in the rural farming provinces.

Figure 3.9 plots various quantiles of household income from the Townsend data for the period from 1997 to 2010. Similar to the result obtained for the LFS data, these quantile plots suggest that income for the Townsend household has grown at a different rate, with households at the top seeing their incomes increase by more than those at the bottom leading to a divergence in the income distribution. This again provides a support for the use of quantile regression framework to analyse income dynamics beyond the mean.

Figure 3.10 shows the APC decomposition obtained using the maximum entropy method. In all three panels we observed large standard errors, which resulted in a relatively wide confidence intervals in all three graphs.

In contrast with the result of the LFS sample, the age and cohort profiles of the Townsend Thai panel sample are very flat and lie around zero, with the exception of a spike at the extreme, as a result of imprecise estimations. This finding of a flat and zero age and cohort effects may be consistent with the usual experience of self-employed agricultural workers

and farmers. Farmers' incomes tend to vary a lot from year to year as crop prices and the general economic condition change. These macroeconomic factors are not related to the age or the period in which farming households were born, which could explain why these latter factors have less or almost no effect on farming household earnings as observed above. Lastly, similar to before, the top right panel which shows the period effect produces a graph that resembles the real GDP from 1997 onwards to 2010. As with the LFS, we will also attempt to gauge an estimate of income persistence of the Townsend Thai data later in the chapter while controlling for the APC effects all at once.

Table 3.5: Summary Statistics for Townsend Panel Data

	count	mean	sd
Real Equivalised Income	9032	3459.711	5367.051
Age	9100	53.03231	12.80565
Female-headed household	9100	.3476923	.4762639
Household size	9100	4.388242	1.814915
Years of Education	9100	4.281538	2.807145
Proportion in Agrictural Sector	9100	.7459341	.4353588
Proportion in Manufacturing Sector	9100	.0781319	.2683937
Proportion in Service Sector	9100	.1187912	.3235604
Proportion Inactive	9100	.0571429	.2321281

Figure 3.9: Quantile Plot: Townsend

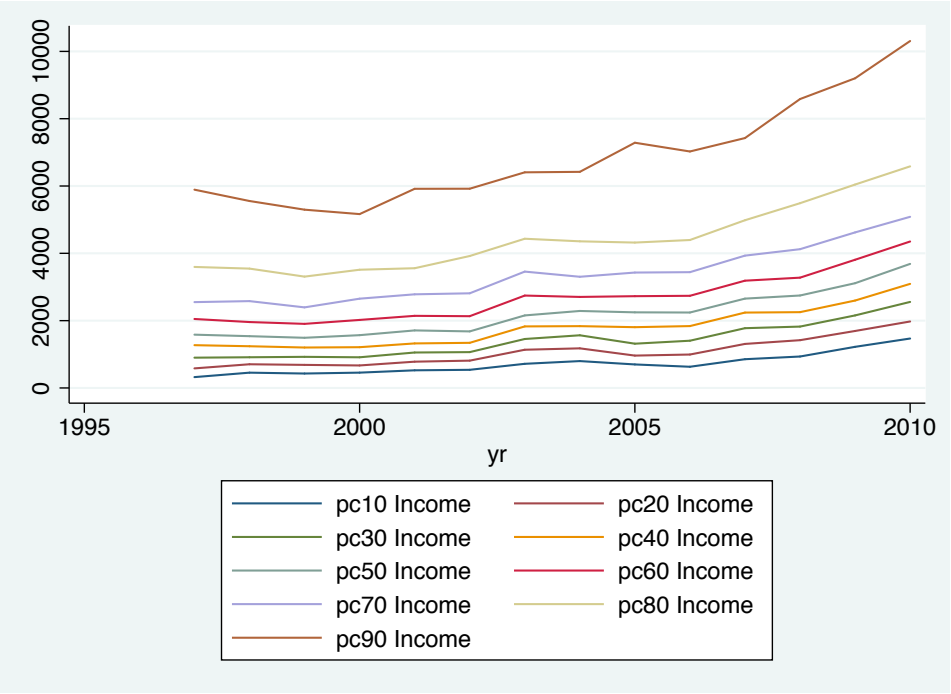
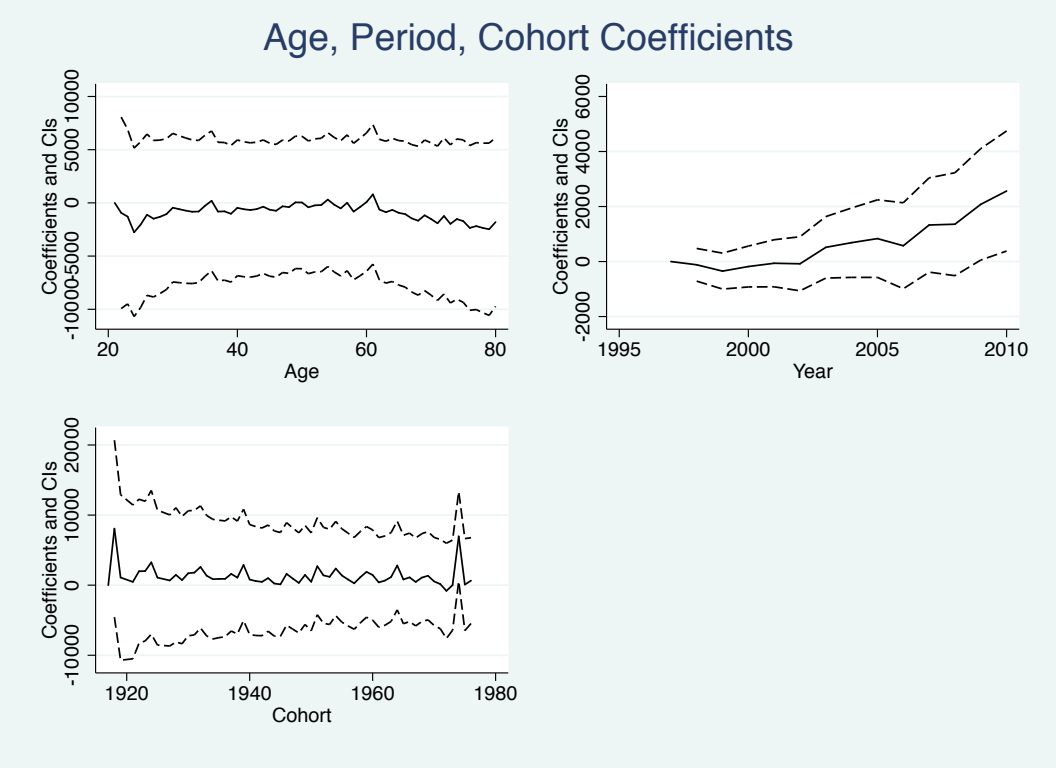


Figure 3.10: Age Period Cohort Decomposition for Townsend Thai Data



3.4.3 Summary

While it was originally intended that the panel estimation of income persistence using the Townsend Thai panel data may serve as a validation for the pseudo-panel estimates, we have observed above that the two data sets are distinctly different. The LFS, which contains rich earnings data for wage-employed workers, lacks data on the incomes of self-employed agricultural workers. On the other hand, the Townsend Thai panel sample focuses almost exclusively on rural self-employed agricultural workers. The result from the Kolmogorov-Smirnov equality-of-distributions test between income data in the Townsend and the LFS samples in Table 3.6 shows that the two samples are statistically different, with the majority of the Townsend incomes having values smaller than those in the LFS.

Table 3.6: Kolmogorov-Smirnov Test of Townsend and LFS Sample

	D	P-value
1:	0.1944	0.000
2:	-0.0002	0.999
Combined K-S:	0.1944	0.000

The first row (1:) tests the hypothesis that incomes for Townsend contains smaller values than for LFS.
The second row (2:) tests the hypothesis that incomes for Townsend contains larger values than for LFS.

These fundamental differences may not allow the two persistence estimates obtained using the two different samples to be compared directly with each other but it would still be interesting to see what we can learn about income dynamics in Thailand and the differences that may exist between urban wage-employed workers and rural self-employed agricultural workers.

In the next section we present the results of all the estimation of income and earnings persistence in Thailand using both LFS pseudo-panel and Townsend Thai panel data.

3.5 Results

This section presents the estimates of earnings and income persistence obtained using methods discussed in Section 3.3. We first present the OLS and quantile estimates of the Galton-Markov earnings model from the Townsend Thai panel data and the LFS pseudo-panel data. We then proceed to present the findings on income mobility from the transition matrix constructed from the Townsend panel data. Lastly we present the estimates of the permanent earnings model for both the Townsend and the LFS pseudo-panel data.

3.5.1 Galton-Markov Earnings Model: OLS Estimates

3.5.1.1 Townsend Thai Panel Data

Table 3.7: Townsend Panel: Income Persistence

	(1) (OLS) Income	(2) (OLS) Income	(3) (FE) Income	(4) (IV) Income	(5) (IV) Income	(6) (AB-IV) Income	(7) (APC) Income
Lag Income	**0.463*****	**0.426*****	**0.108****	**1.021*****	**0.918*****	**0.192*****	**0.435*****
	(0.088)	(0.089)	(0.050)	(0.049)	(0.026)	(0.057)	(0.086)
Education		272.247***					
		(54.684)					
Household Size		-176.009***					
		(65.170)					
Constant	1960.961***	1683.587***	3144.573***	102.747	477.317***	2672.000***	1960.961***
	(289.741)	(399.403)	(165.993)	(174.994)	(107.533)	(218.201)	(62.127)
Observations	8319	8319	8319	8319	7614	8319	8319
R^2	0.208	0.230	0.011				
Hansen Test						**0.332**	
AR1 Test						0.004***	
AR2 Test						0.017***	
AR3 Test						**0.176**	

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Townsend Panel: Age, Period and Cohort Effects

	(1) Income	(2) Income	(3) Income	(4) Income	(5) Income	(6) Income	(7) Income
Lag Income	**0.461*****	**0.453*****	**0.446*****	**0.451*****	**0.436*****	**0.435*****	**0.435*****
	(0.089)	(0.089)	(0.087)	(0.090)	(0.089)	(0.088)	(0.086)
Age	Yes			Yes	Yes		Yes
Period		Yes		Yes		Yes	Yes
Cohort			Yes		Yes	Yes	Yes
Constant	17485.119***	1498.010***	1621.739***	2218.153***	17515.974***	1139.889***	-3.337
	(107.051)	(228.757)	(229.827)	(269.598)	(107.051)	(175.849)	(51.688)
Observations	8319	8319	8319	8319	8319	8319	8319
R^2	0.216	0.219	0.220	0.226	0.235	0.231	0.186

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7 presents estimates of income persistence obtained from the Townsend Thai panel data. Column (1) provides the OLS estimate β of unconditional income persistence corresponding to Equation (3.20) in Subsection 3.3.2. The estimate $\beta = 0.463$ is significantly less than unity and suggests that unconditional income persistence is moderate. Column (2) gives the OLS estimate of conditional income persistence, corresponding to Equation (3.22) in Subsection 3.3.2, by controlling for observable household characteristics. Controlling for education level and household size in Column (2) does little to change the magnitude of the persistence parameter indicating that income remains equally persistence even for households with similar observable characteristics.

Column (3) attempts to estimate conditional persistence by controlling for unobservable household fixed effect yielding the within estimate β of 0.108, which although statistically different from zero, is close to zero in magnitude. This suggests that conditional persistence in households' income is low, in the sense that their incomes exhibit a great level of independence over time, once the household fixed effect is accounted for. Comparing the estimate in Columns (1) and (3), we can deduce that the household fixed effect is positive. This coincides with the positive ability bias story, where the higher level of persistence observed may be due to the difference in individual's ability. Those with higher ability will tend to have a higher and more persistent level of wages over time than those with lower level of ability, which explains why the degree of persistence falls once these ability differences are controlled for by the within model.

However, the within estimate in column (3) will be bias downward due to the introduction of the average terms that cause the transformed lagged income to be negatively correlated with the transformed error term (Nickell, 1981). This negative correlation only converges to zero as $(T-1) \to \infty$. In addition, performing the within transformation will magnify the relative importance of the measurement errors hence, accentuating the attenuation bias (Griliches and Hausman, 1986). Therefore, the within $\hat{\beta}$ in Column (3) will overstate the degree of income persistence by some magnitude.

Columns (4) and (5) attempt to fix this attenuation bias problem using the traditional instrumental variable approach. Column (4) estimates the degree of unconditional income persistence by using the head of household's years of education as an instrument for lagged income. Education is believed to be correlated with income while being uncorrelated with

the measurement error in income that caused attenuation bias in persistence estimates. However, as shown in Equation (3.41) in Subsection 3.3.2.4, a necessary assumption for a valid instrument in the dynamic panel setting with measurement error also requires that the degree of autocorrelation in the instrument match the degree of income persistence perfectly. When all other conditions except this one are satisfied, the instrumental variable estimate will be biased and converge in probability to the degree of autocorrelation of the instrument. As the instrument used in this case is education, which is highly autocorrelated over time, our estimate of income persistence in Column (4) is 1.021, which mimics the high persistence in education in accordance with our theoretical prediction.

Column (5) uses two-period-lagged income as an instrument of the one-period-lagged income. Using further lags of income as instruments ensured that the stringent assumption that the degree of autocorrelation in the instrument must match the degree of income persistence perfectly is satisfied, as lagged income follows the same Galton-Markov process. However, as discussed in Section 3.3.2.5, the IV estimates obtained in this case will be inconsistent as two-period-lagged income will not be completely exogenous from the error term that contains the unobserved individual fixed effect. Given the positive correlation between lagged income and the unobserved individual fixed effect term established from the result in Column (3), IV estimates in Column (5) of 0.918 will be an over-estimate of the true degree of persistence in the sample. These two exercises illustrate that past instruments that were used to study income dynamics such as education or anthropometric measures will not provide consistent estimates of the degree of persistence. In fact, given the very stringent assumptions required, it may not be possible to find a valid external instrument for this process at all in practice. Internal instruments are also unlikely to provide a plausible alternative, especially in the presence of the unobserved fixed effect that renders the instrument invalid.

The estimate in Column (6) employed the AB method using three-period-lagged income as instruments. We showed in Subsection 3.3.2.5 that this should provide a consistent estimate of conditional income persistence in the presence of classical measurement error. It is not possible to estimate unconditional persistence using the AB method as the technique involved taking the first difference to remove the unobserved household fixed effect. The estimate of conditional persistence in Column (6) is 0.192 and it is statistically different

from zero. The Hansen test of over-identifying restrictions statistic cannot reject the null hypothesis that the instruments used are exogenous. The result obtained in Column (6) shows AB estimate to be larger in magnitude than the within estimate, which confirms our theoretical prediction that the conditional persistence estimates in Column (3) suffered from attenuation bias as a result of measurement error. In addition the test for autocorrelation in the differenced residual rejects the null hypothesis that there is no AR(1) and AR(2) process but fails to reject the null hypothesis that there is no AR(3) process. This result provides support for the assumption of classical measurement errors in income, which will only introduce an AR(2) process for the differenced error term and no higher degree of autocorrelations, and also provides support for the use of three-period-lagged income as an instrument. However, the magnitude of the coefficient is still relatively close to zero, suggesting that conditional income persistence in rural Thailand is low.

Column (7) presents an estimate of income persistence conditioned on APC effects all at same time using the maximum entropy method corresponding to Equation (3.39) in Subsection 3.3.2.3. This gives an estimate of 0.435, which still suggests a moderately high degree of conditional income persistence but is not much different from the OLS estimate of unconditional income persistence in Column (1). Table 3.8 presents regression equations with various combinations of APC effects as control variables. Estimates from Table 3.8 are all similar in magnitude and suggest that APC effects do not have a major impact on the degree of income persistence. While this may seem surprising, it is consistent with the APC decomposition of income in the Townsend panel data in Section 3.4, where the cohort and age effects on income are flat and negligible. Similar to other estimates, APC persistence estimates suffered from attenuation bias as a result of measurement error.

3.5.1.2 LFS Pseudo-panel Data

Table 3.9 presents estimates of earnings persistence obtained from the LFS pseudo-panel data. Column (1) gives the estimate of unconditional earnings persistence, 0.727, which indicates a high degree of persistence in cohort earnings. As an estimate from the pseudo-panel does not suffer from measurement error (both classical and non-classical) and non-random attrition, the estimate of persistence is higher than that obtained using the Townsend Thai panel as we would expect. This high degree of persistence observed indicates that

earnings differences between rich and poor households in Thailand are likely to remain over time.

Column (2) extends the analysis to conditional earnings persistence corresponding to Equation (3.48) in Subsection 3.3.3 where additional variables including education, household size, urban dummy, sector dummy and year dummy variables are added to the regression but not the cohort fixed effect. The direction of the effect of each conditioning variable is as expected. The effect of education on earnings is positive, while household size is negatively related to real equivalised earnings. However, both variables are not statistically significant and their inclusion in the earnings regression on their own does little to lower the high degree of persistence observed.[13] The urban dummy variable is positive and strongly significant, indicating that those who lived in urban areas are better off than those in rural areas. The sector dummy also has the expected sign as those people who work in agriculture and manufacturing are on average worse off than those working in service sector, the reference category in this case. The estimate of conditional earnings persistence in Column (2) is significantly lower than that in Column (1), suggesting that on controlling for differences in household characteristics, the degree of earnings persistence in Thailand is low. However, it must be noted that quite a substantial part of the lowering in the magnitude of the persistence estimates was as a result of the inclusion of the year effect; the estimated regression that contains the same set of variables except the year dummies gives a value of $\beta = 0.603$, only a slight decrease from the unconditional estimates in Column (1).

Columns (3) to (9) estimate earnings persistence for various combinations of APC effects. As we are using pseudo-panel data which tracks cohorts over time, including the cohort dummy in the regression is the same as controlling for the household fixed effect to obtain the within estimator in standard panel data. Including just the age, period or cohort effect on its own reduces the degree of persistence in earnings over time but not by much, except for the estimate in Column (4), where persistence in income has decreased reasonably once the year effect is taken in to account. In Column (5), where only the cohort fixed effect was controlled for, the degree of unconditional earnings persistence reduces only slightly to 0.69, a moderately high degree of persistence still. Controlling for each pair of APC effects reduces the conditional earnings persistence to a reasonably low degree, especially in the case of the pairs that contain the period effect. The estimate in Column (9), which is the estimate of

[13]The regression is not shown here but the coefficient β obtained from this regression is 0.716.

earnings persistence controlling for the APC effects all at once, indicates a very low but significant degree of conditional earnings persistence in the LFS data. If we believed that age, period and cohort each has a distinct and independent effect on the earnings path that should be taken in to account, the estimate in Column (9) would suggest that controlling for these effects, conditional earnings persistence in Thailand is low and that earnings differences between initially rich and initially poor households with similar characteristics are likely to diminish over time.

Table 3.10 presents a set of similar estimated equations to those in Table 3.9 but focussing on income as opposed to earnings. Income in the LFS is defined to include those individuals who receive income from farming or profit from owning businesses in addition to those whose labour market earnings are the main source of income. As data on farm and business profits were only collected up to the year 2000, we have a smaller number of observations than in the case of the earnings regression.

The unconditional income persistence estimated in Column (1) is found to be high and is similar in magnitude to earnings persistence (0.731 vs. 0.727). Introducing the same set of control variables as in the conditional earnings regression to the income regression in Column (2) reduces the degree of persistence observed substantially, with most of the lowering effect driven by the year dummies. The estimate of β in Column (2) is not statistically different from zero. Columns (3) to (9) again give estimates of income persistence for various combinations of APC effects. Including the age, period or cohort effect on its own in the regression reduces the degree of persistence in income over time, with the period effect reducing the persistence the most. This observation is the same as what was observed earlier for earnings dynamics but the stronger reduction in the degree of persistence once the year effect is taken into account would suggest that income is more affected by macroeconomic shocks than earnings. This is expected as labour market earnings of waged employees will be less likely to be affected by macroeconomic shocks than incomes, which also constitute profit from farming and businesses. The magnitude of APC estimate in Column (9) is small and statistically insignificant from zero, which possibly indicates a case of no income persistence in Thailand once the APC effects are controlled for.

Antman and McKenzies's estimates of unconditional persistence correspond to our estimate in Column (1) of Table 3.9 and 3.10, while their estimates of conditional persistence

correspond to our estimates in Column (5) of Table 3.9 and 3.10, where only the cohort fixed effect is controlled for. The authors find that while unconditional earnings persistence in Mexico is high, conditional earnings persistence is low. They interpret their findings as indicating that while households may be able to recover from shocks to labour earnings (due to low conditional persistence), the high level of income inequality will likely persist for a long time (due to high unconditional persistence) as the fixed effect, which captures everything specific to households including education, institutional environment or innate ability, is an important factor in determining the degree of persistence. Furthermore, the differences in fixed effects between households will take a long time to disappear if at all.

Our results can be interpreted in a similar manner. We observed that both of our corresponding estimates of unconditional and conditional earnings and income persistence, controlling for the cohort fixed effect, are high. However, our own estimate of conditional persistence once APC effects are taken into account shows conditional persistence to be low in Thailand, which would suggest that households are likely to recover quickly from shocks to labour earnings that may cause their earnings to temporarily deviate from their path. Nevertheless, it must be noted that all three effects – age, period and cohort – can be viewed as exogenous in the sense that they are all beyond the control of the household. If these three effects are the main effects that govern the dynamic process of income and earnings in Thailand, then it is plausible to interpret from the high unconditional persistence estimate that inequality will persist over a long time in Thailand, which fits in with the stylised fact showed earlier in Figure3.1.

Lastly, the use of pseudo-panel analysis allows us to examine persistence over longer time periods than with standard panel data. Table 3.11 provides estimates of the persistence coefficient over one-year, two-year, and five-year time periods. Columns (1) to (3) provide the estimates of unconditional earnings persistence, while Columns (4) to (6) include cohort fixed effects and therefore give measures of conditional persistence. Unconditional and conditional earnings persistence decreases as the time frame increases. A household that has THB 100 higher earnings than another household today is estimated to only have THB 32 higher earnings five years later. Similarly, a THB 100 difference in earnings between two households with the same fixed effect is reduced to a THB 69 difference after one year, and only a THB 27.5 difference after five years.

Table 3.9: Pseudo-Panel: Earnings Persistence

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Lag Earnings	**0.727*****	**0.252*****	**0.669*****	**0.446*****	**0.690*****	**0.202*****	**0.400*****	**0.322*****	**0.126*****
	(0.031)	(0.043)	(0.034)	(0.047)	(0.032)	(0.048)	(0.038)	(0.048)	(0.049)
Education		41.567							
		(38.678)							
Household Size		-40.243							
		(74.628)							
Urban Dummy		2213.160***							
		(353.388)							
Agriculture		-1571.719***							
		(217.957)							
Manufacturing		-2271.807***							
		(347.814)							
Age			yes			Yes	Yes		Yes
Period		Yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	1291.468***	936.993	2586.938	1670.203***	1190.429***	3462.270**	1740.619	1567.974***	0.000
	(120.561)	(644.242)	(1739.937)	(194.644)	(169.658)	(1475.839)	(1633.799)	(215.000)	(0.000)
Observations	650	650	650	650	650	650	650	650	650
R^2	0.536	0.726	0.620	0.627	0.554	0.744	0.682	0.666	

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.10: Pseudo-Panel: Income Persistence

	(1) Income	(2) Income	(3) Income	(4) Income	(5) Income	(6) Income	(7) Income	(8) Income	(9) Income
Lag Income	**0.731*****	**0.058**	**0.704*****	**0.083***	**0.729*****	**0.101***	**0.383*****	**0.101***	**0.015**
	(0.034)	(0.048)	(0.038)	(0.050)	(0.036)	(0.053)	(0.054)	(0.051)	(0.054)
Education		211.238***							
		(51.454)							
Household Size		251.672***							
		(78.164)							
Urban Dummy		2841.837***							
		(467.950)							
Agriculture		1170.928**							
		(542.339)							
Manufacturing		-1387.034***							
		(378.594)							
Age			yes			Yes	Yes		Yes
Period		yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	1440.023***	-1906.839**	89.992	2855.346***	1392.661***	2134.715**	-1268.097	2134.715**	0.000
	(153.914)	(789.178)	(1119.058)	(221.402)	(235.356)	(914.349)	(1109.535)	(914.349)	(1.091)
Observations	375	375	375	375	375	375	375	375	375
R^2	0.600	0.804	0.658	0.760	0.605	0.793	0.705	0.793	0.697

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Pseudo-Panel: Earnings Persistence over Different Time Interval

	(1) (1-Year) Earnings	(2) (2-Year) Earnings	(3) (5-Year) Earnings	(4) (1-Year) Earnings	(5) (2-Year) Earnings	(6) (5-Year) Earnings
Lag Earnings	**0.727*****	**0.635*****	**0.320*****	**0.690*****	**0.591*****	**0.275*****
	(0.032)	(0.028)	(0.029)	(0.033)	(0.028)	(0.029)
Cohort Effect				Yes	Yes	Yes
Constant	1291.468***	1765.390***	3301.890***	1190.429***	1642.687***	2952.325***
	(120.561)	(131.108)	(133.838)	(169.658)	(184.134)	(194.833)
Observations	650	625	550	650	625	550
R^2	0.536	0.430	0.161	0.554	0.461	0.262

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.5.2 Galton-Markov Earnings Model: Quantile Estimates

Estimates of persistence obtained in Subsection 3.5.1 only concentrate on one part of the income/earnings distribution, the conditional mean. It is possible that the income dynamic process will differ at different parts of the conditional distribution; hence, requiring the degree of persistence to be the same across all individuals in the entire distribution may be too restrictive and unrealistic.

In addition, as shown in Subsection 3.3.1.3, estimates of persistence obtained from OLS regressions do not provide any extra information on the shape of the conditional distribution. In this subsection, we employed the method of quantile regression for the Townsend Thai panel data and LFS pseudo-panel data to study the dynamic process at different parts of the earnings distribution and to be able to investigate its shape.

3.5.2.1 Townsend Thai Panel Data

Table 3.12 presents the estimates of unconditional persistence from quantile regressions at selected values of quantiles. Columns (2) – (6) of Table 3.12 presents the estimates of unconditional persistence at the 10th, 25th, 50th, 75th and 90th quantiles respectively. Column (1) of the same table presents the OLS estimate of unconditional persistence for comparison with the quantile regression results. It is clear from the estimates in Table 3.12 that the degree of persistence varies across different parts of the income distribution. At the lower end of the income distribution persistence is low, with estimates of 0.144 and 0.349 at the 10th and the 25th quantiles respectively. These estimates are lower than the OLS estimate. As we move across the income distribution we find an increasing level of

income persistence, with the estimates being close to the value one at the 75th and the 90th quantiles. T-tests on these quantile coefficients indicate that they are all statistically different from each other.

Table 3.13 and 3.14 presents the estimate of quantile regressions of conditional and APC persistence corresponding to Equation (3.23) in Subsection 3.3.2 respectively. We refer to Table 3.14 as quantile APC persistence regression as the quantile estimates are used to provide direct comparison with the maximum entropy APC estimate. Even though performing a maximum entropy APC estimate directly on quantile regression is not currently possible, we use information drawn from the Townsend income decomposition in Section 3.4 to assist in the decision to include relevant control variables. The result of the income decomposition in Figure 3.10 clearly indicates that only the year effect seems to matter for the income process in the Townsend sample; hence, we include only the year effect in our quantile APC estimation. The results from both Table 3.13 and 3.14 gave a similar picture to the quantile estimates of unconditional persistence in Table 3.12. Persistence is low at the bottom end of the income distribution but gradually increases to a high level as we move to the higher end of the income distribution.

Figures 3.11, 3.12 and 3.13, which plot the quantile persistence coefficients over the 5th to the 95th quantiles for unconditional, conditional and conditional APC respectively, reinforce earlier findings with the observed wide range of values of the quantile estimates. These quantile estimates provide strong evidence against the assumption that the degree of persistence is homogeneous across all individuals as imposed by the OLS estimate of the Galton-Markov model discussed earlier.

In addition, we can use the results obtained from the quantile regression to study the shape of the income distribution. From Table 3.12 we can see that an increase in past income by THB 100 would lead to only a THB 14.4 increase in current income at the 10th quantile of the distribution, median income would go up by roughly THB 65 and income at the 90th quantile would increase by even more to roughly THB 115. In fact, what we are observing here can be interpreted as a divergence of the conditional income distribution where those at the top are getting richer and moving away from those at the bottom end of the distribution as time passes. This result as well as the descriptive quantile plot of the Townsend income data in Figure 3.9 illustrate clearly Galton's fallacy discussed earlier in

Section 3.3.1.3, that observing a value of $\hat{\beta} < 1$ does not provide us with any information about the unconditional or conditional convergence of the income distribution as is often falsely claimed in some studies. It is clear from the result obtained here that while the OLS estimates of both unconditional and conditional income persistence in the Townsend panel are significantly less than 1, we are observing a clear divergence instead of convergence in both the conditional and unconditional income distribution.

However, the quantile estimates of income persistence are affected by measurement errors and attrition bias, which caused the estimates to be biased. We therefore propose using quantile regression on dynamic pseudo-panel data, which is free from both measurement errors and attrition bias, to consistently estimate income and earnings persistence at different parts of the income distribution.

Table 3.12: Unconditional Quantile Regression: Townsend Panel

	(1) (OLS) Income	(2) (10th) Income	(3) (25th) Income	(4) (50th) Income	(5) (75th) Income	(6) (90th) Income
Lag Income	**0.463*****	**0.144*****	**0.349*****	**0.653*****	**0.935*****	**1.152*****
	(0.088)	(0.028)	(0.023)	(0.026)	(0.021)	(0.045)
Constant	1960.961***	453.617***	574.388***	746.506***	1164.892***	2285.172***
	(289.741)	(38.906)	(39.728)	(53.840)	(46.737)	(105.293)
Observations	8319	8319	8319	8319	8319	8319
R^2	0.208					

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.13: Conditional Quantile Regression: Townsend Panel

	(1) (OLS) Income	(2) (10th) Income	(3) (25th) Income	(4) (50th) Income	(5) (75th) Income	(6) (90th) Income
Lag Income	**0.426*****	**0.123*****	**0.329*****	**0.621*****	**0.882*****	**1.135*****
	(0.089)	(0.024)	(0.025)	(0.027)	(0.026)	(0.047)
Education	272.247***	70.227***	81.929***	90.468***	91.247***	140.424***
	(54.684)	(8.077)	(8.300)	(8.199)	(14.749)	(23.221)
Household Size	-176.009***	-45.348***	-59.209***	-77.958***	-147.614***	-259.394***
	(65.170)	(6.842)	(7.674)	(9.359)	(12.135)	(27.919)
Constant	1683.587***	422.007***	576.668***	808.492***	1600.588***	2830.461***
	(399.403)	(49.952)	(63.878)	(66.173)	(83.620)	(179.895)
Observations	8319	8319	8319	8319	8319	8319
R^2	0.230					

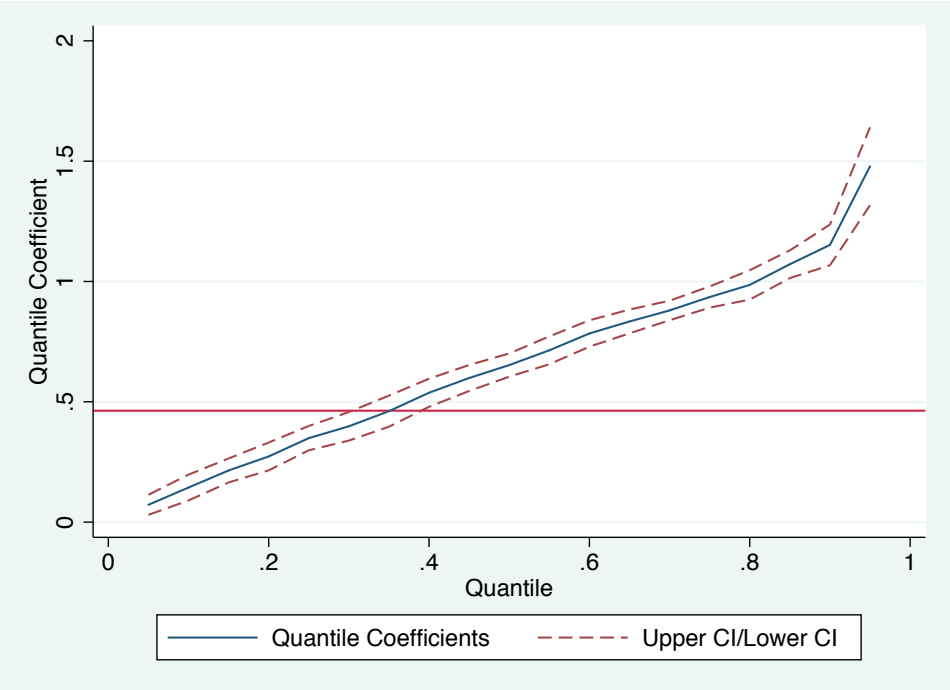
b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.14: APC Quantile Regression: Townsend Panel

	(1) (10th) Income	(2) (25th) Income	(3) (50th) Income	(4) (75th) Income	(5) (90th) Income
Lag Income	**0.130*****	**0.318*****	**0.615*****	**0.914*****	**1.144*****
	(0.024)	(0.029)	(0.028)	(0.024)	(0.034)
Age					
Year	Yes	Yes	Yes	Yes	Yes
Cohort					
Constant	301.588***	423.602***	544.167***	911.527***	1759.246***
	(40.238)	(43.700)	(58.381)	(85.038)	(227.209)
Observations	8319	8319	8319	8319	8319

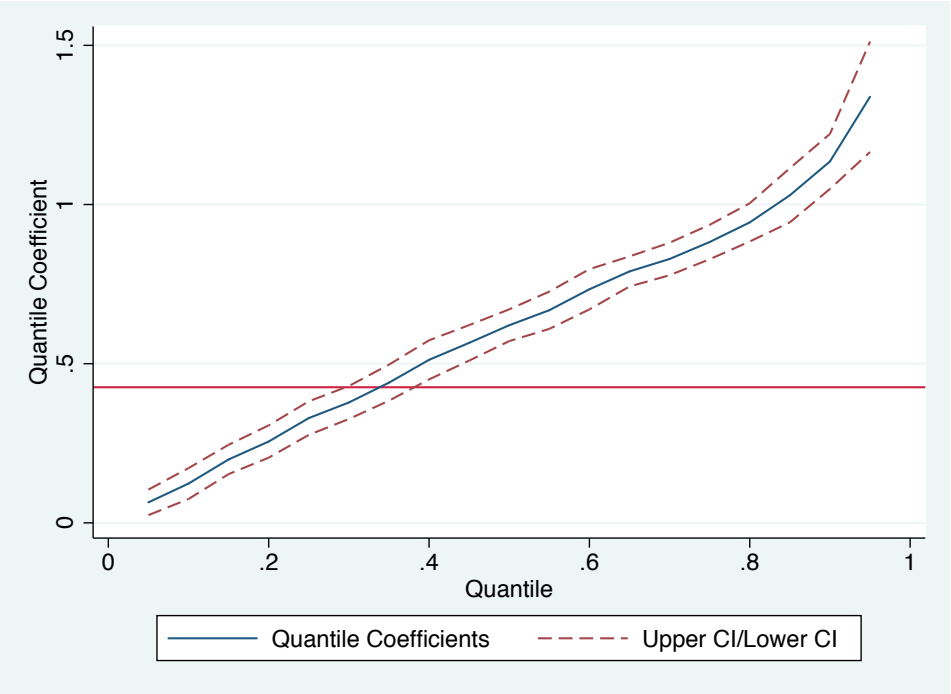
b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.11: Unconditional Beta vs. Quantile Plot: Townsend



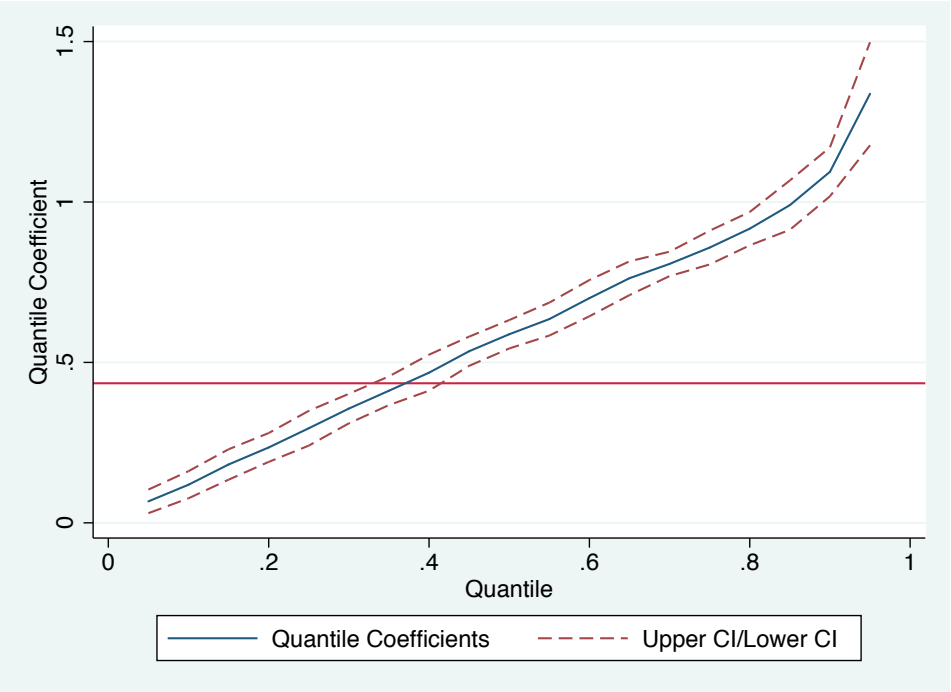
Note: The solid horizontal line is the corresponding OLS estimate

Figure 3.12: Conditional Beta vs. Quantile Plot: Townsend



Note: The solid horizontal line is the corresponding OLS estimate

Figure 3.13: Conditional Beta vs. Quantile Plot: Townsend



Note: The solid horizontal line is the corresponding OLS estimate

3.5.2.2 LFS Pseudo-Panel Data

Table 3.15 and 3.16 shows estimates from quantile regressions on LFS pseudo-panel data at selected values of quantiles. Similarly to the results obtained using OLS regression, estimates of unconditional and conditional persistence are greater in magnitude than the estimates obtained from the Townsend Thai panel data, which seems to suggest that measurement errors and attrition bias have resulted in attenuation biases in the quantile estimates for the Townsend panel.

Table 3.15 shows that the quantile estimates of unconditional persistence at various levels of quantiles are moderately high and importantly are not that different from each other, with most lying within the range of roughly 0.66 to 0.78. In addition, all the quantile estimates obtained have values that are close to the OLS estimate. While formal t-tests show that these coefficients are in fact statistically different from each other, the differences in the magnitude are small and the case against the OLS regression cannot be made as strongly here as in the case with the Townsend panel. In this case, it would seem that there is moderately high persistence throughout the entire earnings distribution but persistence is less at the bottom end and at the top end of the earnings distribution than in the middle of the distribution.

Table 3.16 presents the quantile estimates of conditional persistence corresponding to Equation (3.51) in Subsection 3.3.3. The finding here is very similar to findings of unconditional regression, with high persistence observed throughout the entire conditional earnings distribution and values of quantile estimates that are similar to the corresponding OLS estimates. Figure 3.14 and Figure 3.15, which plot the quantile coefficients over the 5th to the 95th quantiles for unconditional and conditional earnings persistence respectively, reinforce earlier findings. Both plots have a much narrower range of quantile estimates across the entire earnings distribution than similar plots for the Townsend Thai panel, which shows a very wide range.

Table 3.17 contains the conditional APC quantile regression estimates of the LFS pseudo-panel. Similar to the APC quantile regression of the Townsend data, the maximum entropy method was not applied directly but instead, information was drawn from the earnings decomposition in Section 3.4 to aid in the selection of APC control variables. The APC earnings decomposition in Figure 3.6 clearly indicates that all three effects of age, period and cohort are important and should be accounted for in the quantile regression framework.

We circumvent the problem of perfect multicollinearity here, by only capturing one aspect of the age effect, the retirement effect. This provides a plausible alternative to the full APC specification as the age effect observed is mostly flat and is roughly constant except for the sharp decline that was brought about by retirement. The estimates of APC quantile regression produce a finding that is more in agreement with findings from the Townsend panel than earlier estimates from unconditional and conditional quantile specification. Persistence is low at the bottom and gradually increases as we move up the earnings distribution. The quantile estimate plot in Figure 3.16 provides the same illustration.

Investigating the shape of the conditional earnings distribution provides mixed evidence. On the one hand, estimates from unconditional quantile regression seem to indicate that earnings are converging over time, with earnings growth at the top lower than earnings growth at the bottom end of the conditional distribution. On the other hand results from conditional and especially those from APC specifications show that the conditional earnings distribution is diverging over time, much like the income distribution in the Townsend sample. From Table 3.17, it can be seen that a THB 100 increase in past earnings would lead to an increase of THB 2 for those at the 10th quantile of the earnings distribution, median earnings would increase by THB 7, and earnings at the 90th quantile of the distribution would increase by THB 31. While this divergence in the earnings distribution is not observed as strongly as that in the Townsend sample, it still shows that conditional on some characteristics, those at the top are pulling away from those at the bottom, albeit at a smaller margin. Galton's fallacy also applies here. While the OLS estimate of unconditional, conditional and APC earnings persistence is below one, the conditional earnings distribution in the latter two cases shows divergence. So does the descriptive quantile plot of the LFS earnings data in Figure 3.5 as opposed to the false convergence interpretation based only on observing $\hat{\beta} < 1$.

Such drastic differences between the estimates of the two samples could possibly be reconciled by the fact that the two samples are made up of completely different sections of the Thai population as was discussed in Section 3.4. Being set in a rural area where most households work in the agricultural sector, the Townsend Thai sample is only representative of the rural farming households in Thailand whose income quickly fluctuates from year to year. On the other hand the LFS, especially after the year 2000, has an over-representation of people in waged employment earning salaries, which are more likely to remain stable over

time. In addition to the fundamental difference between the samples, the different choice of measurements of living standard may play a key role here. Income in the Townsend panel, which is made up of profits from farming, owning businesses and working as seasonal workers, is less likely to show persistence across time when compared against earnings which are mostly made up of salaries received from full-time waged employment. Despite these differences, studying both samples has led to greater understanding of the dynamics of living standards in the rural agricultural and wage-employed sectors in Thailand.

Table 3.15: Unconditional Quantile Regression: LFS

	(1) (OLS) Earnings	(2) (10th) Earnings	(3) (25th) Earnings	(4) (50th) Earnings	(5) (75th) Earnings	(6) (90th) Earnings
Lag Earnings	**0.727*****	**0.723*****	**0.741*****	**0.783*****	**0.719*****	**0.662*****
	(0.001)	(0.031)	(0.027)	(0.030)	(0.045)	(0.072)
Constant	1291.468***	648.752***	892.143***	1022.676***	1689.267***	2227.269***
	(5.354)	(236.095)	(114.873)	(152.549)	(128.052)	(203.603)
Observations	328599	650	650	650	650	650
R^2	0.536					

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.16: Conditional Quantile Regression: LFS

	(1) (OLS) Earnings	(2) (10th) Earnings	(3) (25th) Earnings	(4) (50th) Earnings	(5) (75th) Earnings	(6) (90th) Earnings
Lag Earnings	**0.603*****	**0.563*****	**0.542*****	**0.621*****	**0.615*****	**0.621*****
	(0.032)	(0.040)	(0.035)	(0.037)	(0.051)	(0.083)
Education	151.548***	214.328***	159.787***	157.300***	125.573**	107.561
	(39.338)	(57.251)	(51.582)	(39.654)	(58.649)	(71.913)
Household Size	-142.254*	-104.714	-238.232**	-179.509**	-115.461	-125.672
	(85.840)	(129.823)	(98.149)	(74.883)	(109.022)	(113.801)
Urban	-837.580***	-1400.442***	-1590.164***	-818.721***	-420.039	102.684
	(280.526)	(435.310)	(375.106)	(295.154)	(438.993)	(439.666)
Agriculture	-503.585**	-736.616**	-992.430***	-739.781***	-443.985	-136.331
	(226.863)	(315.349)	(260.587)	(208.275)	(303.986)	(365.069)
Manufacturing	-2142.613***	-2099.966***	-3086.743***	-2764.457***	-2293.503***	-2364.376***
	(376.436)	(492.142)	(438.554)	(369.575)	(626.944)	(717.782)
Constant	2458.048***	1857.298	3482.740***	2604.640***	2518.079**	2564.333**
	(716.451)	(1146.561)	(903.136)	(695.003)	(1040.459)	(1039.173)
Observations	650	650	650	650	650	650
R^2	0.591					

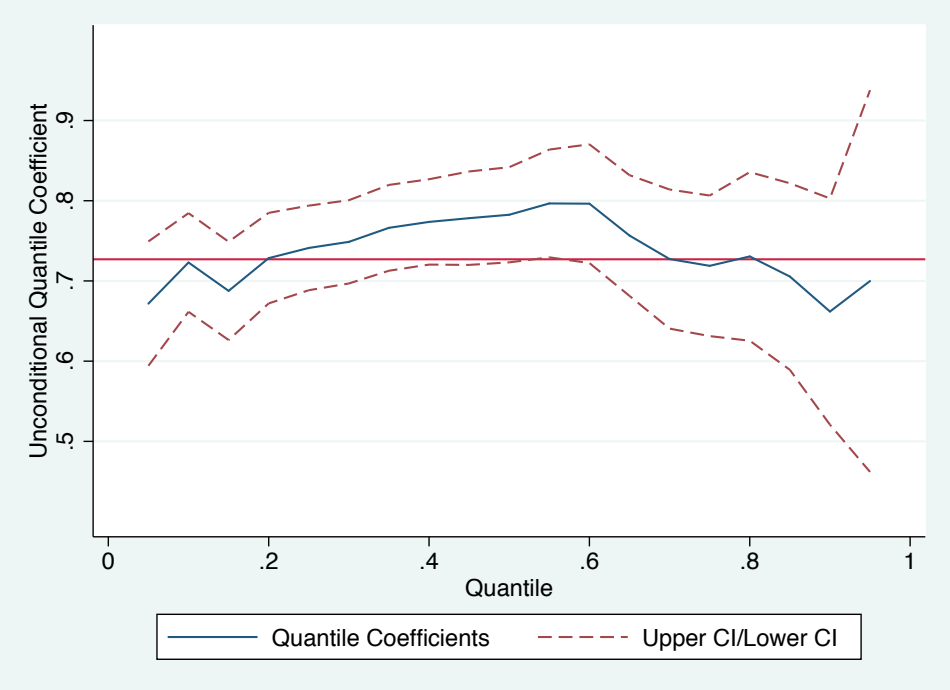
b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.17: APC Quantile Regression: LFS

	(1) (10th) Earnings	(2) (25th) Earnings	(3) (50th) Earnings	(4) (75th) Earnings	(5) (90th) Earnings
Lag Earnings	**0.020**	**0.065**	**0.070**	**0.196*****	**0.313*****
	(0.055)	(0.048)	(0.050)	(0.064)	(0.095)
Retire	-1055.989***	-1110.378***	-1141.900***	-952.961***	-478.720***
	(0.000)	(0.000)	(172.684)	(24.945)	(0.000)
Age					
Year	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
Constant	2172.765***	2236.168***	2470.874***	2289.103***	2244.082***
	(0.000)	(0.000)	(306.527)	(44.258)	(0.000)
Observations	650	650	650	650	650

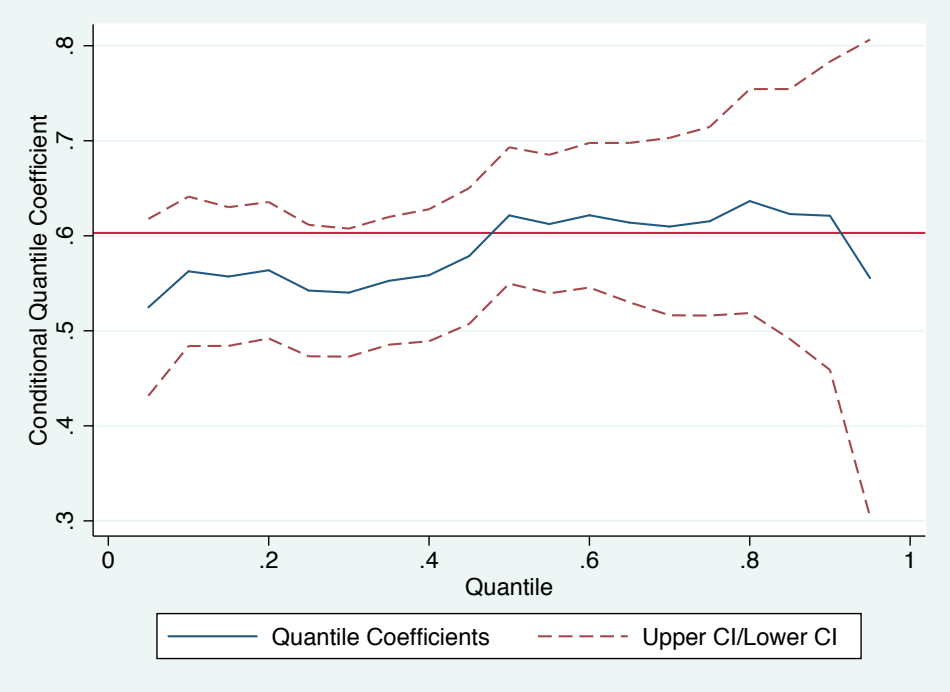
b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.14: Unconditional Beta vs. Quantile Plot: LFS



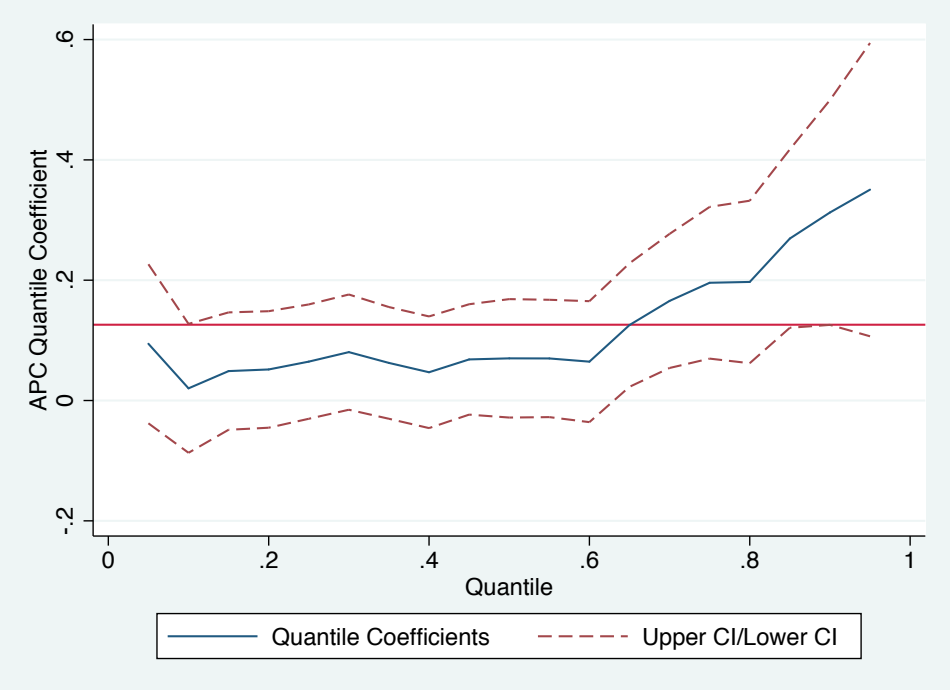
Note: The solid horizontal line is the corresponding OLS estimate

Figure 3.15: Conditional Beta vs. Quantile Plot: LFS



Note: The solid horizontal line is the corresponding OLS estimate

Figure 3.16: Conditional Beta vs. Quantile Plot: LFS



Note: The solid horizontal line is the corresponding OLS estimate

3.5.3 Transition Matrix

In this subsection we present the results from the transition matrix analysis. Unlike the regression-based method, a transition matrix offers additional information on the movement of individual across income quintiles, thus providing a measure of income mobility as well as a measure of income persistence complementing our earlier regression estimates. As previously mentioned in Subsection 3.2.2, it is not meaningful to construct a transition matrix for the LFS pseudo-panel data so the result in this subsection is solely based on data from the Townsend Thai panel.

Table 3.18: Transition Matrix for Townsend Panel Data

		2010				
		I	II	III	IV	V
1997	I	0.3769	0.2077	0.2231	0.1154	0.0769
	II	0.1846	0.2308	0.2154	0.2231	0.1462
	III	0.2154	0.2385	0.2308	0.2077	0.1077
	IV	0.1538	0.1538	0.2077	0.2615	0.2231
	V	0.0692	0.1692	0.1231	0.1923	0.4462

Summary Statistics		Reference Statistics	
		Perfect Mobility	Perfect Immobility
Average Quintile Move	**1.2215**	1.6	0
Immobility Ratio	**0.3092**	0.20	1
Adjusted Immobility Ratio	**0.6446**	0.52	1

The top panel of Table 3.18 presents the quintile transition matrix for Townsend households that were observed in the first year (1997) and the final year (2010) of the survey. Household are split equally into five income quintiles where the number "I" refers to the bottom quintile and the number "V" refers to the top income quintile. As discussed in Subsection 3.2.2, each element of the transition matrix p_{ij} represents the conditional probabilities of a household moving to income quintile j given that it has started off in income quintile i.

The transition matrix above looks closer to the case of perfect origin independence, with many entries being close to 0.2[14] except for some persistence at the bottom-bottom and top-top cells. Summary statistics presented in the bottom panel of Table 3.18 confirm this observation. Summary statistics that were used here are:

[14]Recall from Subsection 3.2.2 that a quintile transition matrix that exhibits perfect origin independence (taken to be the case of maximum mobility) will have all entries equal to 0.2. For the case of perfect origin dependence (a case of minimum mobility), the transition matrix is the identity matrix with entries equal to 1 on the diagonal and 0 elsewhere.

- Average Quintile Move (AQM):
  - This can be computed from the formula $AQM = \frac{1}{5}\{\sum_{i=1}^{5}\sum_{j=1}^{5}(|i-j|)p_{ij}$. The reference statistics in Table 3.18 provide the value of the AQM statistics for cases of perfectly origin-independent and perfectly origin-dependent mobility. If the transition matrix is perfectly origin-independent AQM would take the value of 1.6, while in the case of perfectly origin-dependent AQM will be equal to zero. Higher AQM would suggest a higher level of mobility.
- Immobility Ratio (IM)
  - This measures the fraction of households that remain in the same income quintile and is defined as $\frac{1}{5}\sum_{i=1}^{5}\sum_{j=1}^{5}(p_{ij})$ when $i = j$. If the transition matrix is perfectly origin-independent IM would take the value of 0.2, while in the case of perfectly origin-dependent IM will be equal to 1. Unlike AQM, a higher value of IM would suggest lower level of mobility.
- Adjusted Immobility Ratio (A-IM)
  - This measure extends the IM measure to include those who remain in the same quintile as well as those who move one quintile and is formally defined as $\frac{1}{5}\sum_{i=1}^{5}\sum_{j=1}^{5}(p_{ij})$ when $|i-j| \leq 1$. If the transition matrix is perfectly origin-independent A-IM would take the value of 0.52, while in the case of perfectly origin-dependent IM will be equal to 1. Like IM, a higher value of A-IM would suggest a lower level of mobility.

Summary statistics for the quintile transition matrix of the Townsend panel data seem to indicate a high level of origin independence mobility. The three summary statistics of the transition matrix all have values closer to the reference case of perfect mobility than perfect immobility. Interpreting mobility as being the converse of persistence, these results agree with our result obtained using the regression methods earlier in Subsection 3.5.1.1, which also suggests a relatively low level of income persistence in Townsend panel data.

The income persistence observed at the top and the bottom of the transition matrix provides an additional evidence that the degree of persistence may actually differ across the income distribution. However, it is important to note that the persistence observed

in the top-top and the bottom-bottom cells may be an artefact of the construction of the transition matrix, the so-called floor and ceiling effect, where an individual in the top quintile cannot move upwards and an individual in the bottom quintile cannot move downwards, thus creating a misleading picture of persistence.

Results from the quantile regression in Subsection 3.5.2.1 helps verify the observation of persistence at the top of the transition matrix. However, it would seem that the observation of persistence at the bottom end of the transition matrix is a false illusion. Results from the quantile regression suggest that income persistence is low at the bottom 10 percent of the income distribution but as the quintile matrix bunches up the bottom 20 percent of the income distribution together, this low degree of persistence is overlooked. If a decile transition matrix were used instead of the quintile transition matrix, one would observe a higher level of mobility (lower persistence) at the bottom end but the persistence at the top would still remain.[15]

In Subsection 3.5.1, 3.5.2 and 3.5.3 we have provided a detailed description of the dynamic process of income, earnings and theirs distribution based on the Galton-Markov earnings model and the transition matrix. In the next subsection, we move beyond the statistical model of earnings dynamics, to apply the greater economic structure of the permanent earnings model.

3.5.4 Permanent Earnings Model

In this subsection we move beyond the estimation of the statistical model of earning dynamics to the estimation of a model with a more structural permanent earnings model. The permanent earnings model estimated in this section provides a decomposition of income and earnings into permanent and transitory components. Observing the roles that these elements have in the earnings process will lead to a greater understanding of dynamic of living standards in Thailand.

3.5.4.1 Townsend Thai Panel Data

We first present the results of the estimated parameters of the various permanent earning models studied in this chapter as discussed in Subsection 3.3.4 for the Townsend panel data.

[15]The decile transition matrix is not shown here.

Table 3.19: Townsend: Permanent Earnings Model

	(1) AR1	(2) AR1	(3) ARMA(1,1)	(4) RG-AR(1)	(5) RG-AR(1)
σ_δ^2	0.2848***	0.2844***	0.2781***	0.4714***	0.5253***
	(0.0231)	(0.0230)	(0.0232)	(0.0532)	(0.1341
$\sigma_{\delta\lambda}$				-0.0169***	-0.0189***
				(0.0042)	(0.0065)
σ_λ^2				0.0006	0.0006
				(0.0005)	(0.0005)
σ_v^2	0.9529***	0.9436***	0.3153***	0.9273***	0.8888***
	(0.0621)	(0.0669)	(0.0411)	(0.0651)	(0.1107)
σ_{v0}^2		3..6365			
		(2.6131)			
γ	0.2132***	0.2152***	0.3286***	0.1833***	0.1815***
	(0.0241)	(0.0246)	(0.0288)	(0.0253)	(0.0287)
θ			0.9935***		
			(0.0756)		
π_1^P					1.00
π_2^P					0.9866
π_3^P					0.9217
π_4^P					0.7133
π_1^T					1.00
π_2^T					1.1647
π_3^T					1.0114
π_4^T					1.0077
Observations	9100	9100	9100	9100	9100

Bootstrap standard error in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) of Table 3.19 presents the estimate of the simplest version of the permanent earnings model with a fixed permanent element and a transitory component that follows an AR(1) process corresponding to the covariance structure in Equation (3.59). This is the same earnings model that was studied by Lillard and Willis (1978) in their pioneer work on modelling the covariance structure of earnings. The result from Column (1) indicates that both the permanent and the transitory elements are important parts of the income process as both are statistically significant. The significance of the fixed permanent component is consistent with the finding of the importance of the household fixed effect in the Galton-Markov earnings model for the Townsend sample discussed earlier. The transitory component of income dominates over the permanent component and represents around 75% of the total variation in income. This would suggest that the initial position in the income distribution is less important and that there is movement within the distribution to offset initial differences in income. In addition, the persistence of parameter γ, though statistically different from zero, is of sufficiently low value, which implies that transitory shock will tend not to persist strongly over time and will exhibit a strong degree of mean reversion. Overall, the result from

Column (1) supports the conclusion of a high degree of mobility in the Townsend sample

Column (2) provides the estimates of the same earnings model as Column (1) except that the initial variance of the transitory element σ^2_{v0} is also being estimated as part of the model corresponding to the covariance structure in Equation (3.61). The initial variance term is not significantly different from zero, while the magnitude and the statistical significance of the remaining parameters remains very close to those of Column (1). This evidence would suggest that the initial variance assumption made in the model of Column (1) is adequate and there is no need to extend the model further to include the estimation of the initial variance term.

Column (3) extends the basic model in Column (1) to allow the transitory element of income to follow an ARMA(1,1) process corresponding to the covariance structure in Equation (3.63). The moving average parameter θ is statistically significant and has a value very close to one, indicating that shocks to the dynamic process of the transitory component are not random and may represent deviation from the mean reverting effect. In this model the variance of the transitory component, while still statistically significant, has decreased substantially in magnitude. The size of the transitory component is now roughly equal to the size of the permanent component. In addition, the persistence parameter has increased in value from 0.2132 to 0.3286. The evidence in this model indicates a less mobile picture of the income process than the model in Column (1). On economic grounds, there is no reason to prefer the AR(1) or the ARMA(1,1) over the other. However, on examining the output of the non-linear optimisation program it is clear that the estimates from the AR(1) model achieve better convergence than those from the ARMA(1,1) model. For this reason we have chosen to model the transitory element as following an AR(1) process in a more advanced specification to ensure that we do not ask too much of the data.

Column (4) presents the estimates from the heterogeneous or random growth model corresponding to the covariance structure in Equation (3.66). This model relaxes the constancy assumption for the permanent element and allows some part of it to vary with age. The additional parameter σ^2_λ that was supposed to capture the random growth with age is not statistically different from zero. While the covariance term $\sigma^2_{\delta\lambda}$ is significant and negative, which could be interpreted as indicating convergence in the permanent component over the life cycle, the magnitude of this parameter is extremely small and is unlikely to have

any economic significance. The remaining parameters are all statistically significant and closely resemble their corresponding estimates in Column (1) except for the fixed permanent component, which is larger in magnitude.

Column (5) provides estimates of the same random growth model as Column (4) with an addition of the cohort factor loading parameters on the permanent (π^P) and transitory (π^T) components corresponding to the covariance structure in Equation (3.69). The factor loading parameters for cohort one (π_1^P and π_1^T), the oldest cohort, are normalised to one to allow identification. Results from the factor loading parameter estimates suggested two things: 1) variation of the permanent income component is lower for the younger cohorts, and 2) the exact opposite applies to the transitory income component, where variation are higher among the younger cohorts than the oldest cohort. Nevertheless, none of the cohort factor loading parameters is significantly different from the others, which indicates that the cohort effect is very small or non-existent. This is consistent with the APC decomposition of the Townsend income in Figure 3.10, which shows that the cohort effect in the Townsend sample is flat and negligible. This evidence from the random growth model in Columns (4) and (5) point to the basic AR(1) process as providing a superior model of the income dynamic process in rural Thailand, where the transitory element dominates over the permanent component and income shock persistence is low. This finding of high mobility in the Townsend sample and the preference for the AR(1) model provide positive support for both the Galton-Markov model used and the results found in the earlier section.

3.5.4.2 LFS Pseudo-Panel Data

This subsection presents the result of the estimation of various specifications of the permanent earnings model on the LFS pseudo-panel data.

Column (1) of Table 3.20 presents the estimates from the simple AR(1) model where the permanent component is assumed to be fixed over the life cycle corresponding to the covariance structure in Equation (3.59). The variance parameters for the permanent and transitory components are both very small, with the permanent component being statistically insignificant. While the transitory component is statistically significant and dominates over the permanent effect, its magnitude of 0.008 is too small to be considered economically significant. The degree of persistence in the transitory shocks over time is greater than in

the Townsend sample, which suggests that transitory shocks are more persistent for formal workers in Thailand than transitory shocks faced by rural workers. The result obtained in Column (1) is consistent with the evidence obtained earlier in the Galton-Markov regression. There, the cohort fixed effect was also found to be small and insignificant while persistence, once the year effect is taken into account in addition to the cohort effect, is also relatively low.[16] This corresponds to the result of Column (8) in the Galton-Markov regression of earnings in Table 3.9, where $\hat{\beta} = 0.322$.

Column (2) extends the simple AR(1) model in Column (1) by also estimating the initial variance term corresponding to the covariance structure in Equation (3.61). The initial variance term is not precisely estimated, pointing to a possible identification problem with this specification. Here again, the initial variance term is not statistically significant. Column (3) extends the model in Column (1) to allow the transitory element to follow ARMA(1,1) process corresponding to the covariance structure in Equation (3.63). The moving average parameter θ is not statistically significant, and as a result of this specification the persistence parameter γ also loses its significance. However, the permanent component is now statistically significant, while the the transitory element loses its significance for the first time despite being much larger in magnitude than the permanent component. The result in Column (3) would seem to indicate the failure of the ARMA(1,1) specification in capturing dynamics in the earnings process.

Column (4) provides estimates of the random growth specification corresponding to the covariance structure in Equation (3.66). Given the findings of the very small and insignificant magnitude of the permanent component in both the simple AR(1) model in Column (1) and the Galton-Markov estimation described earlier, it is not surprising that none of the permanent variance components in the random growth model is statistically significant. In this specification, the persistence parameter γ also takes a negative value for the first time though it is not statistically different from zero. The result from this model seems to indicate that the random growth model is not a suitable model of earnings dynamics in Thailand. All these results from Table 3.20 led to us to prefer the simple AR(1) model in Column (1) over the initial variance, the ARMA(1,1) and the random growth specification.

Lastly, it must be noted that the use of pseudo-panel data to identify parameters of the

[16] Recall that year effects were taken into account in the first stage regression of earnings on the year dummies before the construction of the empirical covariance matrix from the earnings residual.

permanent earnings model is not the same as its application to the Galton-Markov earnings model. If the earnings process in Thailand is truly generated by the Galton-Markov process with all its assumptions satisfied, utilising a pseudo-panel as a substitute for a genuine panel will identify the parameter of interest β as shown in Subsection 3.3.3. While the basic AR(1) model of the permanent earning model bears a close resemblance to the Galton-Markov process, other more complex specifications of the permanent earnings model violate the basic assumption of the Galton-Markov model and ask too much of the pseudo-panel data created from repeated cross-sections. These complex specifications such as the random growth model need to be studied with genuine panel data or at least with pseudo-panel data that is constructed from the same set of individuals to be able to achieve credible identification of the relevant parameters.

Table 3.20: LFS Pseudo-Panel: Permanent Earnings Model

	(1) AR1	(2) AR1	(3) ARMA(1,1)	(4) RG-AR(1)
σ^2_δ	0.0000	0.0000	0.0054***	0.0008
	(0.0000)	(0.0000)	(0.0014)	(0.0007)
$\sigma_{\delta\lambda}$				-0.0006
				(0.0000)
σ^2_λ				0.0000
				(0.0000)
σ^2_v	0.0080***	0.0061***	4.6520	0.0052***
	(0.0007)	(0.0012)	(2.9789)	(0.0006)
σ^2_{v0}		0.0427		
		(84426)		
γ	0.4844***	0.6357***	-0.0015	-0.0473
	(0.0581)	(0.1280)	(0.0011)	(0.0445)
θ			0.0007	
			(0.0005)	
Observations	675	675	675	675

Bootstrap standard error in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.5.4.3 Summary

In this subsection we extend our analysis of the dynamics process of income and earnings in Thailand beyond that of the statistical model of earnings by looking at various specifications of the permanent earnings model that seek to identify variances of permanent and transitory components in the earning process.

The results from this analysis are in agreement with those from our earlier analysis and help strengthen the finding of low conditional income and earnings persistence in Thailand for both the rural workers represented in the Townsend sample and the formal wage-employed

workers in the LFS pseudo-panel sample. The preference for the simple AR(1) model, which bears a close resemblance to the Galton-Markov model, also provides positive partial support for the use of the Galton-Markov process to model the dynamics of living standards in Thailand.

The next section concludes on all our findings and discusses the shortcomings.

3.6 Conclusions

In this study we investigate the dynamics of living standards in Thailand first through the statistical Galton-Markov model of earnings dynamics before proceeding to apply the greater economic structure of the permanent earnings model, which seeks to model the covariance structure of the earnings process by decomposing earnings into permanent and transitory components and studying their evolution over the life cycle and over time.

Empirical estimations of income and earnings persistence in Thailand were performed on constructed LFS pseudo-panel data and the Townsend Thai panel data. However, simple OLS estimates of the degree of persistence are likely to be biased as a result of measurement error in income, non-random attrition and non-representativeness of the sample. Various methods such as instrumental variable estimation, Arellano-Bond dynamic panel estimation, APC maximum entropy estimation and dynamic pseudo-panel methods were employed to ensure the consistent estimation of the degree of income and earnings persistence in Thailand.

Galton-Markov estimates obtained from the Townsend panel data found both unconditional and conditional income persistence to be moderately low. Pseudo-panel estimates indicate a high level of unconditional persistence in Thailand. However, the introduction of socioeconomic factors and APC effects lowers the persistence estimates significantly.

Extending the analysis to other parts of the income distribution using the quantile regression technique finds that households at different parts of the distribution experience different degrees of income persistence. In both samples, persistence is found to be low at the bottom of the distribution but high at the top, indicating a divergence in income over time. This finding questions the low degree of persistence observed on average found earlier, and clearly show the Galton's fallacy in previous works which interpret the observation of $\hat{\beta} < 1$ as indicating a convergence in the income distribution. This illustrates the importance of exploring other parts of the income and earnings distribution beyond the conditional mean to obtain a more complete picture of the income and earnings process.

A study of the covariance structure of earnings following the permanent earnings model finds that total variation in the earnings process is predominantly driven by moderately persistent transitory components following the AR(1) process. Permanent components are found to be statistically insignificant in the LFS sample, and while they are statistically significant for the Townsend sample, they account for only a quarter of total variation in

income in the sample. More complex models of the covariance structure such as the random growth model is rejected in favour of the simple AR(1) model with fixed permanent elements. These findings from the variance decomposition analysis lend support to the result obtained in earlier analysis and the use of the Galton-Markov process to model earnings dynamics in Thailand.

Overall, these findings of low conditional persistence in both samples suggest that households are able to recover quickly from shocks to their labour earnings. Therefore, the high level of inequality found in Thailand is not due to earnings shocks having long-term effects. However, the fact that unconditional persistence remains high, especially at the top of the income and earnings distribution, suggests that household characteristics and fixed effects are important and that the high level of inequality found in the cross-section is likely to persist over time.

Lastly, it must be noted that differences between the persistence estimates obtained from the two samples stem mainly from the fact that they are made up of completely different sections of the Thai population. The Townsend Thai sample is only representative of the rural farming households in Thailand, whose income quickly fluctuates from year to year, while the LFS sample, especially after the year 2000, has an over-representation of people in waged employment earning stable salaries. One major limitation of this study is that while each of the samples used has relative advantages over the other, neither possesses a complete set of the desired characteristics that allow for a comprehensive study of income dynamics. An ideal data set needs to be in a panel structure, be representative of the population studied and include earnings and income variables that are measured with a high degree of accuracy. Not only would this facilitate consistent estimations of persistence parameters, but it would also allow for a credible examination of income and earnings mobility of households in Thailand. Despite these shortcomings, studying both set of data has led to greater understanding of the dynamics of living standards in the rural agricultural and wage-employed sectors in Thailand.

3.A Appendix

3.A.1 Robustness Test

3.A.1.1 Regional Price

The results obtained in Section 3.5 were all based on the assumption that households in Thailand all faced the same prices and equivalised household earnings were all deflated by one general CPI of Thailand. This assumption will not hold true in practice as households in different regions will ultimately face different price levels depending on their living locations. A household in metropolitan areas such as Bangkok will face a higher price level than a similar household living in one of the rural and less developed provinces. In this subsection we perform a robustness test to check whether making this assumption has had a significant impact on the final result obtained.

We perform this robustness test by deflating households' equivalised earnings and income in both the LFS and the Townsend data by the Thailand Regional CPI instead. The Thailand Regional CPI provides a different CPI for each of the five regions in Thailand: Bangkok, North, Northeast, Central and South. Table 3.A.2 and Table 3.A.1 present the re-estimation of the Galton-Markov regression using this price adjustment for the LFS and the Townsend data respectively. The results from both tables are extremely similar to those obtained without the regional price adjustment. This indicates that the original results are robust to regional price variation.

Table 3.A.1: Townsend Panel: Regional Price Adjustment

	(1) (OLS) Income	(2) (OLS) Income	(3) (FE) Income	(4) (IV) Income	(5) (IV) Income	(6) (AB-IV) Income	(7) (APC) Income
Lag Income	**0.464*****	**0.427*****	**0.107****	**1.017*****	**0.911*****	**0.209*****	**0.437*****
	(0.088)	(0.089)	(0.050)	(0.049)	(0.026)	(0.056)	(0.088)
Education		275.827***					
		(55.790)					
Household Size		-175.650***					
		(66.368)					
Constant	1988.628***	1696.706***	3203.680***	104.360	498.240***	2658.436***	-3.600
	(293.947)	(405.797)	(169.050)	(177.012)	(108.255)	(216.936)	(52.310)
Observations	8319	8319	8319	8319	7614	8319	8319
R^2	0.210	0.232	0.011				

b coefficients; se in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.A.2: LFS Pseudo-Panel: Regional Price Adjustment

	(1) Earning	(2) Earning	(3) Earning	(4) Earning	(5) Earning	(6) Earning	(7) Earning	(8) Earning	(9) Earning
Lag Earning	**0.738*****	**0.259*****	**0.682*****	**0.462*****	**0.702*****	**0.211*****	**0.441*****	**0.333*****	**0.127*****
	(0.030)	(0.044)	(0.033)	(0.046)	(0.031)	(0.045)	(0.036)	(0.048)	(0.044)
Education		36.367							
		(38.670)							
Household Size		-62.402							
		(74.582)							
Urban Dummy		2055.544***							
		(352.951)							
Agriculture		-1602.164***							
		(218.724)							
Manufacturing		-2355.197***							
		(347.703)							
Age			yes			Yes	Yes		Yes
Period		Yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	1259.541***	1220.497*	2498.766	1661.389***	1155.884***	3482.360**	1703.476	1552.897***	0.000***
	(119.232)	(643.767)	(1730.260)	(192.301)	(168.297)	(1464.496)	(1641.251)	(212.465)	(0.657)
Observations	650	650	650	650	650	650	650	650	650
R^2	0.555	0.736	0.637	0.641	0.573	0.756	0.689	0.680	

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.A.1.2 Formal vs. Informal

As noted in Section 3.4 on data description, our final LFS sample is over-represented by earnings data from formal wage-employed workers as the survey stopped collecting data on income from informal workers in the year 2001. However, in our analysis, the real equivalised earnings variable is created by deflating household total earnings by the total number of people living in the household regardless of their labour market status. This could potentially cause a problem as members who are informal workers are included in the household size variable without making any contribution to total household earnings as their earnings are not collected (not because they do not have any earnings).

For example, we could observe in the sample two households that reported the same amount of total household earnings and the same household structure. Assuming that the household size is equal to five for both households, the first household consists of two wage-employed workers and three non-active dependents. The second household, on the other hand, consists of two wage-employed workers as before but is now made up of two active informal workers and only one dependent. While the final calculation of real equivalised income between the two households will be exactly the same, we would realistically expect the second household to have a higher level of real equivalised earnings as two more members are active participants in the labour market. This particular effect cannot be captured because earnings of informal workers are not collected by the survey. This problem could lead to a potential bias in the earnings estimates as households with more informal workers among their members are being relegated to a lower earnings status than households of the same size that have more members working in the formal sector.

In this subsection we correct for this bias by excluding all informal workers from the analysis. In the cases of our two households used as examples, the first household retains all five members while the second household would only be left with three members. Equivalised earnings of the second household will now be greater than those of the first household as we would expect. It must be noted, however, that this adjustment is not a perfect solution as it is based on the assumption that the informal workers earn enough income to support themselves. If the reality is that households pooled their incomes together and shared their earnings, with formal workers providing a greater contribution than informal workers, then the above adjustment will result in a positive bias towards households with lots of informal

workers in comparison to households with more members in the formal sector or inactive.

We perform this robustness test to check whether including informal workers in our original analysis of the LFS pseudo-panel has caused any significant bias. Table 3.A.3 shows that the result obtained from this robustness test is very similar to the original estimates without the adjustment. As the strategy to exclude all informal workers can present a bias in the opposite direction that now favours households with more members working in the informal sector, obtaining estimates that are not very different from our original result could be taken as indicating that the original biases are small and the Galton-Markov earnings parameters were estimated quite accurately.

Table 3.A.3: LFS Pseudo-Panel: Informal Workers Adjustment

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Lag Earnings	**0.788*****	**0.358*****	**0.600*****	**0.587*****	**0.770*****	**0.199*****	**0.319*****	**0.484*****	**0.094****
	(0.025)	(0.035)	(0.033)	(0.039)	(0.026)	(0.038)	(0.034)	(0.041)	(0.041)
Education		123.174***							
		(42.867)							
Household Size		58.090							
		(97.213)							
Urban Dummy		3269.294***							
		(407.865)							
Agriculture		-2503.919***							
		(378.790)							
Manufacturing		-3560.517***							
		(349.168)							
Age			yes			Yes	Yes		Yes
Period		Yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	1298.811***	-904.530	2910.054	1299.432***	1117.324***	3498.870*	1675.897	1196.707***	0.000
	(133.456)	(822.628)	(2056.540)	(252.044)	(226.163)	(1804.811)	(1916.772)	(299.911)	(0.982)
Observations	650	650	650	650	650	650	650	650	650
R^2	0.637	0.791	0.730	0.683	0.646	0.805	0.778	0.706	

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.A.1.3 Representative Member of the Household

As discussed in Section 3.4 on the creation of the LFS pseudo-panel data, it was noted that the pseudo-panel was created based on the characteristics of the household head following a conventional practice in the literature. In this subsection, we attempt to test the robustness of this practice by forming an alternative pseudo-panel based on the characteristics of a member of the household who contributes the most to household earnings.[17]

Table 3.A.4 provides the result of the estimation of the Galton-Markov earnings parameter applied to this alternative pseudo-panel. While the persistence estimates are not as close to the original in comparison to the regional prices and the informal worker adjustment cases discussed above, all new estimates with the alternative pseudo-panel are reasonably close to the originals and the trend observed as we move across different specifications is retained. This observation is to be expected given that the highest earner and the household head are usually the same person. Around 67% of highest earners are also household heads in our sample.

[17] We do not perform a similar robustness test for the Townsend panel data as income data in the Townsend survey are given at the household level and it is impossible to identify a particular member of the household who would be the maximum earner.

Table 3.A.4: LFS Pseudo-Panel: Highest Earners

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Lag Earnings	**0.885*****	**0.546*****	**0.555*****	**0.770*****	**0.865*****	**0.217*****	**0.285*****	**0.545*****	**0.013**
	(0.019)	(0.040)	(0.038)	(0.033)	(0.021)	(0.049)	(0.041)	(0.048)	(0.045)
Education		80.215							
		(51.086)							
Household Size		-364.512***							
		(126.760)							
Urban Dummy		3119.459***							
		(470.389)							
Agriculture		-92.577							
		(722.512)							
Manufacturing		-3798.400***							
		(422.390)							
Age			yes			Yes	Yes		Yes
Period		Yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	757.991***	753.149	1840.871	599.115***	730.222***	1204.687	2729.652	1117.291***	0.000
	(100.726)	(953.101)	(1821.464)	(214.684)	(275.239)	(1618.434)	(1710.031)	(319.194)	(0.869)
Observations	650	650	650	650	650	650	650	650	650
R^2	0.770	0.847	0.843	0.794	0.774	0.883	0.869	0.817	

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.A.1.4 Two-Year Birth Cohort

The results obtained so far for the LFS pseudo-panel data are all based on a pseudo-panel that is constructed by tracking 25 single year of birth cohorts (1941–1965) over a 27-year period (1985–2011). In this subsection, we present the estimates of the Galton-Markov earnings regression for a pseudo-panel data set that is constructed from 13 birth cohorts spanning two years each (1941–1966) for the same 27-year period, yielding 351 cohort-year observations in total.[18] This robustness test is performed to check the sensitivity of the persistence estimates to cell sizes. With the pseudo-panel of two-year birth cohorts, cell sizes mostly exceed 100 observations per cell (in 348 out of 351 cells), and are greater than 50 in all cohort-year cells.

Table 3.A.5 shows the result of the estimation of the Galton-Markov earnings model for the pseudo-panel data using two-year birth cohorts. The persistence parameter estimates obtained from this robustness test is very similar to the original estimates obtained from the pseudo-panel with single-year birth cohorts. This confirms that the bias from having small cell sizes in some cohort-year observations is negligible.

[18] The years of birth used in this estimation also include the 1966 cohort in order to make the year of birth cohorts an even number for the construction of the pseudo-panel with birth cohorts spanning two years each.

Table 3.A.5: LFS Pseudo-Panel: Two-Year Birth Cohort

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Lag Earnings	**0.748*****	**0.143***	**0.688*****	**0.459*****	**0.714*****	**0.113***	**0.400*****	**0.323*****	**0.053**
	(0.081)	(0.076)	(0.088)	(0.096)	(0.082)	(0.072)	(0.094)	(0.084)	(0.061)
Education		-148.797***							
		(53.173)							
Household Size		-186.956*							
		(101.773)							
Urban Dummy		2953.666***							
		(576.786)							
Agriculture		-2465.624***							
		(286.515)							
Manufacturing		-2325.944***							
		(502.140)							
Age			yes			Yes	Yes		Yes
Period		Yes		Yes		Yes		Yes	Yes
Cohort					Yes		Yes	Yes	Yes
Constant	1201.899***	2705.719***	452.866	1448.544***	1125.928***	2751.947***	1187.803	1424.526***	0.000
	(151.250)	(903.167)	(1676.489)	(243.575)	(176.963)	(596.899)	(1588.855)	(245.053)	(0.522)
Observations	338	338	338	338	338	338	338	338	338
R^2	0.600	0.795	0.669	0.696	0.615	0.816	0.728	0.732	

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 4

Power Law in the Income Distribution: Evidence from Thailand

4.1 Introduction

One the most famous distributions in statistics is the bell-shaped normal distribution. This distribution is referred to as "normal" because many empirical quantities measured usually cluster around a typical value, leaving only small probabilities of observing values far from this central value. One example of a quantity that exhibits this characteristic is the height of the adult population. Most adults are around 175cm tall, and while there are some variations around this value, it is extremely rare to observe large deviations that are more than a factor of three from the mean in either direction. Therefore, the height distribution can be summarised by stating just its mean and standard deviation, both of which are finite. The underlying processes that generate this type of distribution are described well by the central limit theorem (CLT).

However, not all distributions fit this pattern. In some cases the departure from normality is not a defect, but an indication of the possibility that the system may be generated by interesting and complex mechanisms. In particular, the past decades have produced multiple examples of "non-normal" distributions from complex social, biological and technological systems (Newman, 2005). One of the most studied distributions of this type is the power-law distribution.

A quantity x that follows a power-law distribution has a probability distribution $p(x) \propto$

$x^{-\alpha}$, where $\alpha > 0$ is the scaling parameter known as the Pareto or power-law exponent. The power-law distribution has many special mathematical properties. Unlike the normal distribution, power-law distribution, under certain parameter values, has infinite mean and variance due to its heavy-tailed characteristic that is a result from the slow decay at the extreme as $x \to \infty$. Power-law distribution can be produced by interesting endogenous processes like feedback loops, self-organisation, network effects, etc. (Newman, 2005). Therefore, the emergence of power-law in socio-economic data provides an initial signal that the system may have a complex underlying generative mechanism that could challenge the prevailing economic thinking.

For example, Kuznets (1955) provided a mechanism that linked the evolution of the level of income inequality to economic development. Kuznet asserted that inequality tends to rise initially at the early stage of development as a result of rapid urbanisation that caused more migration from the agrarian society to the less equal urban industrial society. Inequality then falls later as the economy develops further. Rising wages from urbanisation and technological advancement that create new industries raise the income level of the lower class at a faster rate than for the upper class, which is less willing to adopt new changes. This gives rise to Kuznet's inverted U curve relationship between inequality and economic growth. However, Kuznet's dynamic is inconsistent with the emergence of power law in the income distribution. In a power-law setting, where one of the possible generating mechanisms is preferential attachment (Yule, 1924), there may never be an equalising mechanism. Under preferential attachment, where there is a mechanism that distributes more wealth to individuals who are already better off, the rich just get richer and stark inequality remains even after the economy has gone through the later stages of development. This seems to be the case in developed economies such as the US, where recent findings show that the share of income going to the top 1% and 0.1% of the population has risen sharply since 1980, leading to a substantial increase in top income inequality (Atkinson et al., 2011). If there are other mechanisms at play, a better understanding of this will assist policymakers to better manage the problem of rising global inequality.

The first findings of power law in economics came from the work of Vilfredo Pareto (1897), who studied income data for England, a number of Italian cities, several German states, Paris and Peru. Plotting the cumulative distributions of income for these countries

on double logarithmic paper, Pareto claimed that in each case the result was a straight line with approximately the same slope (Persky, 1992). Pareto observed that the values of α clustered around 1.5. While Pareto used no quantitative measures of goodness of fit, visual inspection suggested that these linear equations worked quite well. The constancy of α and the quality of the fit encouraged Pareto to assert a law of income distribution. From the start, Pareto emphasised the heavily asymmetric character of his distribution and hence its fundamental difference from a normal curve.

Since Pareto's seminal work, there has been great interest in both theoretical and empirical works concerning power law. The so-called econo-physics literature is engaged with both finding further empirical evidence of power law in both economic and non-economic data, and searching for possible mechanisms that might explain the emergence of power-law in these data.

In a quest to find universality in the income and wealth distribution in the spirit of Pareto's original work, further studies of the income distribution find evidence of power law in various countries including Australia (Clementi et al., 2006), Italy (Clementi and Gallegati, 2005), India (Jayadev, 2008), the UK (Dragulescu and Yakovenko, 2001), Japan (Nirei and Souma, 2006) and the US (Levy and Solomon, 1997; Klass et al., 2007). In most cases, the findings indicate that the majority of the wealth and income distribution is either lognormal or exponential with a power-law tail for those with high income/wealth. However, the power-law exponents are not the same across countries and time, as originally asserted by Pareto. In the US, Klass et al. (2007) updated Levy and Solomon's earlier work using the rank-wealth regression on Forbes 400 rich list data. In addition to finding the power law, they look at the development of the power-law exponents over the years and find a negative correlation between the exponent and the average wealth. This means that inequality is highest when average wealth is highest, contrasting evidence to Kuznet's dynamic described earlier.

Empirical evidence of the power-law distribution is also found in other areas such as the studies of city and firm size. Power-law exponents close to 1 ($\alpha = 1$), also referred to as Zipf's law, are usually found in these studies. As early as Auerbach (1913), it was proposed that the city size distribution could be closely approximated by a power-law distribution. Abundant empirical evidence that followed shows the resilience of Zipf's law. Gabaix (1999,

2009), performing a rank and city size regression, finds that the size distribution of US cities follows power law with an exponent of 1. Power law is also found in the Netherlands (Brakman et al., 1999), Japan (Davis and Weinstein, 2002), Germany (Bosker et al., 2008), and the UK (Rozenfeld et al., 2011).

Zipf's law has also been shown to hold for firm size, measured by the number of employees, assets, or market capitalisation in the US (Axtell, 2001; Gabaix and Landier, 2008), Europe (Fujiwaraa et al., 2004), and Japan (Okuyama et al., 1999). Axtell (2001) uses data on all firms in the US census to show that Zipf's law describes firm size measured by the number of employees (regression of log frequency and firm sizes). More examples of all empirical findings of power law in various fields can be found in Pinto et al. (2012) and Gabaix (2009).

Several mechanisms have been proposed to theoretically explain the existence of power law in these settings. One of the most popular explanations is proportional random growth, which originated with Yule (1924), and was developed in economics by Champernowne (1953) and Simon (1955) and rigorously studied by Kesten (1973). The key feature of this process is the proportional random growth in the spirit of Gibrat's law and the addition of some friction that ensures that a steady state exists (Gabaix, 2009). Jones and Kim (2014), in their study of top income inequality, propose exponential growth with exponential arrival as a process that might generate power law. In their context, heterogeneous entrepreneurs exert effort to generate exponential growth in their income, while creative destruction represents the friction that causes the steady-state distribution to obey a Pareto distribution.

Other mechanisms include the rich getting richer or preferential attachment (Yule, 1924; Simon, 1955), and physics-based self-organised criticality (Bak et al., 1987). While several of these mechanisms can explain the emergence of power law, none explain why the power-law exponent should be in the range of 1.5 – 2 as observed in the literature. This remains an area of active research.

Despite the abundance of empirical evidence of power law, especially in the income and wealth distribution in many countries, the method employed in most of these studies is unreliable. The widespread method used to identify power law is based on performing an OLS regression of log rank and income and a highly subjective visual inspection of the linear fit. Employing this method can lead to a false conclusion on the existence of power law in the data. A recent study by Gan et al. (2006) using Monte Carlo simulations argued

that the observed Zipf's law for city size distribution is a spurious statistical phenomenon. OLS regression of the rank-size relationship from data randomly generated from various probability distributions shows a strong linear fit with high R^2, despite the lack of any underlying economic theory to justify the result.

Clauset et al. (2009) also argue that the least squares-based method for estimating the power-law exponent generates significant systematic error and cannot be trusted. The authors propose that the power-law exponent be estimated using a maximum likelihood method equivalent to the Hill estimator. Evidence from simulations shows that the maximum-likelihood estimators were able to more accurately estimate the true α than the OLS estimator based on rank-frequency or histogram regression. The authors also propose a way to identify x_{min} (the minimum cut-off value above which the power-law distribution applies), and perform a hypothesis test of the validity of power law, both on an absolute basis and relative to other competing distributions such as the lognormal. Under certain conditions, this could also produce a straight line plot on the log scale that is hard to differentiate from that of the power law (Mitzenmacher, 2003).

Researchers who have employed alternative methods mentioned above find that the evidence of power law is not as strong as previously claimed. Other "heavy-tailed" distributions, especially the lognormal, fit the wealth distribution equally well (Brzezinski, 2014; Ogwang, 2013). Given the similarity in the generation mechanism, the similarity of the two distributions is not surprising. This make it all the more important to employ a rigorous method when trying to fit a power-law distribution to data.

This chapter contributes to the existing literature by further promoting the correct fitting of the power-law distribution, providing evidence based on the income and consumption distribution of Thailand. The second objective is to use the estimated power-law exponent to study and provide additional evidence of the dynamics of inequality, especially at the upper end of the income and consumption distribution, in Thailand.

Thailand is a good candidate for this type of study as it is an example of an emerging economy that has been experiencing rapid economic growth. In this type of economy, one expects the rich getting richer effect to be particularly strong; hence the wealth and income distribution may exhibit power-law behaviour where concentration at the top end is at a higher level than predicted by the normal or lognormal framework. Discovery of power law

may also provide an indication that some other underlying complex mechanism may be at work in Thailand. However, finding the underlying mechanism that explains the emergence of power law is beyond the scope of this chapter.

This chapter is the first study of the power-law fit to the income and consumption distribution in Thailand. We find that using the popular but incorrect method based on the linear regression approach will lead to researchers drawing a wrong conclusion. OLS estimates of the power-law exponent, α, provide strong evidence of power-law fit in Thailand, with R^2 exceeding 95% for data in all the years. However, with implementation of the Clauset et al. method, the evidence in support of the power-law fit is much weaker. Only slightly more than half of the data pass the more stringent power-law goodness-of-fit test on an absolute basis and none passes the test when the power-law fit is compared to the competing lognormal fit. This illustrates the superiority of Clauset et al.'s approach, which should be the default starting point for similar studies of this type going forward.

Our estimates of the power-law exponents for both income and consumption are greater than 3, larger than what is usually observed in the literature ($2 < \alpha < 3$), suggesting that there is less inequality at the top end of the income and consumption distribution in Thailand. Upon further inspection we find that while there is less inequality at the top in Thailand, there is a high level of persistent between-group inequality between the top and bottom ends of the distribution. This helps explain and reconcile the fact that we observe high α (high α signals lower inequality) but persistently high Gini coefficients in Thailand.

In addition, we also follow the work of Battistin et al. (2009) to further investigate the underlying income-generating process in Thailand. According to the authors, if Gibrat's law of proportional random growth holds, then we should observe permanent income following the lognormal distribution closely. However, we find that reported income is closer to the lognormal distribution than consumption, a proxy for permanent income, in most years. One possible explanation could be that the volatility of income in Thailand is low and Gibrat's law applies to it instead. Consumption then deviates further from lognormal than income as a result of consumption smoothing. Nevertheless, the pattern of the data does not follow Gibrat's predictions in most instances, so these findings do cast serious doubt on the validity of Gibrat's law of proportional random growth as the income-generating process in Thailand.

The structure of this chapter is as follows. Section 4.2 presents a discussion of the

similarities and differences between the two competing distributions in the distribution-fitting literature, the power law and the lognormal. Section 4.3 discusses the methodology used in this chapter to fit and test the appropriateness of the power-law distribution to the data. Section 4.4 describes the survey data and presents the descriptive statistics of income and consumption data from the Thailand SES. Section 4.5 presents the empirical results and discussion. Section 4.6 offers concluding remarks.

4.2 Power-law and Lognormal Distribution

The question of whether income distribution follows a lognormal or power-law distribution dates back to at least the 1950s (Aitchison and Brown, 1957). The literature has arrived at a consensus that the lognormal provides a better fit to the body of the income distribution, while the tail of the income distribution is better fitted by the Pareto distribution. Several previous works (Gabaix, 2009) have shown that power law and lognormal distributions are intrinsically connected. Very similar basic generative models can lead to either power-law or lognormal distributions, depending on seemingly trivial variations. This close relationship between the two distributions is a reason that arguments as to whether power-law or lognormal distributions provide a more accurate fit have arisen and repeated themselves across a variety of fields (Mitzenmacher, 2003).

This section, following the work of Gabaix (2009) and Battistin et al. (2009), presents one of the basic generative models, the proportional random growth model, that could lead to either a lognormal or a power-law distribution. In this framework we show how small variations in the underlying model can change the result from one distribution to the other. This section will also present the mathematical description and properties of both the power-law and lognormal distributions.

4.2.1 Proportional Random Growth Model

The proportional random growth model is often cited as one of the possible underlying mechanisms that can generate both the lognormal and power-law distributions. The model is closely linked to what is often referred to as Gibrat's law of proportional growth (Gibrat, 1931).

Gibrat's law applies to income claims that the random growth of income is expressed as a percentage of its current level, and is independent of its current size. Given this process, Gibrat's law predicts that the distribution of income should tend to a lognormal distribution, but one that becomes more and more degenerate as time increases (Malevergne et al., 2011).

The equation of income growth following the proportional random growth model and embodying Gibrat's law is

$$Y_{i,t} = g_{i,t} Y_{i,t-1} \tag{4.1}$$

where $Y_{i,t}$ is the income of household i at time t, and $g_{i,t}$ is the random positive growth factor. Taking the logarithm of Equation (4.1) and iterating yields

$$lnY_{i,t} \;=\; lnY_{i,t-1} + \eta_{i,t} \;=\; lnY_{i,0} + \eta_{i,1} + \eta_{i,2} + ... + \eta_{i,t} \tag{4.2}$$

where $\eta_{i,t} = ln(g_{i,t})$. Assuming that the terms $\eta_{i,t}$ are iid random variables with expectation μ and standard deviation σ, the CLT gives

$$\begin{gathered} \frac{lnY_{i,t}}{t} \simeq \frac{\sum_{t=1}^{T} \eta_{i,t}}{t} \sim N(\mu, \frac{\sigma^2}{t}) \\ lnY_{i,t} \simeq t\mu + \sqrt{t}\sigma\xi \end{gathered} \tag{4.3}$$

where ξ is a standard Gaussian random variable $N(0,1)$. Assuming that the strong form of ergodicity holds – that is, that the random growth process of income for a typical household over time is equivalent to sampling the income growth for a cross-section of households – Equation (4.3) ensures that the distribution of income is lognormal $ln(Y_{i,t}) \simeq N(t\mu, t\sigma^2)$, for large t.

Equation (4.2) shows that the growth of income follows a random walk in its log size. This generates a distribution of income $Y_{i,t}$ that approaches a lognormal distribution with an ever-increasing mean and variance as t increases. The model therefore has no real steady state as the parameters of the income distribution keep changing.[1] In the limit, where the variance of the lognormal distribution is infinite, income is undetermined as the probability that income is in a certain interval approaches zero for any income interval.

As demonstrated in Gabaix (2009), a minor modification leads to a steady state and therefore a stationary distribution of income. This modification, which can take many forms (Sornette and Cont, 1997; De Wit, 2005), essentially involve preventing the small income from becoming too small. The corresponding equation of motion of income representing this idea is

$$Y_{i,t} = g_{i,t}Y_{i,t-1} + \epsilon_{i,t} \tag{4.4}$$

where the term $\epsilon_{i,t} > 0$ prevents the accumulation of large numbers of incomes with extremely small values. Without $\epsilon_{i,t}$, Equation (4.4) is the same as the random walk process in log income, leading to the lognormal distribution in Equation (4.1).

[1] For more detail see Clauset et al. (2009)

The presence of $\epsilon_{i,t} > 0$ even if it is arbitrarily small, provided that it is non-zero, together with the conditions that $E\left[lng_{i,t}\right] < 0$ and that there are values of t with $g_{i,t} > 1$ causes the income distribution to switch from a lognormal to power law. This is called a Kesten process (Kesten, 1973). The scaling parameter α of the power-law distribution is the solution to $E\left[(g_{i,t})^{\alpha}\right] = 1$. Gabaix (1999) argued for the validity of the constraint $E\left[(g_{i,t})\right] = 1$ in the steady state, which leads to Zipf's law where $\alpha = 1$.

The intuition behind the Kesten process is that, because of the condition $E\left[lng_{i,t}\right] < 0$, in the absence of $\epsilon_{i,t}$, the income process $Y_{i,t}$ tends to shrink stochastically towards zero becoming an increasingly degenerate lognormal distribution. The inclusion of the additive term $\epsilon_{i,t}$ makes sure that the process does not collapse to the origin. During this phase, the condition that occasionally $g_{i,t} > 1$ can occur with exponentially small probability creates intermittent bursts that form the fat tail. The term $\epsilon_{i,t}$ allows the process to repeatedly exhibit the exponentially rare and exponentially large deviation. The combination of these two exponentials leads to the power-law distribution. The tail exponent of the Kesten process depends on the relative sizes of the additive and multiplicative terms.

4.2.1.1 Permanent Income Model

In the income-generation model proposed by Battistin et al. (2009), the authors make a distinction between permanent and transitory components. The authors argue that it is the permanent component of income that follows the multiplicative process described above and hence Gibrat's law should only apply to the permanent component and not the total income.

In their model, an individual who has been earning an income for τ years has log income $y_{\tau} = lnY_{i,\tau}$ and log permanent income $y_{\tau}^{p} = lnY_{i,\tau}^{p}$ of the following form

$$y_{\tau} \;=\; y_{\tau}^{p} + \nu_{\tau} \tag{4.5}$$

where ν_{τ} is the transitory shock in log income. Permanent income follows Gibrat's law of proportional growth and evolves as

$$y_{\tau}^{p} \;=\; y_{\tau-1}^{p} + \eta_{\tau} \tag{4.6}$$

where η_{τ} is the shock to permanent income. This equation is essentially the same as

Equation (4.2) so η_τ can be interpreted as the random growth rate of permanent income.

Iterating Equation (4.6), similar to before, yields

$$y_\tau^p = y_0^p + \eta_1 + \eta_2 + ... + \eta_\tau \tag{4.7}$$

where τ is the number of time periods that the person has been earning an income. As τ increases, income in the initial period y_0^p becomes relatively small compare to the accumulation of random shocks. Assuming as before that the shock η_t is an iid random variable, then by application of the CLT, assuming standard regularity conditions, there exists moments μ_p and σ_p^2 such that $y_\tau^p \approx N(\tau\mu_p, \tau\sigma_p^2)$ for large τ.

Therefore, the income-generation model above suggests that permanent income scaled by age τ should be approximately lognormally distributed, at least for older individuals who are of sufficient age to have experienced a reasonable number of permanent income shocks. In addition, if the strong form of ergodicity holds, then the above model implies that the distribution of permanent income across individuals within the same age cohort should also be approximately lognormal. The CLT also implies that the dispersion of income within cohorts increases as the cohorts aged as $E(y_\tau^2) = \left[E(y_\tau^p + \nu_\tau)^2\right] = \tau^2\mu_p^2 + \tau\sigma_p^2 + E(\nu_\tau^2)$, which increases with τ. This is consistent with the result from a study by Deaton and Paxson (1994).

Using the model discussed above, Battistin et al. (2009) has shown that it is the permanent component of income that is determined by the accumulation of a series of proportional income shocks, not total income as originally assumed by Gibrat. Therefore, Gibrat's law of proportional growth should hold for permanent income, but not necessarily for total income.

In the case where the transitory shocks ν_τ in Equation (4.5) are small relative to y_τ^p, then log total income y_τ will also be close to normal. However, unless transitory shocks are also normally distributed, the distribution of log permanent income will be closer to the normal distribution than that of total income. When transitory shocks have an appropriately skewed distribution, possibly through temporary unemployment or random large wealth shocks, the total income distribution can take the lognormal form with a Pareto upper tail commonly observed in the literature.

In this section we have discussed one possible mechanisms for the generation of both the power-law and lognormal distributions. As shown, the two distributions are very closely

related to each other, partly due to the similarity in their generative process. In the next section, we present the descriptions and properties of both distributions.

4.2.2 Power-law and Lognormal Distribution

4.2.2.1 Power-Law Distribution: Definitions and Properties

Mathematically, a quantity x follows a power law if it is drawn from a probability distribution

$$p(x) \propto x^{-\alpha} \tag{4.8}$$

where the exponent α is a constant scaling parameter of the distribution. The scaling parameter normally lies in the range $2 < \alpha < 3$, although there are occasional exceptions. In practice, it is difficult to find empirical occurrences that obey power laws for all values of x. It is usually the case that the power law behaviour applies only for values greater than some minimum x_{min}. In such cases only the tail of the distribution follows a power law.

There are two basic forms of power-law distribution: continuous and discrete power-law distributions. In this chapter, as we are working with income and consumption, we will only be focussing on the continuous case.[2]

Let X represent household income or consumption, the quantity whose distribution we are interested in modelling. The probability density $p_X(x)$ of a continuous power-law distribution is

$$p_X(x)dx = Pr(x \leq X < x + dx) = Cx^{-\alpha}dx \tag{4.9}$$

where C is a normalisation constant. It is clear from Equation (4.9) that the probability density diverges as $x \to 0$, so this relationship cannot hold for all values of $x \geq 0$. For this reason, there must exist a lower bound to the power-law behaviour, which we denote as x_{min}. Provided $\alpha > 1$, substituting in the normalising constant gives[3]

$$p_X(x) = \frac{\alpha - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha} \tag{4.10}$$

It is also useful to consider the complementary cumulative distribution function (CDF) of

[2] for the discrete case see Clauset et al. (2009)

[3] the normalising constant is a value of C, which ensures that $\int_{x_{min}}^{\infty} p(x)dx = 1$

a power-law distributed variable in some cases. The complementary CDF ($P_X(x)$) is defined as $P_X(x) = Pr[X \geq x]$. For the continuous power-law case the CDF is

$$P_X(x) = \int_x^\infty p(z)dz = \left(\frac{x}{x_{min}}\right)^{-\alpha+1} \tag{4.11}$$

If X has a power-law distribution, the log-log plot of the probability density function (PDF) or the complementary CDF will yield a straight line. With real data, such straightness is a necessary but not sufficient condition for the data following a power-law relation.

One of the properties of the power-law distribution is that its moment may be infinite. The mean of the power-law distribution is given by

$$E[X] = \int_{x_{min}}^\infty xp(x)dx = C\int_{x_{min}}^\infty x^{-\alpha+1}dx = \frac{C}{2-\alpha}\left[x^{-\alpha+2}\right]_{x_{min}}^\infty \tag{4.12}$$

This expression becomes infinite if $\alpha \leq 2$. For $\alpha > 2$, the mean is well defined and is equal to $\frac{\alpha-1}{\alpha-2}x_{min}$. The second moment is given by

$$E\left[X^2\right] = \frac{C}{3-\alpha}\left[x^{-\alpha+3}\right]_{x_{min}}^\infty \tag{4.13}$$

This diverges and becomes infinite if $\alpha \leq 3$. Thus power-law distributions in this range have no meaningful variance or standard deviation. The implication of this is that there is a very large fluctuation in the value of the standard deviation and there are higher chances of observing very high values of x in the tail of the distribution than is usually the case with a distribution that has finite variance.

It can be shown that in general for all moments $E[X^m]$ exist for $m < \alpha - 1$ and all higher moments diverge. As stated earlier, most of the observed power law has a scaling parameter that lies in the range $2 < \alpha < 3$. This suggests that these commonly observed power laws have a well-defined mean but an infinite variance.

Another property of the power law is that it is a top-heavy distribution. For any power law with exponent $\alpha > 1$, the median is well defined. There exist a point $x_{1/2}$ that divides the distribution in half. This point is given by

$$\int_{x_{1/2}}^\infty p(x)dx = \frac{1}{2}\int_{x_{min}}^\infty p(x)dx \tag{4.14}$$

So $x_{1/2} = 2^{1/(\alpha-1)} x_{min}$. The median divides the income distribution into the richer half and the poorer half, where each half contain the same number of individuals. The fraction of income possessed by the richer half is given by

$$\frac{\int_{x_{1/2}}^{\infty} xp(x)dx}{\int_{x_{min}}^{\infty} xp(x)dx} = \left(\frac{x_{1/2}}{x_{min}}\right)^{-\alpha+2} = 2^{-(\alpha-2)/(\alpha-1)} \tag{4.15}$$

The integrals converge as long as $\alpha > 2$. If we find that $\alpha = 2.15$ for the income distribution, this will imply that a fraction $2^{-0.15/1.15} \simeq 91\%$ of income is in the hands of the richer half (top 50%) of the distribution, making the distribution relatively top-heavy. More generally, the fraction W of total income in the hands of the fraction P of the population whose income exceeds x is

$$W = \frac{\int_{x}^{\infty} xp(x)dx}{\int_{x_{min}}^{\infty} xp(x)dx} = \left(\frac{x}{x_{min}}\right)^{-\alpha+2} = P^{(\alpha-2)/(\alpha-1)} \tag{4.16}$$

where $P = \int_{x}^{\infty} p(x)dx = \left(\frac{x}{x_{min}}\right)^{-\alpha+1}$. Assuming that $\alpha = 2.15$ as before, we can calculate that the richest 20% of the population hold roughly 80% of total income. This is basically what is commonly referred to as the Pareto rule or the 80/20 rule. The lower the scaling parameter α, the fatter the tail of the power-law distribution and the greater the inequality in income.

A related property is that the power-law distribution is regularly varying. A positive function is regularly varying in the limit if there exist a finite real number α such that

$$\lim_{x\to\infty} \frac{f(t \cdot x)}{f(x)} = t^{\alpha}, \ \forall t > 0 \tag{4.17}$$

The power-law or Pareto distribution satisfies this property and therefore, in the limit where $x \to \infty$ the distribution does not decay as fast as the exponential distribution (normal distribution). This property results in the Pareto distribution being a heavy-tailed distribution.

Another interesting property of the power-law distribution is that it is a scale-free distribution. A scale-free distribution is one where

$$p(kx) = g(k)p(x) \tag{4.18}$$

for any k. That is, if we increase the scale or units by which we measure x by a factor of k, the shape of the distribution $p(x)$ is unchanged, except for an overall multiplicative constant. Not only does the power-law distribution satisfy the scale-free condition above, it can be proven that it is the only distribution that exhibits this property. This scale-free property gives rise to an interesting concept of universality. That is, on observing two quantities with the same power-law exponent, it may be possible that a similar underlying mechanism is at work for both.

Lastly, one of the reasons that power laws are ubiquitous is because of their strong invariance under aggregation (Gabaix, 2009). Power-law distribution is preserved under addition, multiplication, and polynomial transformation. When combining two independent power law-distributed variables, the one with the fattest tail (smallest exponent) dominates. When a power law-distributed variable is raised to a non-zero power, it remains a power law but with an altered exponent.

Letting α_X be the tail exponent of the random variable X. For $X_1, ..., X_n$ independent random variables we can write these transformation properties as:

$$\alpha_{X_1+...+X_n} = min\left(\alpha_{X_1}, ..., \alpha_{X_n}\right) \tag{4.19}$$

$$\alpha_{X_1 \mathrm{x} ... \mathrm{x} X_n} = min\left(\alpha_{X_1}, ..., \alpha_{X_n}\right) \tag{4.20}$$

$$\alpha_{bX} = \alpha_X \tag{4.21}$$

$$\alpha_{X^b} = \frac{\alpha_X}{b} \tag{4.22}$$

Therefore if X is a power-law distributed variable with $\alpha_X < \infty$, and Y is a power law-distributed variable with an exponent $\alpha_Y \geq \alpha_X$, then $X + Y$ or $X \cdot Y$ is still power law-distributed with the same exponent α_X. This property holds even when Y is normal, lognormal or exponential. Therefore, multiplying a power law-distributed variable by normal variables or adding non-fat tail noise preserves the power-law exponent.

4.2.2.2 Lognormal Distribution: Definitions and Properties

A random variable X has a lognormal distribution if the random variable $Y = lnX$ follows a Gaussian distribution. The density function of the normal distribution Y is given by

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2} \tag{4.23}$$

where μ and σ are the mean and the standard deviation of Y respectively and the range is $-\infty < y < \infty$.

The density function for a lognormal distribution is given by

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-(ln(x)-\mu)^2/2\sigma^2} \tag{4.24}$$

The change of variables introduces an additional $1/x$ term outside of the exponential term. The corresponding complementary CDF for a lognormal distribution is given by

$$Pr\left[X \geq x\right] = \int_{z=x}^{\infty} \frac{1}{\sigma z\sqrt{2\pi}} e^{-(ln(z)-\mu)^2/2\sigma^2} dz \tag{4.25}$$

A lognormal variable X has parameters μ and σ^2 when both parameters represent the mean and the variance of the corresponding normal distribution Y respectively. The lognormal distribution is positively skewed, and has a finite mean and variance. This is unlike the power law distribution, which has finite mean but infinite variance under natural parameters.

Despite these differences, both distributions have extremely similar shape. A log-log plot of the complementary CDF or the PDF of a lognormal distributed variable will appear to be close to a straight line for a large portion of the body of the distribution.

A logarithm of the density function of the lognormal distribution is given by

$$\begin{aligned} ln\, f(x) &= -ln(x) - ln\sqrt{2\pi}\sigma - \tfrac{(ln(x)-\mu)^2}{2\sigma^2} \\ &= -\tfrac{(ln(x))^2}{2\sigma^2} + \left(\tfrac{\mu}{\sigma^2} - 1\right) ln(x) - ln\sqrt{2\pi}\sigma - \tfrac{\mu^2}{2\sigma^2} \end{aligned} \tag{4.26}$$

If the variance of the associated normal distribution σ is sufficiently large, the quadratic term in Equation (4.26) will be small for a large range of x values, and hence, the logarithm of the density function will appear almost linear for a considerable range of values (Mitzenmacher, 2003).

In addition, the limit behaviour of the lognormal distribution is that of a rapidly decreasing function.

$$\lim_{x\to\infty}\frac{f(t\cdot x)}{f(x)} = \lim_{x\to\infty}\frac{1}{t}e^{-\frac{(ln(t))^2}{2\sigma^2}}e^{-ln(t)\cdot\frac{ln(x)-\mu}{\sigma^2}} = \begin{cases} 0,\ t>1 \\ 1,\ t=1 \\ \infty,\ t<1 \end{cases} \tag{4.27}$$

Therefore, the lognormal density goes to zero, in the upper tail, faster than any Pareto density. Effectively, the Pareto distribution has a heavier tail than the lognormal distribution. Intuitively, this means that the probability of extreme events are more likely to occur in a power law-distributed data than lognormally distributed data. If income follows a power law instead of a lognormal distribution, we would expect to observed individuals with very high income more than we would expected under the lognormal distribution.

In this section, we have discussed the definition, the properties and the possible mechanism that generates both the power-law and lognormal distributions. We have shown that the two distributions are closely related and that a minor change in the generative mechanism can cause a switch between the two distributions. This is why it is inherently difficult to distinguish between the two, and the debate whether lognormality or power law provides a better fit for various distributions is still ongoing. However, given sufficient data, the tail behaviour of the two distributions is different and could be differentiated given the right procedure. In the next section, we present methods that allow the correct fitting of the power-law form to the empirical data and allow for a test to compare the fit with an alternative distribution such as the lognormal. In addition, we provide a test of the hypothesis in Battistin et al. (2009) that the permanent component of income, consumption, follows the lognormal distribution more closely than total income.

4.3 Methodology

In this section we discuss the methodology used in this chapter to identify and fit the power-law form to the empirical distributions. We begin by examining the shortcomings of the widely used method based on linear regression, which could lead to false conclusions regarding the existence of power law in many cases. We then present a method recently proposed by Clauset et al. (2009) that provides a superior alternative in correctly identifying and fitting the power-law form to the Thai income and consumption distribution.

4.3.1 Power-law Distribution

Restating the definition from Section 4.2.2, the PDF of the continuous power-law distribution is

$$p(x) = \frac{\alpha - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha} \tag{4.28}$$

The complementary CDF for continuous power law is

$$P(x) = Pr\left[X \geq x\right] = \int_{x}^{\infty} p(z)dz = \left(\frac{x}{x_{min}} \right)^{-\alpha+1} \tag{4.29}$$

In the next two subsections we will discuss two alternative methods for fitting a power law to the Thai empirical income and consumption distribution.

4.3.2 Fitting Power Law: Linear Regression Approach

Fitting a power law to the empirical distribution requires estimation of the scaling parameter α and occasionally also of the lower-bound x_{min}. The tool most often used for this task is the method based on least-squares linear regression.

In this approach, the probability density $p(x)$ is first estimated by constructing a histogram of the data. Taking a logarithm of both sides of Equation (4.28) gives

$$lnp(x) = -\alpha ln(x) + constant \tag{4.30}$$

This produces a straight-line relationship on a doubly logarithmic plot. Equation (4.30) can then be fitted to the linear form by least-squares linear regression. The slope of the fit

is the estimate $\hat{\alpha}$ of the scaling parameter.

Alternatively the complementary CDF, $P(x)$, constructed through a simple rank ordering of the data, can be use instead. This has the added benefit over working with the PDF as there is no need to make an arbitrary choice regarding the size of the bin used in the construction of the histogram. By definition, $P(x)$ is well defined for all values of x and can be plotted as a normal function without binning. The CDF also makes much better use of the data and does not throw away any information in contrast to binning, which often lump data within a given range together. It is important to note that while the CDF also follows the power law, the scaling parameter is $\alpha - 1$. A linear fit to the CDF plot, therefore, will yield a slope parameter that is 1 less than the original scaling parameter α.

Least-squares linear fitting also provides the standard errors for the estimated slope, and calculates the R^2, the fraction of the variance accounted for by the fitted line. The R^2 is often taken as an indicator of the quality of the fit.

Although this method appears frequently in the literature, there are several problems with it. Clauset et al. (2009) show that the estimates of the slope are subject to systematic and potentially large errors. There are other serious problems as well. First, standard errors produced by the least-squares fit are based on the assumption of independent and normally distributed error terms. When fitting to the logarithm of a complementary CDF as in the analysis of power-law data, however, the noise is not normally distributed. The noise of the individual values $P(x)$ is Gaussian as it is the sum of independent Gaussian variables, but the noise in the logarithm does not thus violate the key assumption. Furthermore, the assumption of independence does not hold as $P(x) = P(x+1) + p(x)$ result in the adjacent values of the CDF being strongly correlated. This causes an underestimation of the standard error of the estimated scaling parameter due to a failure to account for the correlations.

Second, a high R^2 value, often used as to validate the goodness-of-fit, cannot be taken as evidence for the power-law form. While a low value of R^2 can be used to reject the power-law hypothesis, it cannot be used to validate the power-law fit. Several other distributions can appear to follow a power law for small samples and some, like the lognormal, can approximate a power law very closely over a considerable range. A linear least-squares fit to these will yield a high value of R^2 also. Thus, the value of R^2 has very little power as a test of the power-law hypothesis, as the probability of successfully detecting a violation from the

power-law assumption based on the observed value of R^2 is low.

Finally, to satisfy the basic requirement of probability distributions, the complementary CDF must take the value 1 at x_{min} and similarly, the PDF must integrate to 1 from x_{min} to ∞ if the probability distribution above x_{min} is properly normalised. Ordinary linear regression, however, does not take account of such constraints. Therefore, the fits obtained from the regression method usually do not satisfy these basic probabilistic requirements. While standard methods exist to include these constraints into the regression analysis, they are not commonly used in the literature on power laws (Clauset et al., 2009).

It is due to these shortcomings of the linear least-squares approach that we elect to follow an alternative method proposed by Clauset et al. (2009) in our analysis of power-law behaviour in the Thai income and consumption distribution.

4.3.3 Fitting Power Law: An Alternative Approach

In this section we present the method of fitting power law to the empirical data using the method proposed by Clauset et al. (2009).

4.3.3.1 Estimating the Scaling Parameter

The correct estimation of the scaling parameter α requires a value for the lower bound x_{min} of power-law behaviour in the data. We first assume that x_{min} is known. In the next section, we will consider a method for estimating x_{min} from the data.

The method proposed for fitting the power-law distributions to observed data is the method of maximum likelihood, which has been proven to give accurate parameter estimates in the limit with a large sample size (Greene, 2011). Assuming that the data follow a power-law distribution for $x > x_{min}$, the PDF (as stated in Equation 4.28) is

$$p(x) = \frac{\alpha - 1}{x_{min}} \left(\frac{x}{x_{min}} \right)^{-\alpha} \tag{4.31}$$

where α is the scaling parameter and x_{min} is the minimum value at which power-law behaviour holds as before. Intuitively, the maximum-likelihood approach seeks to identify the value of α for the power-law model that is most likely to have generated the observed data. Given a data set containing n observations $x_i \geq x_{min}$, the likelihood function is

$$p(x \mid \alpha) = \prod_{i=1}^{n} \frac{\alpha - 1}{x_{min}} \left(\frac{x_i}{x_{min}} \right)^{-\alpha} \tag{4.32}$$

The maximum-likelihood estimate (MLE) of the scaling parameter α is the value of $\hat{\alpha}$ that maximised this function. We usually work with the logarithm $\mathcal{L}$, a monotonic transformation of the likelihood, which has its maximum in the same place:

$$\begin{aligned} \mathcal{L} &= ln \left[\textstyle\prod_{i=1}^{n} \frac{\alpha-1}{x_{min}} (\frac{x}{x_{min}})^{-\alpha} \right] \\ &= \textstyle\sum_{i=1}^{n} \left[ln(\alpha - 1) - ln(x_{min}) - \alpha ln(\frac{x_i}{x_{min}}) \right] \\ &= n\, ln(\alpha - 1) - n\, ln(x_{min}) - \alpha \textstyle\sum_{i=1}^{n} ln(\frac{x_i}{x_{min}}) \end{aligned} \tag{4.33}$$

Setting $\partial\mathcal{L}/\partial\alpha = 0$ and solving for α, we obtain the MLE for the scaling parameter of the continuous power law:

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^{n} ln \frac{x_i}{x_{min}} \right]^{-1} \tag{4.34}$$

Equation (4.34) is equivalent to the Hill estimator (Hill, 1975), which is asymptotically normal and consistent. The variance of $\hat{\alpha}$, is equal to the inverse of Fisher's information $I(\alpha)$, the second derivative of the log likelihood function (Greene, 2011), where

$$I(\alpha) = -E \left[\frac{\delta^2 \mathcal{L}}{\delta \alpha^2} \right] \tag{4.35}$$

The standard error of the MLE $\hat{\alpha}$ is

$$\sigma = \sqrt{I(\hat{\alpha})^{-1}} = \frac{\hat{\alpha} - 1}{\sqrt{n}} + O(1/n) \tag{4.36}$$

where the higher-order correction is positive.

In these calculations we assumed that $\alpha > 1$ as distributions with $\alpha \leq 1$ are not normalisable and does not usually exist in practice.[4] The MLEs are unbiased only in the asymptotic limit of large sample size. For finite datasets, biases are present but decay as $O(1/n)$ for any choice of x_{min}. The biases can be significant for small sample size, but they can mostly be ignored in practice as they are much smaller than the statistical error of the estimator,

[4]It is possible for a probability distribution $x^{-\alpha}$ to have $\alpha \leq 1$ if the range of x is bounded above by some cut-off, but different maximum-likelihood estimators are needed to fit such a distribution (Clauset et al., 2009).

which decays as $O(1/\sqrt{n})$.[5] A reasonable rule of thumb for the size of data needed to obtain reliable parameter estimates is suggested to be approximately $n \geq 50$.

4.3.3.2 Estimating the Lower Bound of Power Law

As stated, it is typically the case that empirical data only follow a power-law distribution for values of x above some lower bound x_{min}. Therefore, before calculating the estimate of the scaling parameter α, it is necessary to discard all samples below x_{min} so that only those for which the power-law model is a valid model remain. In the last section we take the value x_{min} as given. In this section we present a method for estimating this lower-bound parameter from the data.

To obtain an accurate estimate of α we will also need an accurate method for estimating x_{min}. Selecting a value for x_{min} that is too low will result in a biased estimate of α as we will be attempting to fit a power-law model to a section of the data that does not follow a power law. On the other hand, selecting a value for x_{min} that is too high will result in the loss of legitimate data points, which increases both the statistical error of α and the bias from finite size effects.

The most common methods of choosing $\hat{x}_{min}$ are either via visual inspection of the cut-off point above which the PDF or CDF of the distribution looks approximately straight on a log-log plot, or by plotting $\hat{\alpha}$ as a function of the corresponding $\hat{x}_{min}$ and identify a point beyond which the $\hat{\alpha}$ appears to be stable. These approaches, however, are very subjective and can be sensitive to fluctuation in the tail of the distribution.

Clauset et al. (2009) proposed a more objective method for estimating x_{min} that can be applied to both discrete and continuous data. The fundamental idea behind this method is based on selecting $\hat{x}_{min}$ that makes the probability distribution of the observed data and the best-fit power-law model as similar as possible above $\hat{x}_{min}$. If the $\hat{x}_{min}$ selected is higher than the true value x_{min}, then valuable information is being discarded, which will make the probability distributions a poorer fit because of statistical fluctuation. Conversely, if we select $\hat{x}_{min}$ that is smaller than the true x_{min}, the distribution will differ because of the fundamental difference between the data and model. In between lies the best estimate.

The measures used by the authors for quantifying the distance between two probability distributions in this method is the Kolmogorov-Smirnov or KS statistic, which is the

[5] For more detail see Clauset et al. (2009)

maximum distance between the CDFs of the data and the fitted model:

$$D = \max_{x \geq x_{min}} |S(x) - P(x)| \tag{4.37}$$

where $S(x)$ is the empirical CDF of the observed data with value greater than and equal to x_{min}, and $P(x)$ is the CDF for the power-law model that best fits the data in the region $x \geq x_{min}$. The estimate $\hat{x}_{min}$ is the value of x_{min} that minimises D.

This method is particularly sensitive to slight deviations of the data from the power-law model around x_{min} because, for this kind of heavy-tailed distribution, this region is where most of the data lie. This sensitivity allows the method to be able to select a reasonably accurate value of x_{min}. Simulation results by the authors give excellent results in support of the method and generally outperform alternative approaches.

The accuracy of $\hat{x}_{min}$ is sensitive to the number of observations in the power-law part of the distribution n_{tail}. The number of observation n_{tail} does depend on the particular form of the non-power-law part of the distribution. If this part of the distribution shows a pronounced departure from the power law below x_{min} then the task of estimating $\hat{x}_{min}$ would be easier and fewer observations n_{tail} would be needed to obtain an accurate result.

Finally, as with α, it is essential to quantify the uncertainty of the estimate for x_{min}. The authors make use of a nonparametric bootstrap method. Given n measurements, a synthetic data set with a similar distribution to the original is generated by drawing a new sequence of points x_i, $i = 1, ..., n$ uniformly with replacement at random from the original data. Using the method described above, $\hat{x}_{min}$ and $\hat{\alpha}$ are then estimated for this generated dataset. Estimates of the uncertainty in the original estimated parameters are then obtained by taking the standard deviation of these estimates ($\hat{x}_{min}$ and $\hat{\alpha}$) over a large number of repetitions of this process.

4.3.3.3 Testing the Power-law Hypothesis

In previous sections we presented methods for fitting a power-law distribution to observed data and for estimating the parameters α and x_{min}. These methods, however, provide no information about whether the power law is a credible fit to the data. A power law can always be fitted to any dataset irrespective of the true distribution from which the data were drawn. In this section, we present a method that quantifies whether the power-law fit is a

good one.

Many empirical studies in the literature that find power law-distributed data have not attempted to quantitatively test the the goodness of fit of the power law. Instead, these studies typically use qualitative assessments such as visual inspections of the fit. One semi-quantitative approach is the use of the R^2 of the linear regression to measure the goodness of fit. These methods, as discussed earlier, can be misleading and lead to a false claim of power-law existence.

It is often difficult to say with certainty whether a particular dataset has a power-law distribution. Observed distribution is unlikely to follow the power-law form precisely even if the data are drawn from a power law distribution due to the random nature of the sampling process. The challenge is therefore in devising an approach of measuring goodness of fit that will be able to distinguish deviations resulting from statistical fluctuation from those that occur because the data are drawn from a non-power law distribution.

A standard approach to this type of problem is to use a goodness-of-fit test, which generates a p-value that quantifies the credibility of the hypothesis. In this case, the KS statistic is, again, selected as the principal goodness-of-fit measure. In detail, Clauset et al. (2009) propose the following procedure.

First, the power-law model is fitted to the empirical data using the methods presented in Section 4.3.3.1 and 4.3.3.2 and the KS statistic for this fit is calculated. At this stage, we obtain $\hat{\alpha}_0$, $\hat{x}_{min0}$ and KS_0. Next, a large number of power law-distributed synthetic datasets with scaling and lower-bound parameters equal to those of the best-fitted power law estimated earlier ($\hat{\alpha}_0$ and $\hat{x}_{min0}$) are generated. We fit each synthetic data set individually to its own power-law model and calculate the KS statistic for each one relative to its own model. If m synthetic data sets were generated, we would have calculated and stored m KS statistics $(KS_1, KS_2, ..., KS_m)$. The p-value is then the fraction of the time the synthetic KS statistics are larger than the value KS_0 obtained for the empirical data. Formally, the p-value is $\sum_{i=1}^{m} 1\{KS_i > KS_0\}/m$, where $1\{\}$ is the indicator function, taking the value 1 if the statement in the bracket is true.

It is important to note that for each synthetic dataset, the KS statistic is computed relative to the best-fit power law for that dataset, not relative to the best-fit power law for the original observed dataset. In this way, the same calculations performed on the real

dataset are also performed on each synthetic dataset, a crucial requirement to obtain an unbiased estimate of p-value. Under the null hypothesis H_0 that the observed data follow the power-law distribution, a large p-value (close to 1) would indicate that the difference between the empirical data and the fitted model can be attributed to statistical variation alone and there is insufficient evidence to reject the null hypothesis. A sufficiently small p-value would lead to the rejection of the null hypothesis and indicates that the power-law distribution is not a credible fit to the data.

To obtain an accurate estimate of p-value, the generated synthetic data need to have a distribution similar to the empirical data below $\hat{x}_{min0}$ but follow the fitted power law above $\hat{x}_{min0}$. Clauset et al. (2009) make use of a semi-parametric approach to achieve this goal. Suppose that observed data have n_{tail} observations for $x \geq \hat{x}_{min0}$ and n observations in total. A synthetic dataset with n observations is generated as follows. With probability n_{tail}/n, a random number x_i drawn from a power law with scaling parameters $\hat{\alpha}_0$ and $x \geq \hat{x}_{min0}$ is generated. Otherwise, with probability $1 - n_{tail}/n$, one element is selected uniformly at random from among the elements of the observed dataset that have $x < \hat{x}_{min0}$ and set x_i equal to that element. The process is repeated for all $i = 1, ..., n$ to generate a final synthetic dataset that follows the power-law distribution above $\hat{x}_{min0}$ but has the same distribution as the observed data below. For each yearly income and consumption dataset we will generate 1,000 synthetic datasets for the calculation of the p-value.

According to Clauset et al. (2009), the criterion for rejecting the power-law hypothesis is when $p \leq 0.1$. Power law is therefore ruled out if 10 percent or less of the synthetic data produce a fit as poorly as the model of the real data. However, a large p-value does not necessarily indicate that the power law is the correct fit for the data as other competing distributions may fit the data equally well if not better over the range of x observed. Additional tests must be implemented to rule out such alternatives. We discuss these tests in Section 4.3.3.4.

Importantly, for small numbers of observations n, it is also possible for the empirical distribution to follow the power law closely, yielding a large p-value, even when the power law is the incorrect model for the data. This is not a shortcoming of the method described above, it simply reflects the fact that it is harder to rule out the power law hypothesis if there is very little data to work with. Consequently, one must be mindful of a high p-value when

n is small. In addition, since we fit the power-law form to only the part of the distribution above x_{min}, the value of x_{min} will determine the number of data points we have to perform the fit. When x_{min} is large, only a small portion of the dataset lies above it. Therefore, the larger the value of x_{min}, the larger the number of total observations n needed to reliably test the power law hypothesis.

4.3.3.4 Comparison of the Power-law Distribution with Alternative Distributions

Though a large p-value, obtained from performing a hypothesis test using the method described in Section 4.3.3.3, provides evidence in support of the plausibility of the power-law hypothesis, the results of such tests are incomplete. Even if a power-law distribution provides a good fit to the data, it is still possible for another distribution, such as an exponential or lognormal one, to give a fit as good or better. It is therefore important to perform further tests to verify this.

One possible method to test for an alternative distribution is through the use of the goodness-of-fit test presented in Section 4.3.3.3. The p-value for a fit to the competing distribution is calculated using the same method as before, and compared to the p-value for the power law. By comparing p-value calculations with respect to the power law with several plausible alternative distributions, it is possible to make a decision for or against the power-law form in the data. If the p-value for the power law is high, while those for competing distributions are sufficiently small, then the alternatives are ruled out. In this case, although we cannot say for certain that the power law is correct, there is greater support for it.

This goodness-of-fit method requires a test of the credibility of the fit to the data for each and every plausible alternative distribution. In practice, these individual goodness-of-fit tests are unnecessary as we only want to know which distribution is a better fit than the power law. It is therefore only necessary that we perform a goodness-of-fit test for the power law. If that test rejects the power-law hypothesis, then there is no need for a comparison test. On the other hand, if the test fails to reject the power-law hypothesis, then our main concern is whether alternative distributions might provide a better fit.

In these cases, we make use of methods that can directly compare two distributions to each other, and are much easier to implement than the individual KS test. One of

these methods is the likelihood ratio test. Likelihood ratio test computes and compares the likelihood of the data under two competing distributions. The distribution with the higher likelihood is then deem a better fit to the data.

Consider two different competing distributions for the data with PDF $p_1(x)$ and $p_2(x)$ respectively. The likelihoods of the data given the two models are

$$L_1 = \prod_{i=1}^{n} p_1(x_i), \quad L_2 = \prod_{i=1}^{n} p_2(x_i) \tag{4.38}$$

The ratio of the likelihood R is

$$R = \frac{L_1}{L_2} = \prod_{i=1}^{n} \frac{p_1(x_i)}{p_2(x_i)} \tag{4.39}$$

Alternatively one can calculate the logarithm $\mathcal{R}$ of the ratio, which is

$$\mathcal{R} = \sum_{i=1}^{n} [lnp_1(x_i) - lnp_2(x_i)] = \sum_{i=1}^{n} \left[\ell_i^1 - \ell_i^2\right] \tag{4.40}$$

The log likelihood ratio $\mathcal{R}$ is positive or negative depending on which distribution is a better fit, and is equal to zero in the event of a tie. The null hypothesis being tested here is

H_0**:** There is no different between the two competing distributions ($\mathcal{R} = 0$)

The alternative hypothesis H_1 is that there is a difference between the two distributions ($\mathcal{R} \neq 0$), and one provides a better fit. We will need to consider the sign of the likelihood ratio $\mathcal{R}$ in order to determine which distribution is superior.

The sign of $\mathcal{R}$ alone is insufficient to indicate which model is the better fit because it is subject to statistical sampling variation. It is entirely possible that the true value of $\mathcal{R}$ is close to zero while the sign of $\mathcal{R}$ observed is different from zero as a result of statistical sampling fluctuations. Therefore, in order to reach a definitive decision between distributions, we need to observe a log likelihood ratio $\mathcal{R}$ that is sufficiently positive or negative that it could not have plausibly arisen as a result of a chance fluctuation.

In order to make a quantitative judgement whether the observed value of $\mathcal{R}$ is sufficiently far from zero, the standard deviation $\sigma_{\mathcal{R}}$ and the distribution of $\mathcal{R}$ must be known. This can be estimated from the data using a method proposed by Vuong (1989). Recalling Equation (4.40), the log likelihood ratio is the sum of the difference between the individual log likelihoods: $\mathcal{R} = \sum_{i=1}^{n} \left[\ell_i^1 - \ell_i^2\right]$. As under the null hypothesis x_i are independent, the

differences $\ell_i^1 - \ell_i^2$ must also be independent. By the CLT, the sum of these independent differences $\mathcal{R}$ must be normally distributed for large n, with expected variance $n\sigma_{\mathcal{R}}^2$, where $\sigma_{\mathcal{R}}^2$ is the expected variance computed from the data and is given by

$$\sigma_{\mathcal{R}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\ell_i^1 - \ell_i^2\right) - \left(\bar{\ell}^1 - \bar{\ell}^2\right)\right]^2 \tag{4.41}$$

where $\bar{\ell}^1 = \frac{1}{n}\sum_{i=1}^{n}\ell_i^1$, and $\bar{\ell}^2 = \frac{1}{n}\sum_{i=1}^{n}\ell_i^2$.

Then the p-value, the probability that the measured log likelihood ratio has a magnitude as large or larger than the observed value $|\mathcal{R}|$, is

$$p = \frac{1}{\sqrt{2\pi n\sigma_R^2}}\left[\int_{-\infty}^{-|\mathcal{R}|} e^{-t^2/2n\sigma_{\mathcal{R}}^2}dt + \int_{|\mathcal{R}|}^{\infty} e^{-t^2/2n\sigma_{\mathcal{R}}^2}dt\right] \tag{4.42}$$

This method gives a p-value that tells us whether the observed value of $\mathcal{R}$ is statistically significant from zero. If the p-value is small, then the observed value is unlikely to be a result of chance fluctuations and the observed value is a reliable indicator of which model is the better fit to the data. Conversely, if the p-value is large, the observed value is not reliable and the test does not favour either model over the other. One of the advantages of this approach over the simple goodness-of-fit test presented earlier is that it indicates which of the two distributions is favoured, as well as informing us when the data are insufficient to favour either of them.

Finally, it is important to note that it is not possible to compare the power-law fit of the data with fits to every competing distribution, as there is an infinite number of these distributions. In fact, it will almost certainly be possible to find an alternative distribution that fits the data better than the power law if a family of curves is defined with a sufficiently large number of parameters. Therefore, a combination of statistical techniques and prior knowledge about what is a sensible model for the data is needed when attempting to fit statistical distributions to the data.

In this chapter, we will be comparing the power-law distribution fitted to the tail of the Thai income and consumption distribution with the exponential and lognormal distribution fitted over the same range. This involves fitting a truncated exponential and lognormal distribution to the tail of the income and consumption distribution and performing a log likelihood ratio test as described above.

4.3.4 Testing the Lognormal Fit

In this section we present the method proposed by Battistin et al. (2009) for testing the lognormal fit against the Thai income and consumption data. Based on the model of permanent income described in Section 4.2.1.1, we would expect consumption to follow the lognormal distribution more closely than income.

We examine the goodness of fit of the lognormal distribution to the observed data by comparing different characteristics of the empirical distributions of log income and log consumption to their theoretical normal counterparts. One of the tools used to detect departures from normality is the quantile-quantile (QQ) plot. QQ plots are scatterplots of empirical data quantiles against theoretical quantiles of the normal distribution. The data points on the QQ plot will lie along the 45-degree line if the data follow a normal $N(\mu, \sigma^2)$ distribution.

Constructing formal test statistics for normality requires estimation of the location and scale parameters μ and σ. As standard estimates of these parameters can be very sensitive to outliers present in both income and consumption data, we follow Battistin et al. (2009) by using estimates and tests based on robust statistics, which help lessen these impacts. The median $M(Y)$ and the population median absolute deviation $MAD(Y) \equiv M\left(|Y - M(Y)|\right)$ are used as the robust measures of location and scale respectively. Normal distributions $M(Y)$ and $MAD(Y)$ are related to the mean and variance by $M(Y) = \mu$ and $MAD(Y) \simeq 0.6745\sigma$ therefore, the corresponding estimators of the location and scale parameters for a normal distribution are

$$\hat{\mu} = \hat{M}(Y), \quad \hat{\sigma} = \frac{\hat{MAD}(Y)}{0.6745} \tag{4.43}$$

where $\hat{M}(Y)$ and $\hat{MAD}(Y)$ are the sample median and sample median absolute deviation respectively.

Goodness-of-fit tests of the normality hypothesis similar to those described in Section **4.3.3.3** are implemented. First, a KS test based on the distance between the empirical distributions of income and consumption and the theoretical distributions distributed under $N(\hat{\mu}, \hat{\sigma}^2)$ is calculated. To account for estimation error in $\hat{\mu}$ and $\hat{\sigma}$, 10,000 random samples are generated under the null hypothesis of normality, $N(\hat{\mu}, \hat{\sigma}^2)$. P-values for the test are then calculated by counting the number of times that the synthetic samples produced a KS

test statistic greater than or equal to that calculated for the actual data.

The null hypothesis is that the lognormal distribution provides a good fit to the data. Observing a low p-value means that there is only a low probability that we would just by chance get data that agree as poorly with the lognormal fit as the data that we have, and the lognormal fit can be rejected.

Two additional tests based on robust indicators of skewness and kurtosis are performed. Hinkley (1975) and Groeneveld and Meeden (1984) suggest skewness measures of the form

$$\frac{[Q_{1-p}(Y) - M(Y)] - [M(Y) - Q_p(Y)]}{Q_{1-p}(Y) - Q_p(Y)} \tag{4.44}$$

where $Q_p(Y)$ is the p-th percentile of the distribution of Y. We follow Battistin et al. (2009) in using quartile skewness, which takes $p = 0.25$ and is zero for normal distributions. The positive (negative) values of this statistic indicate right (left) skewness. Additionally, this coefficient will take values in the interval $(-1, 1)$, with 1 (-1) representing extreme right (left) skewness.

For kurtosis we use

$$\frac{[O_7(Y) - O_5(Y)] + [O_3(Y) - O_1(Y)]}{O_6(Y) - O_2(Y)} \tag{4.45}$$

where $O_\beta(Y)$ is the β-th octile of the distribution of Y. This statistic is non-negative and not very sensitive to the extreme tails of the distribution, and for normal distributions it equals 1.233 (see Moors, 1988).

As before, the sample statistics of both the skewness coefficient (4.44) and the kurtosis coefficient (4.45) are computed and compared against their theoretical value under the assumption of normality. The p-value under the null hypothesis of normality is computed from 10,000 synthetic samples using the same method as previously discussed.

4.3.5 Kullback-Liebler Divergence

The Battistin et al. (2009) method discussed in the previous subsection is able to detect whether the income or consumption distribution follows the lognormal distribution or not.

However, the argument based on the proportional random growth of permanent income does not strictly require consumption to follow the lognormal distribution, and income does not. In this case it is sufficient to show that the distribution of consumption, the permanent component of income, is closer to lognormal than the distribution of total income. Given this, in the case where both consumption and income fail Battistin et al.'s test of lognormality, it is necessary to use an alternative method that provides a measure of the distance between two distributions in order to help determine which distribution is closer to lognormal.

One of the methods that provides a measure of the distance between two distributions is a method based on relative entropy, namely the Kullback-Liebler (KL) divergence. Relative entropy has its roots in information theory literature. Equation (4.46) defines the concept of Shannon entropy, which is a measure of the average amount of information or uncertainty contained in the data.

$$H(X) \;=\; E(I(X)) \;=\; -\sum_{i=1}^{n} p(x_i) log[p(x_i)] \tag{4.46}$$

The KL divergence is defined as

$$D(p\|q) \;=\; \sum_{i=1}^{n} p(x_i) log[\frac{p(x_i)}{q(x_i)}] \tag{4.47}$$

$$D(p\|q) \;=\; -\sum_{i=1}^{n} p(x_i) log[q(x_i)] \;-\; \left\{ -\sum_{i=1}^{n} p(x_i) log[p(x_i)] \right\} \tag{4.48}$$

Where $0log[\frac{0}{0}] = 0$, $log[\frac{0}{q}] = 0$ and $log[\frac{1}{0}] = \infty$. The KL divergence provides a measure of the amount of information, in bits or nats, required to code samples from P when using a code optimised for Q, rather than using the true code optimised for P. Typically P represents the "true" distribution of the data. The measure Q typically represents a theory, a model, or an approximation of P. KL divergence of Q from P is a measure of information lost when Q is used to approximate P. However it is not a proper distance since it is not symmetric and does not satisfy the triangular inequality. However, it shares the properties of weak positivity, $D(p\|q) \geq 0$, and $D(p\|q) = 0 \leftrightarrow p = q$ with a distance.

Therefore, for the purpose of determining which distribution is closer to lognormal, we will be comparing the KL divergence of consumption from lognormal with the KL divergence of income from lognormal later in Section 4.5.

4.4 Data and Descriptive Statistics

In this section, we present descriptions of the dataset used for fitting the power law and the lognormal distribution to income and consumption in Thailand, the SES.

4.4.1 Household Socio-Economic Survey (SES)

The first SES was conducted by the National Statistical Office (NSO) in 1957 and the survey was carried out every five years. Due to rapid economic expansion and the importance of the survey as a source of data for anti-poverty policy, the NSO was commissioned to carry out the survey every two years from September 1987.

The survey is representative of all private, non-institutional households residing permanently in municipal areas, districts, and villages of all regions in Thailand. The survey employed a stratified three-stage sample design from 1975-76 to 1986. The primary sampling units were amphoes (districts). The secondary sampling units were blocks (in municipal areas) or villages (non-municipal areas). The ultimate sampling units were the households. From 1988 survey onwards, a stratified two-stage sampling was adopted instead. The primary sampling units are blocks for municipal areas and villages for non-municipal areas, while the secondary sampling units are private households.

The primary objective of the survey is to collect information on household income, expenditures, consumption, changes in assets and liabilities, durable goods ownership, and living conditions. The survey also included detailed information on household members' characteristics such as sex, age, education, and occupation. Each biennial survey from 1986 to 2006 contains data for 10,000 – 40,000 households, representing around 0.05 – 0.2 percent of the total Thai population.

In this chapter, we used SES data from 1986 to 2006 (11 surveys in total). This covers the period of the Thai economic boom starting from the mid-1980s to 1995, the Asian financial crisis in 1997, and the economic recovery from the early to mid-2000s. The most recent surveys after 2006 are not used in this study as the SES from these periods no longer provide information on total household income.

The preference for the SES over the LFS employed in the previous chapter is based on two main considerations. Firstly, as discussed in Chapter 3, the LFS data had stopped collecting information on individual income from the year 2000 onwards. The SES, on the

other hand, collects income information up to the year 2006 thus allowing for a longer study of the income distribution. In addition, the SES also collect expenditure data, a proxy for consumption, which is not available in the LFS. Secondly, the income and consumption data in the SES is not top-coded unlike the income and earnings data collected in the LFS. Given that we are interested in fitting the power-law distribution to the top-end of the income distribution, the absence of top-coding is an essential requirement.

4.4.2 Variable Construction

In this chapter, the two main variables used for fitting the power-law and lognormal distributions are income and consumption.

Income and consumption both represent measures of the well-being of the household, which are often claimed in the literature to follow a power-law distribution. Pareto's original works actually claim that the wealth distribution, measured by asset holding (land), follows the power-law form. In this spirit, it is households' accumulated wealth that should be investigated. While the SES collected data on household assets and liabilities, it only started to collect these on an absolute level in the 2006 survey. In all previous surveys, the information was only collected on the change in asset and liabilities, and it is not possible to compute total household accumulated wealth from this information. Given this data limitation, the focus in this chapter is therefore on household income and consumption.

The variable income measured by SES is made up of the monthly recall of individual wages and salaries, net profit from businesses, net profit from farming, income from pension, income from work compensation and termination payment, remittances, income from government assistance programmes (for the elderly or the disabled), rental income, interest earned on savings and investment returns. In-kind income such as those from home produced food, food received as part of wages and rental value of owner occupied home are also included in the total income calculation. However, the income variable excludes one-off payments such as income received from inheritance and proceeds from insurance.

The variable consumption consists of both the weekly recall of individual in-cash and in-kind consumption expenditure on items such as housing (rental expense, furnishing expense, etc.), apparel, medical care, transportation, education, entertainment, personal care, tobacco, food and beverages. Consumption, however, excludes items such as taxes, insurance

premium, interest on debts, gift and contributions from its calculation. These expenditures represent the permanent component of income, while the total income variable also includes transitory elements as discussed earlier.

Income and consumption variables used in this chapter consist of the same components and are comparable to those measured in the US Consumer Expenditure Survey published by the US Bureau of Labour Statistics (BLS), which was employed by Battistin et al. (2009) in their study of the income and consumption distribution in the US. This allows for a fair comparison of the results obtained later in Subsection 4.5.2 from fitting the lognormal distribution to both income and consumption data with the results of Battistin et al. (2009).

Both income and consumption are measured at the household level as households tend to pool and share resources, making this a better representation of well-being than individual income or consumption. Both variables are adjusted for differences in the price level faced from year to year using Thailand Consumer Price Index (CPI) data. Demographic adjustments are also made to the data by dividing total household income and consumption by household size to arrive at per-capita income and consumption. These price and demographic adjustments allow income and consumption to be compared across different households and time periods. The final variables used for fitting both the power-law and lognormal distributions are the real per-capita household income and consumption.

Per capita adjustment is employed instead of equivalence scale used in Chapter 3 to facilitate comparison of the final estimates of the power-law exponent with those in the literature. While we acknowledge the possible presence of measurement errors in the data, its impact on the magnitude and the direction of the biases in the estimates of the power-law exponents are difficult to quantify and to correct for given the lack of information on the nature of the measurement error. As the emphasis of this chapter is on fitting the distribution to available data instead of trying to find consistent estimates of the parameter of interests, we will not be too overly concern with the issue of measurement error in the income and consumption variables.

Figure 4.1 and 4.2 shows the histogram of log real per capita income and consumption in all datasets used respectively. It is clear from both figures that log income and log consumption in all the years deviate from the normal fit, thus providing early indication that income and consumption may not be well-fitted by the lognormal distribution. The

histograms also illustrate that both income and consumption data are not top-coded as there are no evidence of spikes at the top end in either of the distribution.

Figure 4.1: Histogram: Log Real Per Capita Income

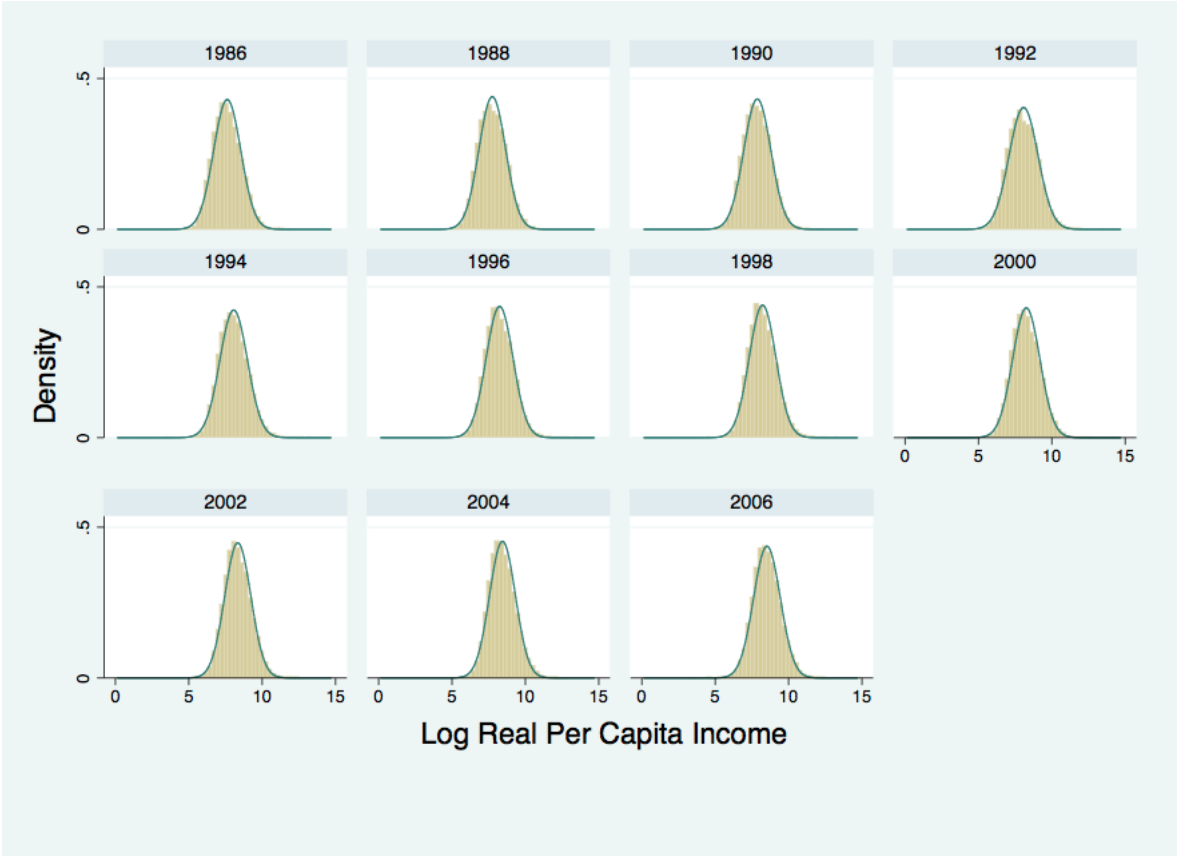


Figure 4.2: Histogram: Log Real Per Capita Consumption

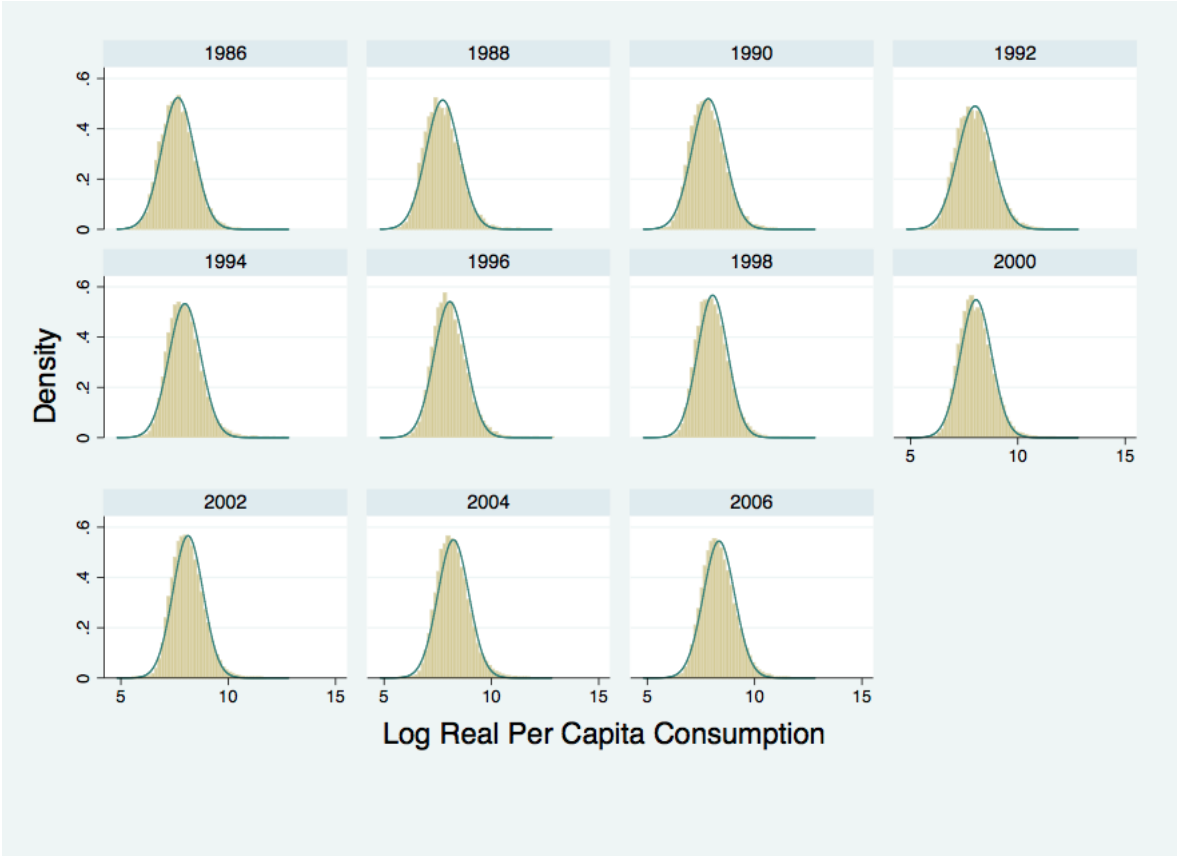


Figure 4.3: Real Per Capita Income and Consumption: 1986-2006

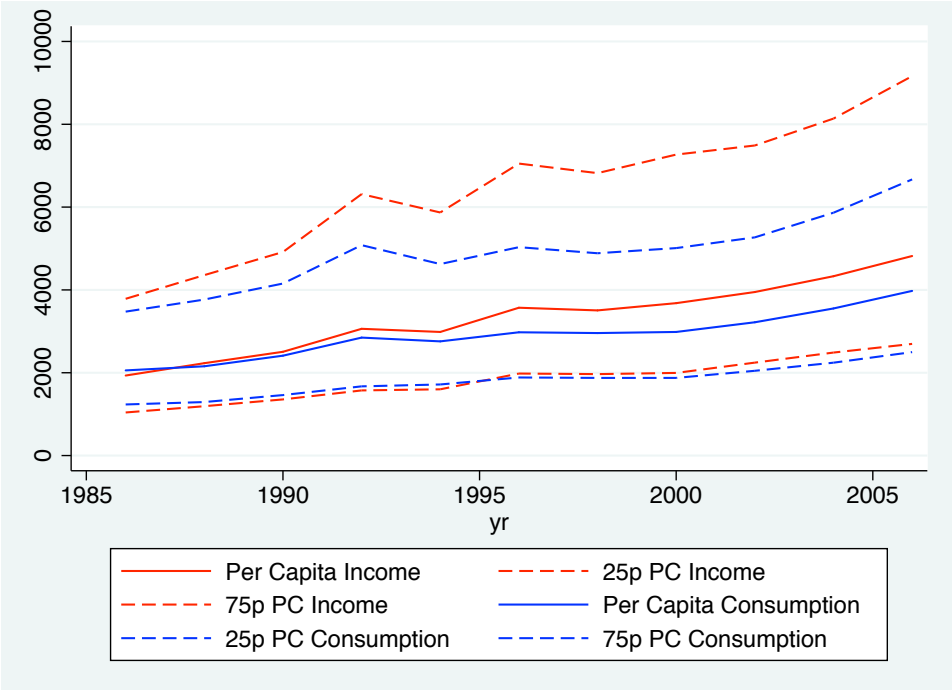


Figure 4.3 plots the median, the 25th percentile and the 75th percentile of the real

per-capita income and consumption. We can see from the plot that the two series show a similar trend during the period of study. Both income and consumption increased during the economic boom period in the late 1980s to the early 1990s and the early to mid-2000s with a dip in both income and consumption between 1996 and 1998, which is the period encompassing the 1997 Asian financial crisis. The decrease observed is not as pronounced because the data for 1997 were not collected. The dip observed in 1994 is due to the great increase in inflation that caused the real per-capita income and consumption to fall slightly even though the nominal amount was increasing. Income and consumption adjustments may lag behind inflation adjustment during this period. This adjustment lag also explains the big increase in both variables in 1996 after sufficient time had passed.

Another important observation from Figure 4.3 is the relative stability of per-capita consumption in comparison to that of total per-capita income. Median per-capita consumption varies less from year to year than income and the range from the 25th to the 75th percentile is also narrower. This fits with the hypothesis that consumption can be taken to represent the permanent component of income.

Figure 4.4: Gini: Income vs. Consumption

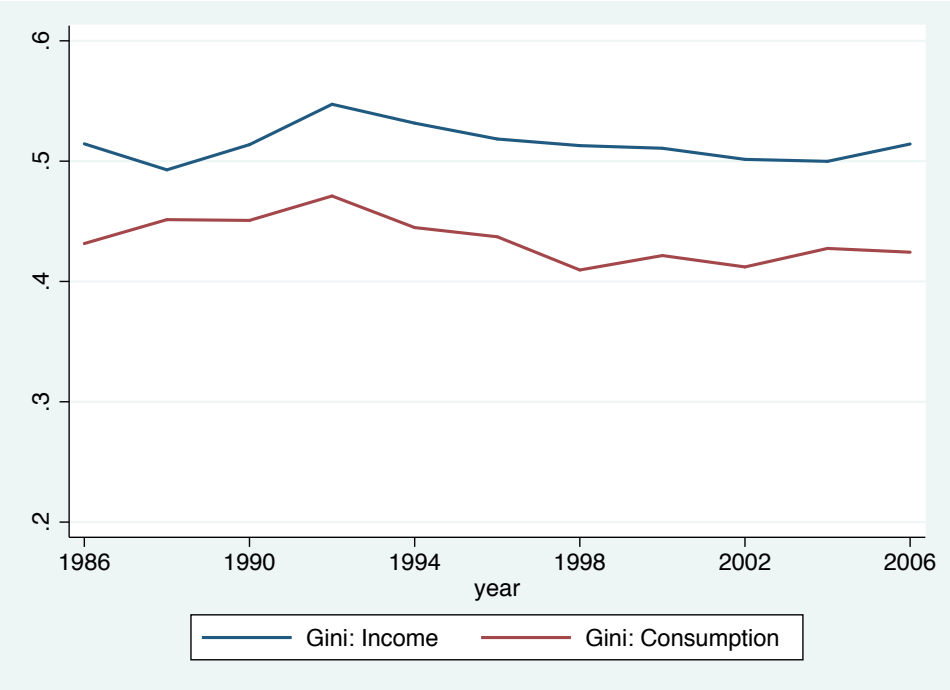


Figure 4.4 plots the Gini coefficient calculated from both the income and consumption data from 1986 to 2006. Both sets of Gini coefficients exhibit similar pattern; inequality

increased slightly from the late 1980s to the early 1990s as Thailand underwent rapid economic growth during this period. Inequality then started to decline before taking noticeable dips during the financial crisis between 1996 and 1998. Despite these movements, the Gini coefficients for both distributions are relatively high and stable throughout the period of study. Consumption is less unequal than income as expected, given redistributive tax and transfers as well as the fact that consumption is less affected by transitory shocks. The Gini coefficients computed here are for the entire distribution, while the power-law exponent estimates that will be discuss in the next section will help shed further light on inequality at the very top of both distributions.

In the next section we present the results of fitting the power-law distribution to the income and consumption distribution in Thailand using both the linear regression and the approach proposed by Clauset et al. (2009). We also discuss findings on the dynamics of inequality at the top end of the distribution and link these back to the overall inequality presented in this section.

4.5 Results

This section presents the estimates of the power-law exponent in the Thai income and consumption distribution. We first present the estimates obtained from the simple log rank-size OLS regression method. We then proceed to compare and contrast these estimates with those obtained using the method proposed by Clauset et al., 2009 (henceforth CNS) based on the MLE framework. Lastly, the findings from lognormality of income and consumption study are discussed.

4.5.1 Power-law Fit: Income and Consumption

4.5.1.1 Linear Regression Method

This section presents the result of an attempt to fit the power-law form to the income and consumption data using the linear regression method.

As discussed in Section 4.3.2, if the data follow the power-law distribution, then the CDF plot will follow a straight line. Estimates of the power-law exponent α and the goodness of fit are therefore obtained from fitting an OLS regression line to the log rank-size plot, which represents the counter CDF. The OLS estimation is only performed on a subsample of the data that lie above the minimum cut-off level, x_{min}, estimated using the KS method proposed by CNS. This step is taken to ensure that the OLS α estimates can be compared on the same basis as the MLE estimates later. Essentially, the scaling parameter α is the OLS estimate of the slope of the line that best fits the section of the data above x_{min}, while the goodness of fit is given by the R^2 of the regression. As stated earlier, a linear fit to the CDF plot will yield a slope parameter that is equivalent to $\alpha - 1$ instead of the original α obtained using the MLE method. Therefore, in this case, we present the estimate of α in Table 4.1 and 4.2 by adding one to the original OLS estimates.

Table 4.1 contains the OLS estimates of α for the income distribution from 1986 to 2006. The scaling parameters are all statistically significant at the 1% level with a magnitude lying approximately between 3 and 3.6. More importantly, the R^2 all exceed 98%, which indicates that the straight line is a very strong fit to the data.

Table 4.2 shows the estimates of α for consumption. The power-law exponents of the consumption distribution lie approximately between 3.2 and 3.8, similar in magnitude to that of the income distribution. Again, the R^2 values for all years are all greater than

0.99, providing strong evidence in support of the power-law fit. Therefore, based on the OLS regression method, it is reasonable to conclude that both income and consumption in Thailand strongly follow the power-law distribution.

Figure 4.5: $\hat{\alpha}$ OLS Estimates: Income vs. Consumption

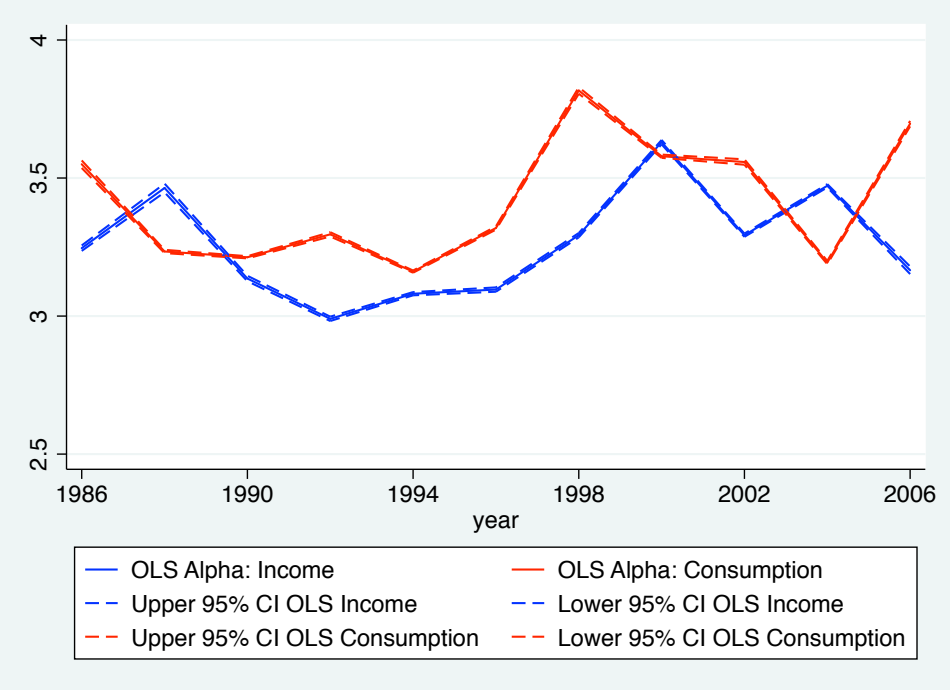


Figure 4.5 shows the plot of $\hat{\alpha}$ of both income and consumption distribution from 1986 to 2006. The magnitude of the exponents is slightly greater for consumption than income in most years, which indicates that the consumption distribution is more equally distributed than the income distribution. This finding is intuitive as we usually observe more inequality in income than consumption. This is usually attributed to the distributive tax and transfers that help equalise consumption when compared against pre-tax income figures. Income also includes more transitory elements such as large positive and negative shocks that increase its volatility when compared to consumption, which is more stable due to the act of consumption smoothing (Meyer and Sullivan, 2003).

We also observed from Figure 4.5 that the scaling exponents are relatively stable over time for both income and consumption. This is consistent with earlier findings of studies on inequality in Thailand, which find the Gini coefficient to be stable over time. However, the Gini coefficient is computed for the entire distribution while the power-law exponents only represent inequality at the top end of the distribution. Figure 4.6 plots income and

Table 4.1: Income: Linear Regression Method

	1986 Rank	1988 Rank	1990 Rank	1992 Rank	1994 Rank	1006 Rank	1998 Rank	2000 Rank	2002 Rank	2004 Rank	2006 Rank
Income	-3.246***	-3.464***	-3.136***	-2.990***	-3.081***	-3.096***	-3.293***	-3.630***	-3.293***	-3.472***	-3.166***
	(0.005)	(0.008)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.007)
Constant	26.928***	29.106***	26.558***	25.879***	27.261***	27.656***	29.590***	33.200***	30.085***	32.260***	29.513***
	(0.047)	(0.075)	(0.044)	(0.042)	(0.032)	(0.042)	(0.044)	(0.042)	(0.035)	(0.031)	(0.074)
Observations	984	1310	1123	659	1189	1099	745	553	1682	1014	1764
R^2	0.995	0.986	0.995	0.997	0.997	0.996	0.998	0.999	0.996	0.999	0.982

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.2: Consumption: Linear Regression Method

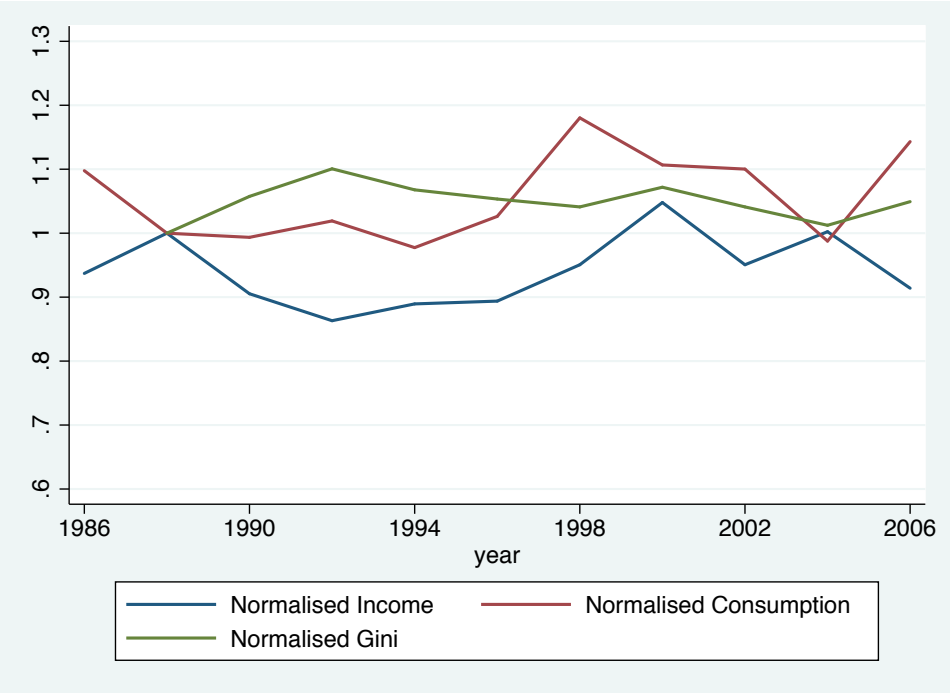
	1986 Rank	1988 Rank	1990 Rank	1992 Rank	1994 Rank	1006 Rank	1998 Rank	2000 Rank	2002 Rank	2004 Rank	2006 Rank
Consumption	-3.550***	-3.234***	-3.213***	-3.296***	-3.161***	-3.319***	-3.817***	-3.579***	-3.558***	-3.194***	-3.697***
	(0.007)	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.006)	(0.003)	(0.005)	(0.003)	(0.005)
Constant	29.117***	26.550***	26.763***	28.181***	27.228***	28.869***	33.337***	31.174***	31.445***	28.270***	33.758***
	(0.060)	(0.028)	(0.023)	(0.040)	(0.022)	(0.025)	(0.059)	(0.030)	(0.045)	(0.024)	(0.047)
Observations	1273	1387	954	712	2239	1558	794	1468	1331	5692	2397
R^2	0.992	0.997	0.999	0.998	0.997	0.998	0.996	0.998	0.996	0.992	0.993

b coefficients; se in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

consumption's $\hat{\alpha}$ with the Thai Gini coefficient. All three variables have been normalised so that they are equal to 1 in year 1988.[6] These plots confirm that all three series are relatively stable, though income and consumption's $\hat{\alpha}$ are more volatile, suggesting that there are more changes in inequality at the top end of both distributions than at the overall distribution level.

Figure 4.6: Income $\hat{\alpha}$, Consumption $\hat{\alpha}$ and Gini Coefficient



It is important to note that smaller value of $\hat{\alpha}$ represents greater inequality; this is why we see the Gini and the consumption and income's $\hat{\alpha}$ series moving in opposite directions in most years. This shows that inequality at the top and the rest of the distribution mostly moves in tandem. The greatest jump in consumption and income inequality at the top occurs around the Asian financial crisis between 1996 and 1998. Consumption and income inequality at the top falls more than overall inequality, suggesting that the top was more affected by the crisis than the rest.

The magnitude of $\hat{\alpha}$ for both consumption and income is greater than what is usually observed ($2 < \alpha < 3$). There are two possible explanations for this. First, it may be that Thai income and consumption are distributed more equally than in other countries. This explanation seems implausible given that Thailand is one of the developing Asian economies,

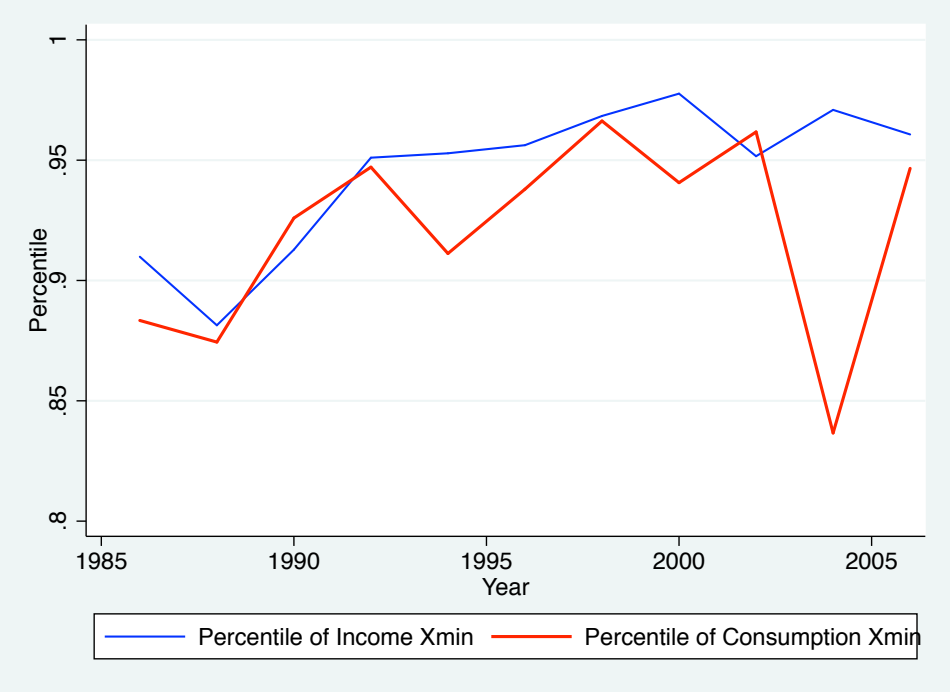
[6]The official Gini coefficient released by the National Economic and Social Development Board (NESDB) for Thailand is only available from 1988 onwards.

which are usually characterised by a high level of inequality and extreme concentration of wealth at the top end of the distribution. Coefficients for overall income inequality in Thailand are approximately 0.5 for all years indicating high level of inequality; the average for OECD countries was 0.32 in 2012.[7] A more likely explanation is the impact on estimates of α from reporting errors and survey non-response by the rich/super-rich at the top-end of the income distribution. Top earners usually misreport their income due to lack of confidence in the anonymity of the data or may not be represented in the survey at all due to the time constraints nature of their work. If top incomes are under-reported, then income would appear more equal causing estimates of α to be higher than its actual value.

Another explanation is of a more technical nature. Past studies of power law usually attempted to fit the power-law distribution to the entire sample or, at the very best, eyeballing the cut-off point x_{min}. It is possible that these studies fitted the power-law form to more data points than they should have. Given that the x_{min} used in the OLS regressions in Table 4.1 and 4.2 are estimated using the KS approach advocated by CNS, the power-law form is fitted to a smaller and better subset of the data. This could possibly explain the observation of greater equality, as more individuals with low income/consumption levels are excluded from the fit. Figure 4.7 shows the plot of $\hat{x}_{min}$ expressed as a percentile of the respective income and consumption distribution over time. This plot shows that only a small portion of the sample in each year is used in the calculation of $\hat{\alpha}$ and that position of $\hat{x}_{min}$ within the respective distribution is not stable over time. Therefore, having a method that can reliably select the correct value of $\hat{x}_{min}$ is crucial to the fitting of the power-law distribution to the data. On average, approximately 6% and 8% of the total sample were above the $\hat{x}_{min}$ cut-off and were included in the OLS estimation for income and consumption respectively.

[7] http://www.oecd.org/social

Figure 4.7: $\hat{X}_{min}$ as a Percentile of Income and Consumption Distribution 1986-2006



To further illustrate the point above, Figure 4.8 and 4.9 provide a plot of OLS estimates of $\hat{\alpha}$ against the number of observations in the tail of the income distribution that the OLS regression was fitted to (N_{tail}) for the year 1990 and 1988 respectively. These types of plot are often used to select the portion of data above which the power-law fit applied. The selection of N_{tail} in this case is just another way of representing the selection of $\hat{x}_{min}$ discussed above. In an ideal case, one should be able to identify the point beyond which the value of $\hat{\alpha}$ appears relatively stable, which allows for the selection of N_{tail} or $\hat{x}_{min}$ and the identification of $\hat{\alpha}$.

Figure 4.8 provides an illustration of this ideal case where the $\hat{\alpha}$ and N_{tail} selected by this method (represented by the green lines), though still subjective in nature, approximate the OLS $\hat{\alpha}$ that was estimated using CNS $\hat{x}_{min}$ (represented by the red lines) very closely. However, Figure 4.9 illustrates the alternative case where it is not clear from the plot which value of $\hat{\alpha}$ and N_{tail} to select. Notice in both cases that the absolute magnitude of $\hat{\alpha}$ kept decreasing as more observations were included in N_{tail}, indicating an increase in inequality as more and more diverse income observations were used in the calculation of $\hat{\alpha}$ as discussed above.

In addition, the existence of clear choices of $\hat{\alpha}$ and N_{tail} does not give any indication whether the data are well-fitted by the power-law. The R^2 for both regressions in the year 1988 and 1990 shown in Table 4.3 exceed 0.96, indicating a strong evidence of power-law

in both samples despite the 1990 sample showing a much better convergence of $\hat{\alpha}$ than the 1988 sample. Later, we show in Table 4.3, the result from employing the CNS approach, that neither of the samples are well-fitted by the power-law distribution. These plots help demonstrate the unreliability of the traditional method used in the selection of $\hat{x}_{min}$ in comparison to the KS method proposed by CNS as well as its shortcomings in identifying the presence of the power-law distribution.

Figure 4.8: Income OLS $\hat{\alpha}$ - N Tail Plot: 1990

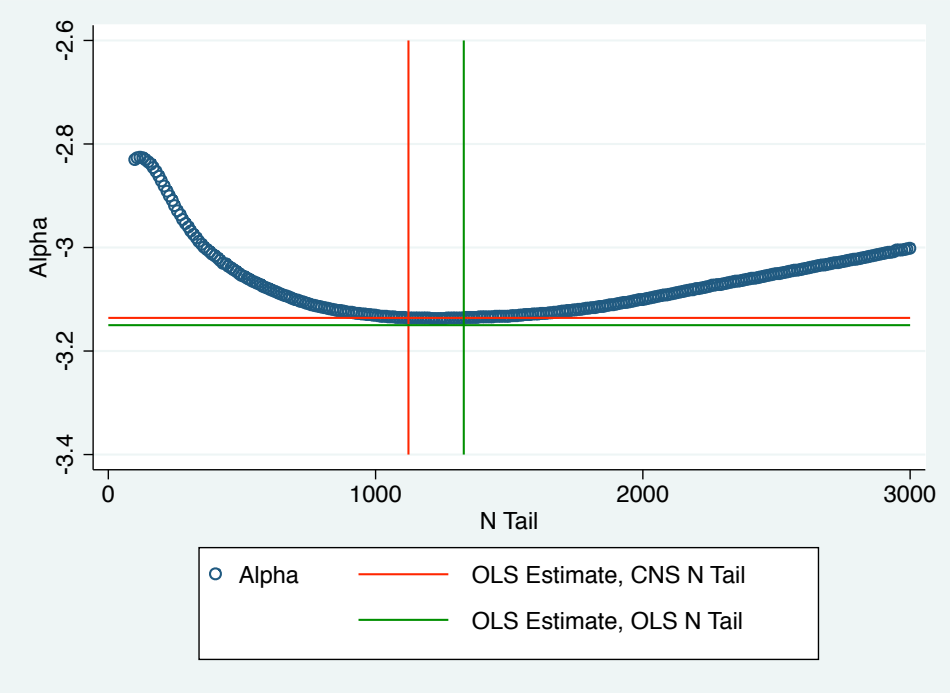
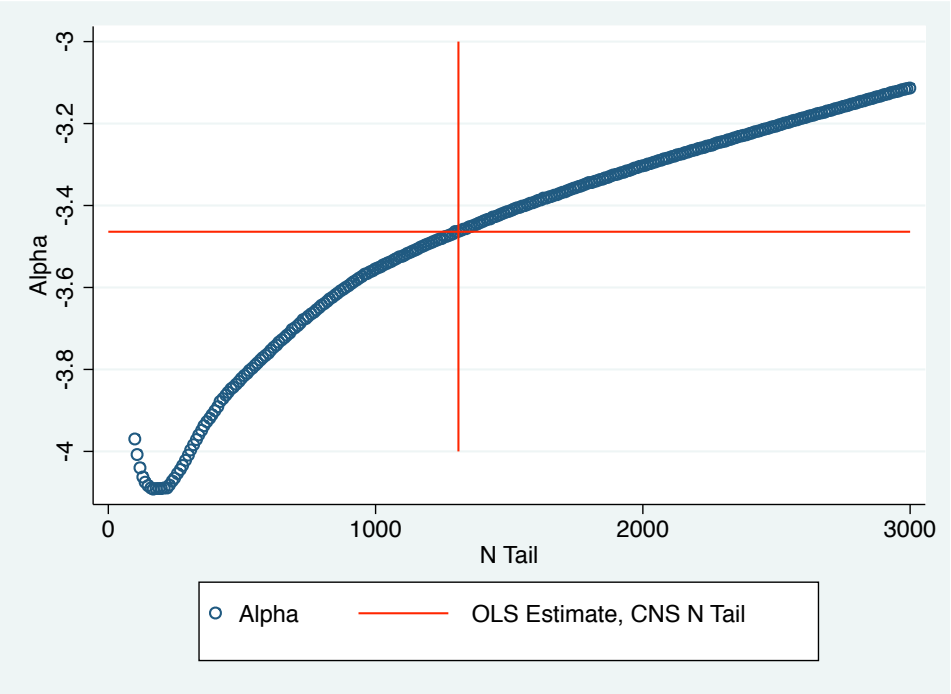


Figure 4.9: Income OLS $\hat{\alpha}$ - N Tail Plot: 1988



4.5.1.2 CNS Method

This section presents the result of fitting the power law to the Thai income and consumption distribution using the CNS method.

Table 4.3 and 4.4, show some descriptive statistics and the results of fitting the power-law distribution to the income and consumption data respectively. The power-law exponents $\hat{\alpha}$ are estimated based on the MLE approach only on the subsample of the data that lie above $\hat{x}_{min}$, estimated using the KS method proposed by CNS. Recall that $\hat{x}_{min}$ here are the same as the cut-off point used in the OLS estimation earlier. The p-values reported are computed using the bootstrap methodology discussed in Section 4.3.3.3. A p-value greater than 0.1 indicates that the null hypothesis, which is that the power-law distribution is a good fit to the empirical data, cannot be rejected.

Table 4.3: Income: Power-law Fit 1986-2006

Year	n	μ	σ	x_{max}	$\hat{x}_{min}$	$\hat{\alpha}$	n_{tail}	p
1986	10917	3245.72	4766.24	142228	7433 ± 961	3.203	984	**0.377**
1988	11044	3553.47	4293.82	138256	6985 ± 2462	3.198	1310	0.000
1990	12881	4184.60	6845.96	269723	9418 ± 1494	3.176	1123	0.061
1992	13458	5533.76	10543.03	631791	17058 ± 3012	2.989	659	**0.611**
1994	25220	5209.76	9356.77	523461	16422 ± 2682	3.137	1189	**0.881**
1996	25108	6102.87	11013.22	706564	19365 ± 4283	3.181	1099	**0.738**
1998	23549	5928.04	9256.55	496000	22730 ± 4825	3.373	745	**0.702**
2000	24732	6141.46	8296.16	286697	27545 ± 5879	3.617	553	**0.682**
2002	34784	6491.20	9855.90	469133	$19675 \pm 3508.$	3.307	1682	0.085
2004	34840	7129.72	10044.98	473640	28335 ± 3653	3.493	1014	**0.855**
2006	44872	8063.26	19792.91	2419921	27062 ± 5355	3.337	1764	0.004

Table 4.4: Consumption: Power-law Fit 1986-2006

Year	n	μ	σ	x_{max}	$\hat{x}_{min}$	$\hat{\alpha}$	n_{tail}	p
1986	10917	2902.26	3208.94	143214	5357 ± 1593	3.378	1273	0.003
1988	11044	3211.04	4030.54	121579	5620 ± 895	3.174	1387	**0.176**
1990	12881	3588.67	4785.37	196505	8060 ± 893	3.222	954	**0.764**
1992	13458	4386.95	5733.02	176048	12204 ± 2286	3.291	712	**0.481**
1994	25220	4070.62	5398.80	293744	8171 ± 1755	3.072	2239	0.009
1996	25108	4375.32	5757.44	369270	10585 ± 3037	3.255	1558	0.012
1998	23549	4114.95	4152.69	107542	12728 ± 2959	3.729	794	**0.456**
2000	24732	4229.98	4503.82	123846	10455 ± 989	3.554	1468	**0.619**
2002	34784	4485.01	4773.06	119262	13023 ± 1918	3.513	1331	**0.499**
2004	34840	5080.00	5733.32	178979	7427 ± 2037	3.053	5692	0.000
2006	44872	5665.01	5930.59	144378	14841 ± 2539	3.532	2397	0.004

Table 4.3 shows the result of fitting the power law to the Thai income distribution. As previously observed, the power-law exponents $\hat{\alpha}$ are quite stable over time ranging, from 3–3.5. The value of $\hat{\alpha}$ estimated here appear to be very similar in magnitude to those estimated using the OLS approach. Figure 4.10 plots income $\hat{\alpha}_{OLS}$ against $\hat{\alpha}_{CNS}$ estimates. The two series track each other very closely.

Figure 4.10: Income $\hat{\alpha}$: CNS vs. OLS

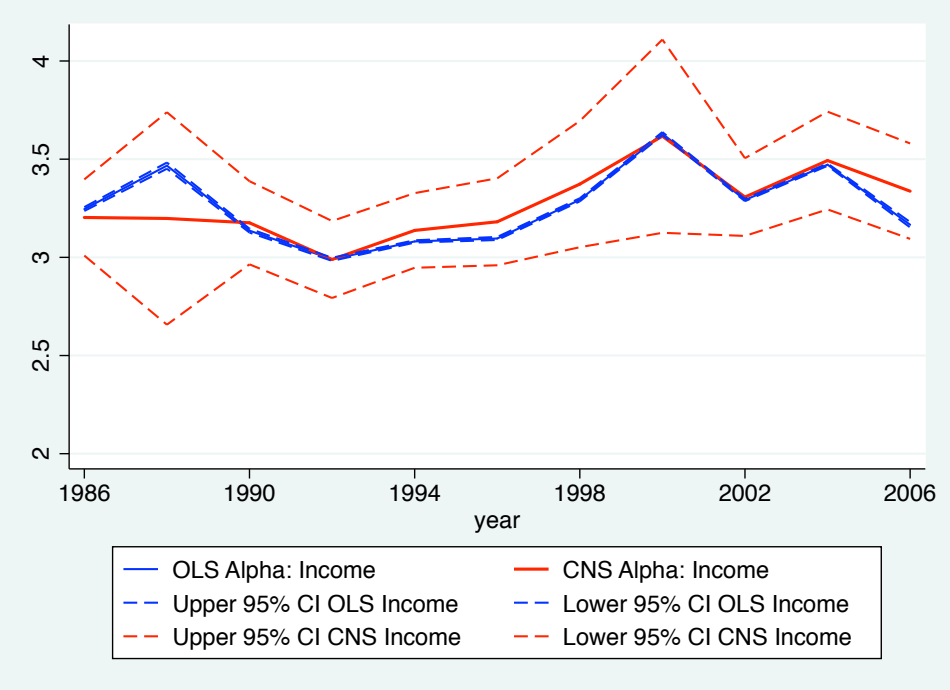
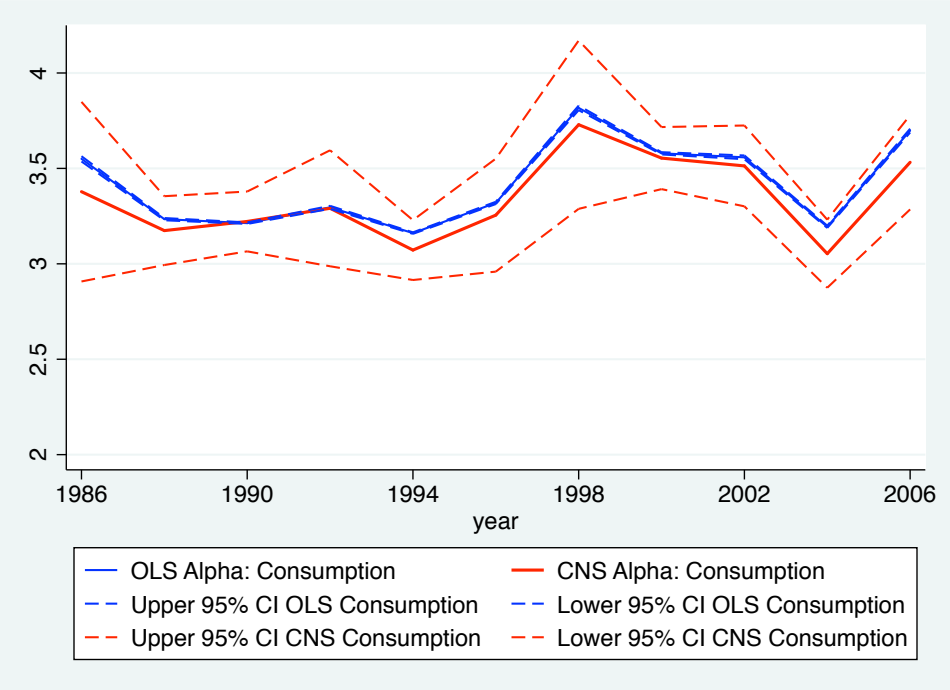


Table 4.4 shows a similar result for consumption. $\hat{\alpha}_{CNS}$ are stable over time ranging from 3-3.7. Again, Figure 4.11, which plots $\hat{\alpha}_{OLS}$ against $\hat{\alpha}_{CNS}$ estimates for consumption, also shows that estimates for the two approaches are very similar.

Figure 4.11: Consumption $\hat{\alpha}$: CNS vs. OLS



From this evidence we may be able to conclude that both OLS and MLE estimators

are capable of producing accurate estimates of the power-law exponents. This is consistent with the evidence from a simulation exercise performed by CNS (Table 3.1, Clauset et al. (2009)), which showed that the least-squares fit to the rank-frequency CDF yields an accurate estimate of real α and is almost as good as the MLE estimator.[8] However, a small caveat from drawing this conclusion is that in estimating $\hat{\alpha}_{OLS}$, $\hat{x}_{min}$, estimates from the KS approach proposed by CNS were used. Without knowing $\hat{x}_{min}$, we would have to fit the linear regression to the entire distribution or, at the very best, eyeball a region of the CDF to find the best cut-off point. Therefore, without using the KS approach to estimate $\hat{x}_{min}$ in advance, it will be difficult for the pure OLS approach used in the literature to produce estimates of α close to those based on the CNS approach.

While the two different approaches produce similar estimates of α given known x_{min}, the major difference lies in the result of the goodness-of-fit test. It is clear from Tables 4.3 and 4.4 that the power-law distribution does not pass the goodness-of-fit test for income and consumption in all the years. For income, the power-law distribution passes the test in 7 out of 11 years. The consumption power-law fit passes the test in 6 out of 11 years. While this still provides some evidence that income and consumption may follow the power-law distribution, the support for power law is much weaker than that provided by the R^2 from the OLS regressions. This clearly illustrates the weakness of the OLS approach, which tends to overstate the fit of the power law to most heavy-tailed dataset, resulting in the misleading conclusion that these datasets follow the power-law distribution when, in fact, most do not. For these reasons, the CNS approach offers a superior method for fitting the power-law distribution accurately to income and consumption data than the OLS approach, and should be the default approach going forward.

Comparing Power-Law Fit with Competing Alternatives

Despite showing weaker evidence in support of power law than the OLS approach, the results from the CNS approach still show that the power-law fit cannot be rejected for the majority of income and consumption data. This in itself, however, is not sufficient to conclude that the data follow the power-law distribution. The results in Tables 4.3 and 4.4 merely show that power law provides a good fit to the Thai income and consumption data in some years,

[8] While the MLE is a superior estimator in terms of accuracy, it is only accurate if the right functional form is specified. It matters if the distribution is discrete or continuous. For the OLS fit to the inverse cumulative CDF, functional form is less of a concern.

but this does not rule out the possibility that other competing distributions could fit the data just as well as, if not better than, the power law distribution.

The CNS approach advocates further tests of the power-law fit against alternative distributions. Tables 4.5 and 4.6 restate the p-value of the power-law fit and report the Vuong likelihood ratio (LR) test statistics and their corresponding p-values for income and consumption respectively. The LR tests are computed to compare the power-law fit with that of alternative distributions. In this case the alternatives are the tail of the lognormal distribution that was fitted to the whole dataset (column "Lognormal 1"), the truncated lognormal distribution fitted to the tail only, and the truncated exponential distribution fitted to the tail only. The cut-off for the tail is $\hat{x}_{min}$, estimated earlier using the KS approach. Recall that a positive Vuong statistic indicates that power law is a better fit than the competing distributions and vice versa.

Table 4.5: Income: Power-law vs. Lognormal vs. Exponential

	Power-law	Lognormal 1		Lognormal 2		Exponential		Support
Year	p-value	LR	p-value	LR	p-value	LR	p-value	Power-law
1986	**0.377**	1.527	0.127	-0.294	0.769	8.811	**0.000**	Moderate
1988	0.000	-1.853	0.064	-2.377	**0.017**	10.295	**0.000**	None
1990	0.061	2.284	**0.022**	0.117	0.907	8.457	**0.000**	None
1992	**0.611**	1.930	0.054	-0.135	0.892	8.013	**0.000**	Moderate
1994	**0.881**	2.613	**0.009**	0.453	0.650	10.892	**0.000**	Moderate
1996	**0.738**	2.671	**0.008**	0.726	0.468	9.985	**0.000**	Moderate
1998	**0.702**	1.874	0.061	0.421	0.674	10.447	**0.000**	Moderate
2000	**0.682**	0.736	0.462	-0.313	0.754	11.126	**0.000**	Moderate
2002	0.085	2.366	**0.018**	0.105	0.917	13.403	**0.000**	None
2004	**0.855**	1.669	0.095	-0.090	0.928	15.143	**0.000**	Moderate
2006	0.004	2.899	**0.004**	1.638	0.101	8.372	**0.000**	None

Table 4.6: Consumption: Power-law vs. Lognormal vs. Exponential

	Power-law	Lognormal 1		Lognormal 2		Exponential		Support
Year	p-value	LR	p-value	LR	p-value	LR	p-value	Power-law
1986	0.003	-0.303	0.762	-1.500	0.133	5.988	**0.000**	None
1988	**0.176**	2.081	**0.037**	-0.707	0.480	10.415	**0.000**	Moderate
1990	**0.764**	1.939	0.052	-0.167	0.867	7.482	**0.000**	Moderate
1992	**0.481**	1.216	0.224	-0.282	0.778	8.714	**0.000**	Moderate
1994	0.009	-1.468	0.142	-1.297	0.195	12.898	**0.000**	None
1996	0.012	-0.228	0.819	-0.681	0.496	10.624	**0.000**	None
1998	**0.456**	1.434	0.152	-0.763	0.445	10.755	**0.000**	Moderate
2000	**0.619**	1.784	0.074	-0.495	0.620	12.233	**0.000**	Moderate
2002	**0.499**	3.866	**0.000**	-0.643	0.520	12.204	**0.000**	Moderate
2004	0.000	1.409	0.159	-3.369	**0.000**	25.597	**0.000**	None
2006	0.004	-1.332	0.183	-2.256	0.024	21.074	**0.000**	None

The final column in both tables states the support for the power-law fit given the observed Vuong test statistics and p-values. In most cases, there is only moderate support for the power-law fit but no strong support, as the lognormal distribution provides a fit to the tail of the income and consumption distribution that is at least as good as the power-law fit in all cases. Nevertheless, we can rule out the truncated exponential distribution as a possible alternative fit to the Thai income and consumption distribution as the power-law distribution provides a better fit that is also highly statistically significant in all years

Power-Law Fit vs. Lognormal Fit to The Entire Distribution

In comparing the power law against the tail of the lognormal distribution that was fitted to the entire dataset, the mostly positive LRs from the column "Lognormal 1" show that the power law generally provides a better fit for both consumption and income data. However, only a few of these LRs are statistically significant at the 5% significant level, while most are not, which suggests that the null hypothesis that the lognormal tail provides an equally good fit as the power-law cannot be rejected. This is particularly true for the consumption distribution, where the LRs are statistically significant in only 2 out of 11 years. For income, the evidence for power law is stronger, with statistically significant positive LRs in 5 out of 11 years. Figures 4.12 and 4.13 show the graphical plot of the power-law fit and the lognormal fit to the entire complementary CDF of income and consumption respectively for the years 1986 to 2006. This confirmed the results stated in the column "Lognormal 1" of Tables 4.5 and 4.6.

Figure 4.12: Income: Power Law vs. Lognormal

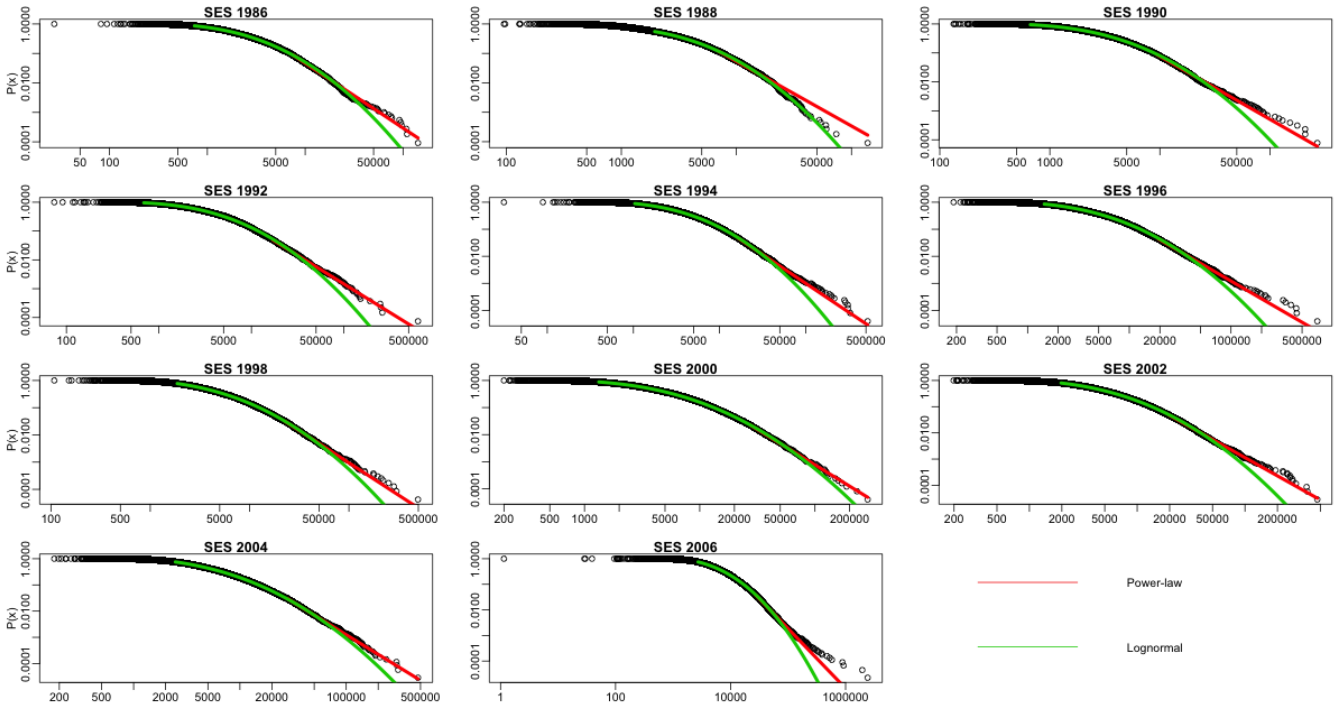
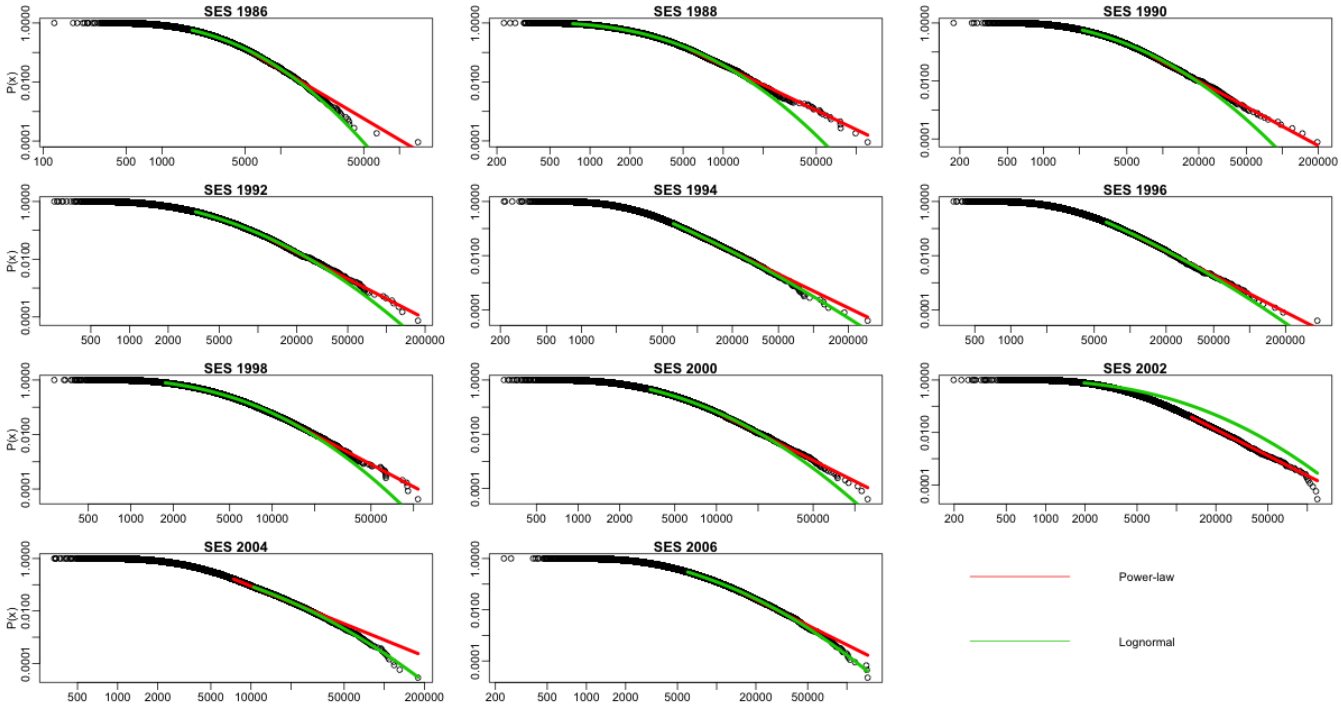


Figure 4.13: Consumption: Power Law vs. Lognormal



Power-Law Fit vs. Lognormal Fit to The Tail

In comparing the power law against the truncated lognormal distribution, the evidence for the power law is even weaker, with most of the LRs in the column "Lognormal 2" showing a negative sign, indicating that the truncated lognormal generally provides a better fit for both income and consumption distributions. But again only a few of these LRs are statistically significant, so the null hypothesis that both distributions fit the data equally well cannot be rejected. Figures 4.14 and 4.15 show the graphical plot comparing the fit of the power law against the truncated lognormal distribution for income and consumption. This confirmed the result in the column "Lognormal 2" of Tables 4.5 and 4.6.

Figure 4.14: Income: Power Law vs. Lognormal Tail

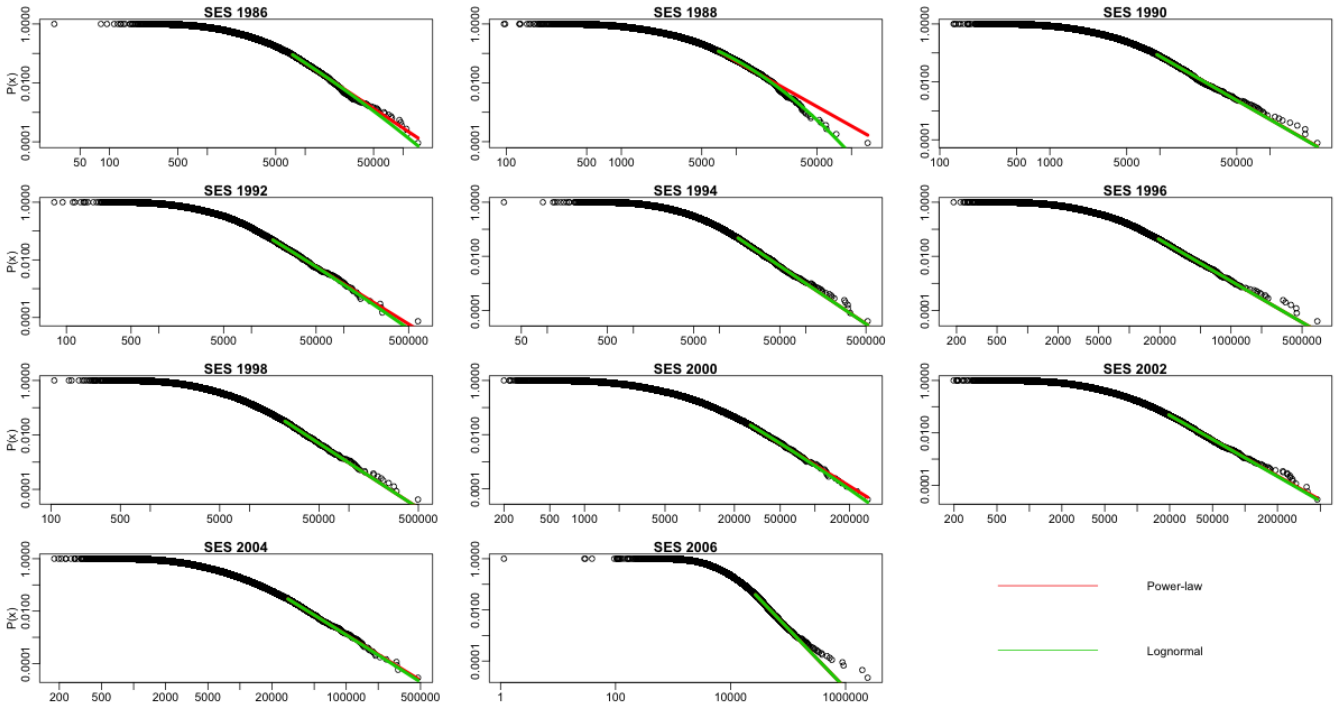
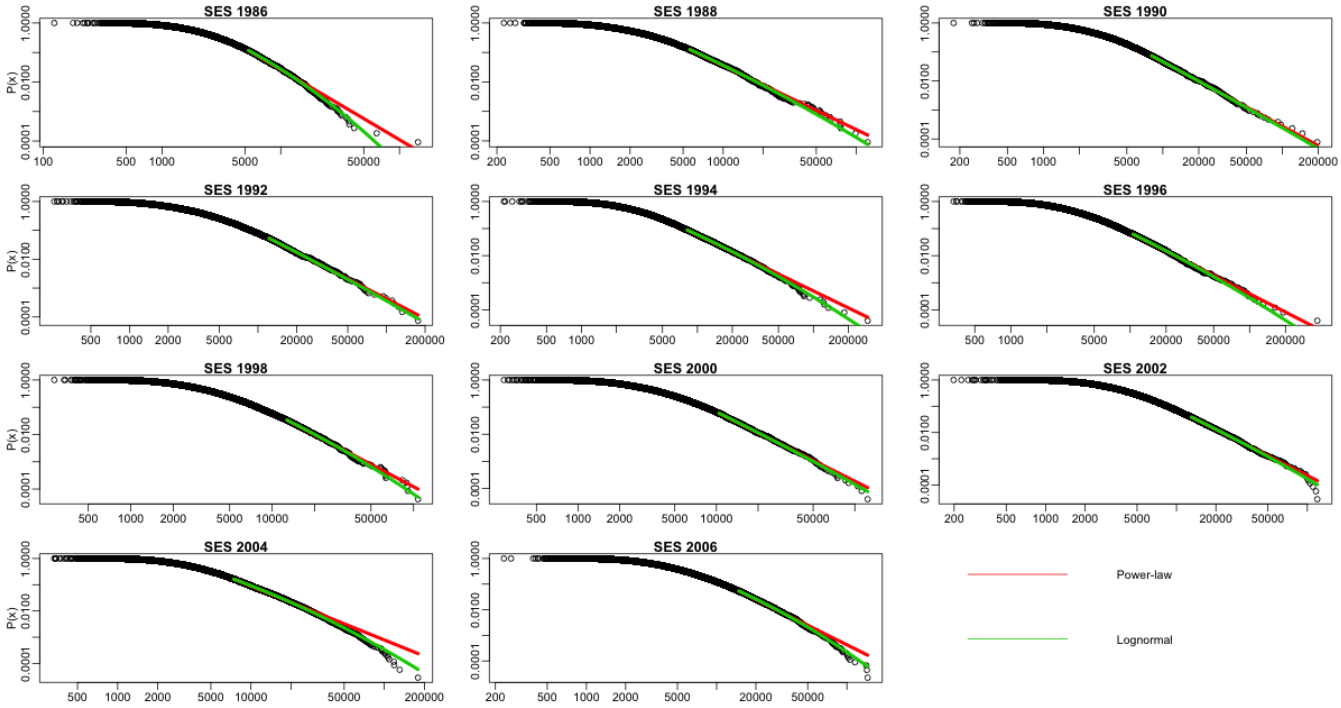


Figure 4.15: Consumption: Power Law vs. Lognormal Tail



4.5.1.3 Key Findings

There are two main conclusions that can be drawn from these results. First, using the OLS approach to fit power law to the data is a bad idea. While it is possible to obtain reasonably accurate parameter estimates, it will be hard to do so without employing a method that can credibly estimate the cut-off point x_{min}. If we had used visual inspection, a common method to pick x_{min} in the linear regression method, our estimate from the OLS approach could be very different. More importantly, the OLS approach will produce high R^2 for most heavy-tailed data, hence overstating the true extent of the goodness of fit of the power law to these data and leading to many false claims as seen in the literature (Brzezinski, 2014). In addition, the pure OLS approach will not allow for the differentiation between the power law and competing alternative distributions that could exhibit similar linear CDFs over a certain range. Therefore, going forward, the CNS approach used here should be the default starting point for studies that wish to identify power-law fit to any dataset.

Secondly, the results from this study have shown that there is only moderate support for power law in the Thai income and consumption distribution. This is not surprising as all estimated $\hat{\alpha} > 3$ would suggest that the best-fitted power-law in this case has a finite mean and variance, contrasting with the usual estimated range of $2 < \alpha < 3$ found in the literature. This finite second moment suggests that the tail of the Thai distribution is not as thick as usually observed in other countries, hence making it even more likely that the data could also be well-fitted by the lognormal distribution.

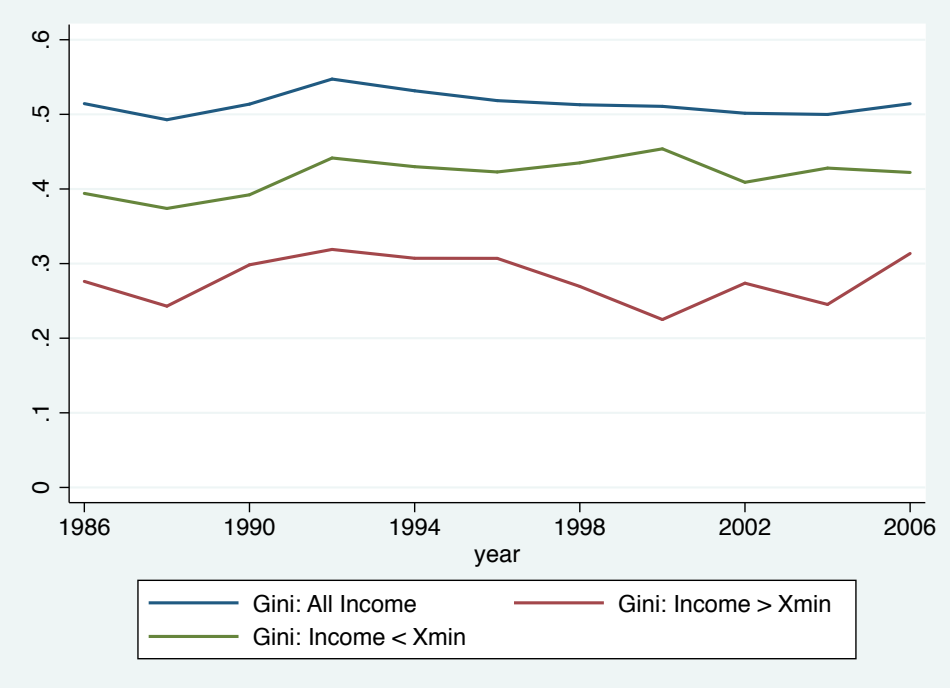
The findings here also suggest that inequality at the top of the income and consumption distribution is not that high. Using the formula that relates the relationship between α and the Gini coefficient, the $\hat{\alpha}$ estimates suggest that top income and consumption Gini coefficients are approximately 0.3 and 0.25 respectively.[9] These results seem to be in disagreement with observed high overall inequality and the conventional belief that there is a high degree of wealth concentration in Thailand. On these points, the explanation lies in a closer inspection of the Thai income and consumption distribution.

The Gini coefficient is a measure of inequality that is computed for the entire distribution, while the power-law exponent only measures inequality at the top end of the distribution. Figures 4.16 and 4.17 plot the Gini coefficient for the entire income and consumption distri-

[9] Gini coefficient $= \frac{1}{2\alpha - 1}$; refer to Appendix 4.A.2 for the derivation.

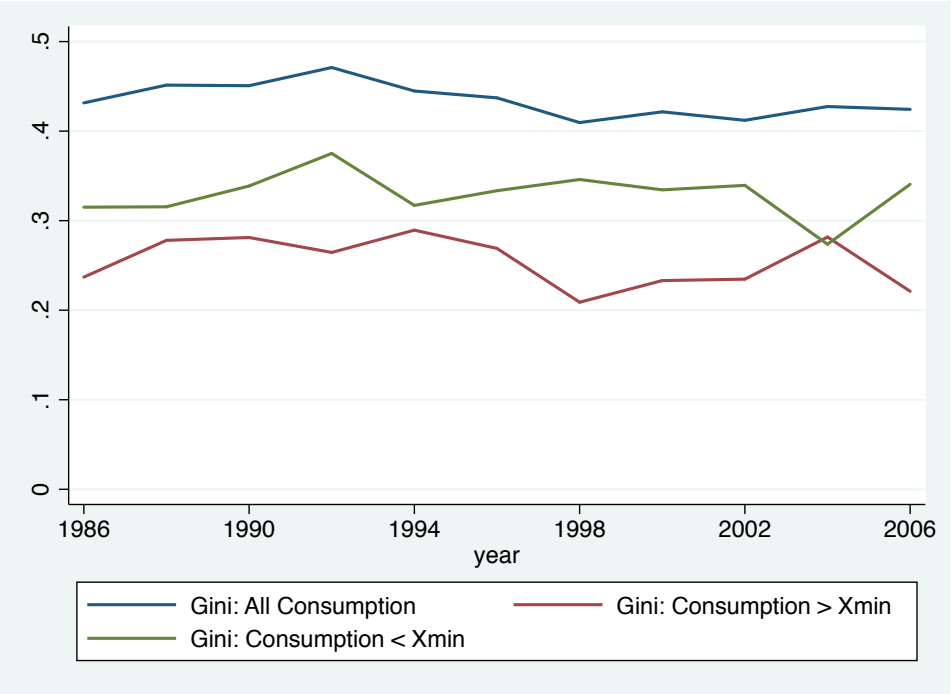
bution against the Gini coefficient for income and consumption above $\hat{x}_{min}$ and below $\hat{x}_{min}$ respectively.[10] Overall, there is greater inequality in income than consumption, consistent with earlier discussions and results. More interestingly, the Gini coefficient at the top end coincides with $\hat{\alpha}$, ranging approximately around 0.2 to 0.3. The magnitude of the top-end Gini coefficients is quite low and even lower than the Gini coefficients at the bottom end of the distribution, suggesting that there is less inequality among the rich than among the poor. Overall inequality captured by whole sample Gini coefficients are higher than both sub-populations' Gini. This means that while within-group inequality in income and consumption might not be that high in Thailand, between-group inequality between the top and the bottom is.

Figure 4.16: Gini: Income



[10]The Gini coefficients used in these plots are computed from the SES data, and is not the official Gini coefficient released by NESDB used earlier. There are some small discrepancies but the overall trend and magnitude are approximately consistent with the official data.

Figure 4.17: Gini: Consumption



Figures 4.18 and 4.19 plot the 90/10 and 50/10 percentile ratios of income and consumption respectively. Firstly, the figures show that the top half of the income distribution holds many more resources than the bottom half. The 50/10 ratio is approximately 2 to 3 times, while 90/10 is more than 10 times. Absolute inequality is lower in consumption than income, with the 50/10 ratio approximately 2 to 2.5 times, while 90/10 is 6–7 times. These estimates show that the major difference between the inequality measured by income and consumption observed lies at the top. While the 50/10 ratio of both income and consumption are comparable, the 90/10 ratio of income is much higher than that of consumption. While progressive taxation can account for some of the differences, the inequality gap observed could signal the underestimation of inequality in living standards as measured by consumption. This result is in contrast to those of (Meyer and Sullivan, 2013b), which find that most of the differences in the income and consumption inequality in the US are at the bottom end of the distribution.

Figure 4.18: Income: Inequality 1986-2006

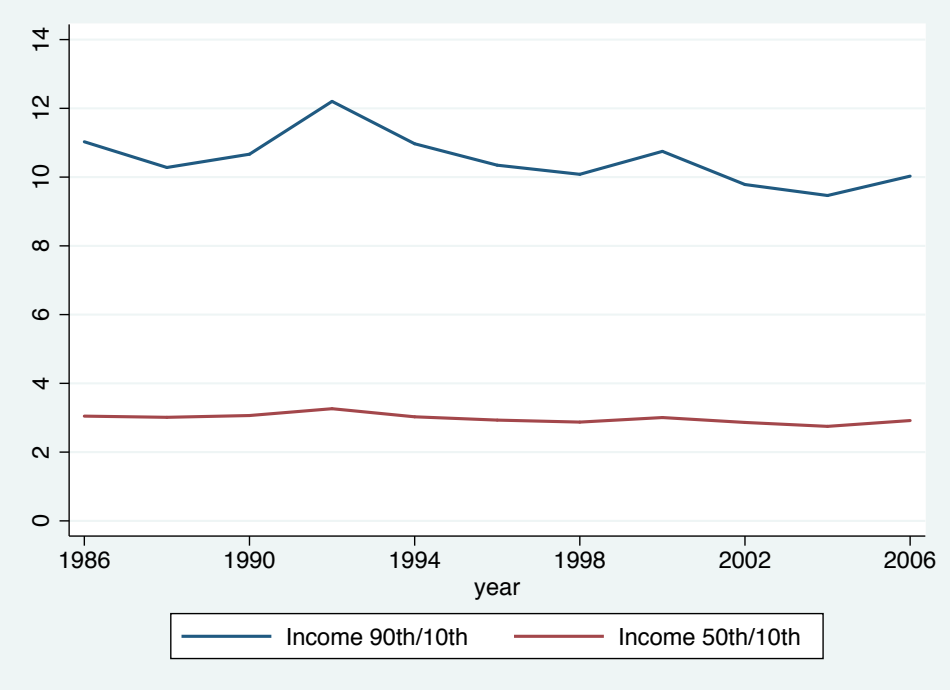
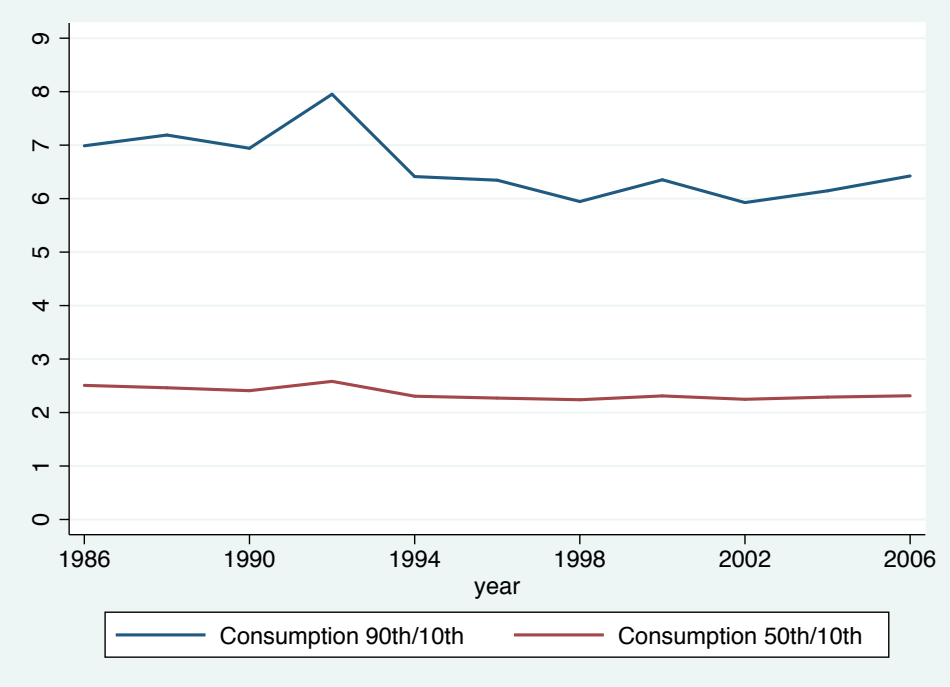


Figure 4.19: Consumption: Inequality 1986-2006



Secondly, Figures 4.18 and 4.19 provide the evidence in support of our conjecture that there is a high level of between-group inequality. One thing to note from these figures is how income and consumption inequality have moved closely together and that the 90/10 and 50/10 ratios have remained relatively stable over time. This reveals that over the past

20 years, while there has been growth in income and consumption, the 10th percentile of the population has not caught up to the median and the median has not caught up to the 90th percentile, keeping relative inequality approximately the same. This result again is in contrast with those observed in the US as reported by (Meyer and Sullivan, 2013a), which showed that consumption and income inequality does not always move together and that in fact, income inequality has risen noticeably higher than consumption inequality over the period of study.

While one cannot immediately draw a conclusion on income mobility, as we cannot really observe individual movement within the distribution from the data, the above statistics do provide a picture of a very static economy where between-group inequality is almost a permanent occurrence. This explains why we observed high values of $\hat{\alpha}$ for the top end at the same time as high Gini coefficients for the entire distribution. Thailand is therefore a society that is characterised by both high absolute and relative inequality.

Despite this grim conclusion it is important to note that, between 1986 and 2006, real per-capita income at the 10th, 50th and 90th percentiles grew by 260%, 249% and 236% respectively. This shows that while relative income inequality may not have improved much, the absolute increase in well-being for the Thai population has been quite significant. In the US, from 1966 to 2011, inflation-adjusted wage and salary income at the 50th and 90th percentiles increased by 11% and 58% respectively. At the 99th percentile, the increase was 121%, while at the 99.99% the increase was 617% (Dew-Becker and Gordon, 2005). Gordon (2009) also finds that income inequality in the bottom 99% of the US has not changed much since the early 1990s and most of the recent rise in inequality observed came from the increasing disparity of income at the top 1% and the rest of the distribution. In a way, this finding is similar to ours that within-group inequality has been stable at the bottom end of the distribution. The major difference is the absence of an explosive growth of income at the very top as that observed in the US, which widened the between-group inequality between the top and the bottom sufficiently to cause the rise in overall inequality. Real income growth at the 99th percentile in Thailand is only 240%.

In conclusion, the findings from this study suggest that the evidence in support of the power-law fit to income and consumption is not strong. Given the lower level of inequality at the top than anticipated, the top ends of both distributions are less heavy and can be

fitted as well by the lognormal distribution. In the next subsection, we attempt to test the lognormal fit to both income and consumption distribution further to assess the validity of Gibrat's law of proportional growth as the generative mechanism for these distributions.

4.5.2 Lognormal: Income vs. Consumption

This section presents and discusses the lognormal fit to the Thai income and consumption distribution using the method of Battistin et al. (2009) discussed in Section 4.3.4. The income and consumption data from years 1986 to 2006 are pooled together and then grouped in to four ten-year-of-birth cohorts: 1960–69, 1950–59, 1940–49, and 1930–39. The data are pooled and grouped into cohorts because the model of proportional random growth in income and consumption should apply to individuals in the sample who are "sufficiently similar in a sense of having similar unconditional moment μ and σ^2" and are going through the same phase of their life-cycle.

Recall from Section 4.2.1.1 that if an individual's permanent income follows Gibrat's law of proportional random growth as the income-generation process, then that individual income scaled by age should be lognormally distributed. Assuming that a strong form of ergodicity holds, the distribution of permanent income for individuals in the same cohort should also be lognormal. Therefore, the income-generation model can be tested by examining whether the shape of cohorts' permanent income distribution is close to the lognormal distribution.

The summary of the cohort data is given in Table 4.7.

Table 4.7: Age Group

	1986-1990	1992-1996	1998-2002	2004-2006
1960-69	21-25	27-31	33-37	39-41
1950-59	31-35	37-41	43-47	49-51
1940-49	41-45	47-51	53-57	59-61
1930-39	51-55	57-61	63-67	69-71

In this section, we will discuss the result of the youngest cohort (1960–69) in detail. Given the similarity of the findings, we will leave the results for the remaining cohort in Appendix 4.A.

4.5.2.1 Cohort 1960-69

This section presents the lognormal fit to the income and consumption of the 1960–69 cohort from years 1986 to 2006.

In their work, Battistin et al. (2009) find that the null hypothesis of lognormal distribution being a good fit to the data cannot be rejected for consumption, while it can be rejected for income in most cohort-year observations. They also find that as the cohorts age, the

deviations in both income and consumption become more dispersed and the distributions move closer to lognormal. These findings are consistent with Gibrat's proportional random growth model. As older households receive more random shocks, the dispersion in their permanent income will increase; at the same time more instances of random shocks allow the CLT to set in and move the distribution closer to the lognormal distribution. Therefore, the authors found strong support for the Gibrat process, especially for consumption, a proxy of permanent income, in the UK and the US.

However, results from Figures 4.20 and 4.21 show that, in the case of Thailand, the null hypothesis of the income and consumption data following the lognormal distribution can be strongly rejected, as the p-values are 0.000 in all the four age groups. This is also true for the three remaining cohorts in all their age groups. A closer look at the QQ plots reveals that the rejections of the lognormal fit mostly result from the top-end of the distribution, where observed incomes are larger than the predicted incomes from theoretical quantiles of the normal distribution. While this could have arisen because of the possible misreporting by individuals with high income, the results also provide evidence that the income distribution may be better represented by distributions with a thicker tail[11].

In addition, we do not observe any particular trends in the dispersion of log income and consumption as the cohorts age. In fact, the opposite seems to be the case, where dispersion lessens as the cohorts get older. However, this could be the result of the fact that we have more observations in later years of the data, which means that log consumption and log income can be estimated more precisely. It is also not possible from the current results to tell whether the departure from the lognormal is narrowing as the cohorts aged. This is because the p-values are 0.000 for most of the tests. We will come back to this point in Subsection 4.8 when we discuss KL divergence.

The only consistent development throughout these results is that consumption is less volatile than income. The standard deviation of log consumption is lower than that of income in all four periods for all four cohorts. This is consistent with the finding from the power-law study in the previous section.

[11] Alternative to the survey non-response issue discussed earlier which led to the under-reporting of income at the top-end, another data quality issue may be the intentional over-reporting by certain individuals at the top-end. Given the difficulty of reaching top income earners and their usual unwillingness to respond to extensive surveys, the remaining sample of those who reported high income may consist of individuals with unusual characteristics that have the tendency to over-report. However, their existence is hard to quantify and should not represent a significant proportion of the sample as to cause meaningful bias to the result in any one year and over time.

It is important to distinguish between the lognormal fit that we are performing here versus the previous section. Here we are fitting the lognormal distribution to the entire income and consumption distribution, while in the previous section we only concentrated on the top-end of the distribution. In any case, the results reported here do not provide a strong support for the lognormal fit to income or even consumption, a proxy for permanent income. This points to the possibility that Gibrat's law of proportional random growth may not be the appropriate income-generation process for Thailand on an entire distribution basis.

4.5.2.2 Kullback-Liebler Divergence

As discussed in Section 4.2.1.1, the claim made by Battistin et al. (2009) is merely that consumption, a proxy of permanent income, should have a distribution that resembles the lognormal distribution more than income. Therefore, while neither income nor consumption follows the lognormal distribution perfectly, as shown in the results of the previous subsection, it will be sufficient to show that consumption is closer to lognormal than income in order to verify the validity of Gibrat's law as an income-generation process in Thailand.

Table 4.8 shows the result of the performing the KL divergence calculation.

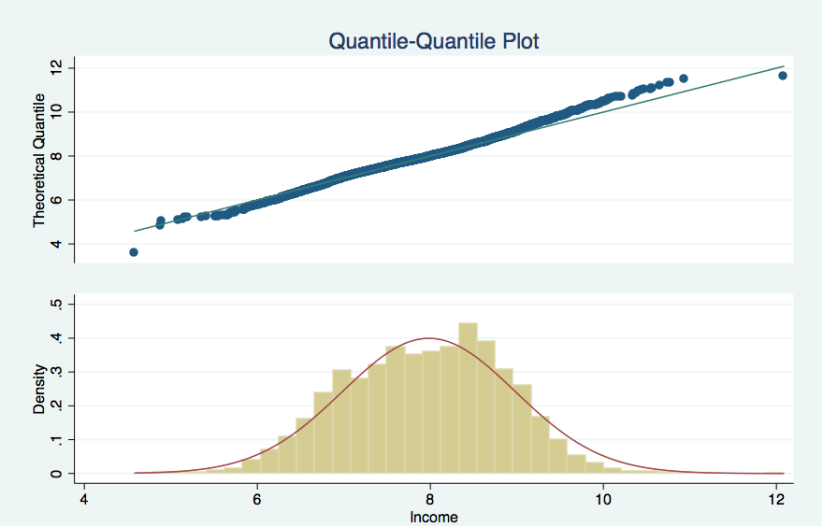
Table 4.8: KL Divergence: Income vs. Consumption

Cohort	Year	Age	N	KL Income	KL Consumption	Closer
1960 -1969	1986 -1990	21 -25	4,355	**0.01632**	0.01017	Consumption
1960 -1969	1992 -1996	27 -31	12,050	0.00813	**0.01572**	Income
1960 -1969	1998 -2002	33 -37	18,701	0.00666	**0.01751**	Income
1960 -1969	2004 -2006	39 -41	19,326	0.00558	**0.01678**	Income
1950 -1959	1986 -1990	31 -35	8,902	0.00914	**0.01449**	Income
1950 -1959	1992 -1996	37 -41	16,066	0.00718	**0.01856**	Income
1950 -1959	1998 -2002	43 -47	21,180	0.00931	**0.01748**	Income
1950 -1959	2004 -2006	49 -51	19,753	0.01557	**0.02246**	Income
1940 -1949	1986 -1990	41 -45	7,450	0.00952	**0.01341**	Income
1940 -1949	1992 -1996	47 -51	12,892	0.01531	**0.02243**	Income
1940 -1949	1998 -2002	53 -57	15,097	0.02288	**0.02624**	Income
1940 -1949	2004 -2006	59 -61	13,244	0.02309	**0.03470**	Income
1930 -1939	1986 -1990	51 -55	6,730	0.01064	**0.01891**	Income
1930 -1939	1992 -1996	57 -61	10,882	0.01940	**0.02692**	Income
1930 -1939	1998 -2002	63 -67	11,750	0.01560	**0.02187**	Income
1930 -1939	2004 -2006	69 -71	8,958	0.01805	**0.02629**	Income

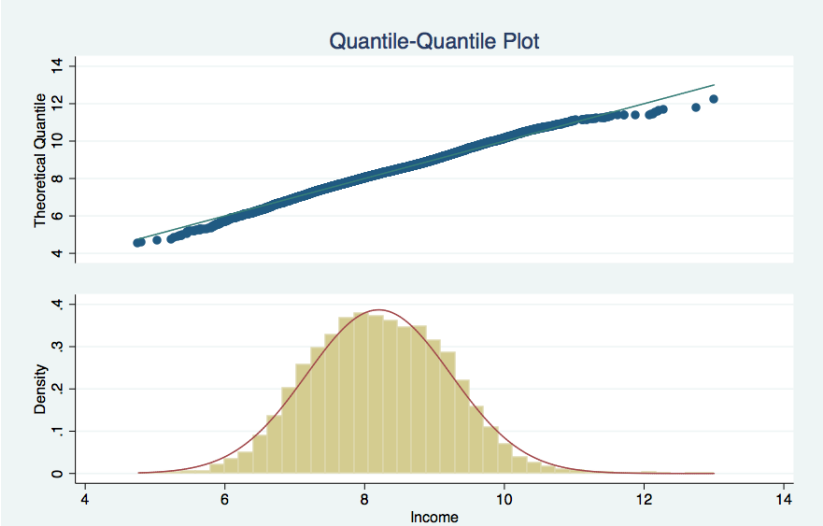
The results are in contrast with those of Battistin et al. The distribution of consumption, which is supposed to be a proxy for permanent income, is found to be further away from the lognormal distribution than reported income.

Figure 4.20: Log Income: Cohort 1960-69

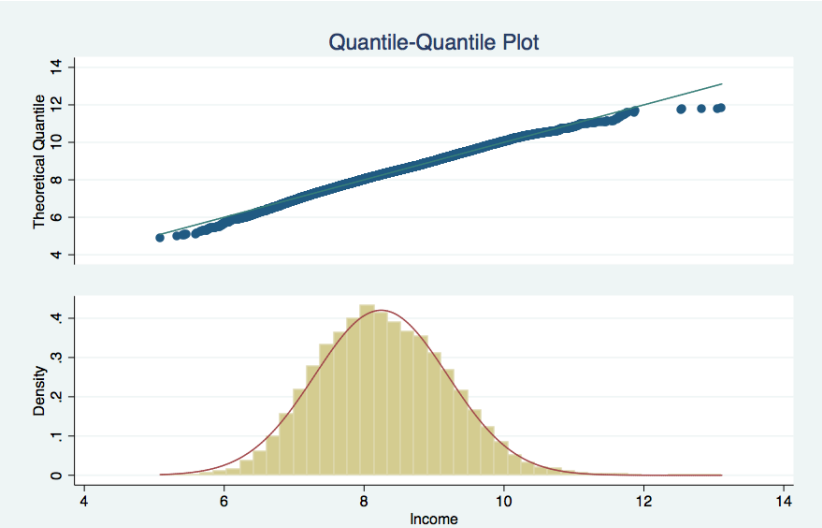
(a) Age 21-25 (Year 1986-1990)



(b) Age 27-31 (Year 1992-1996)



(c) Age 33-37 (Year 1998-2002)



(d) Age 39-41 (Year 2004-2006)

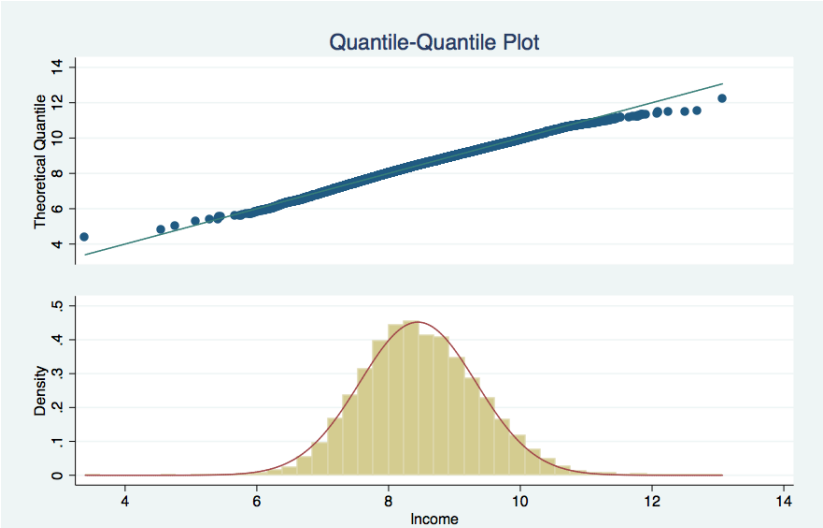
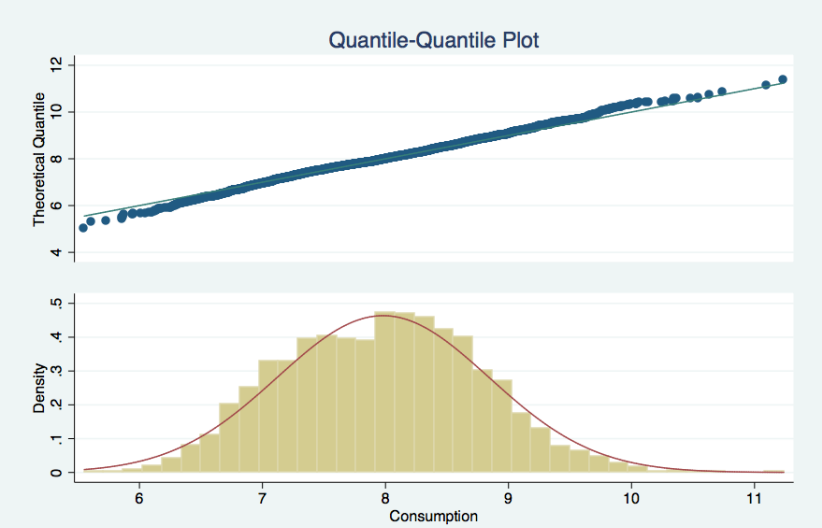
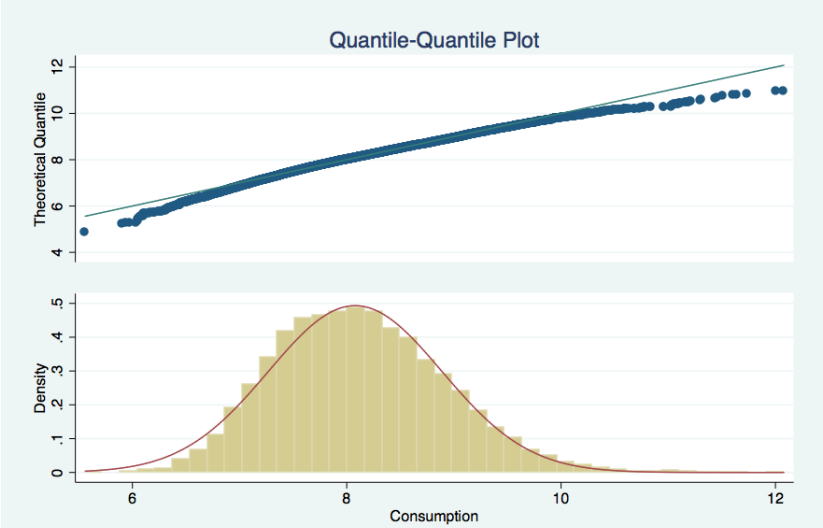


Figure 4.21: Log Consumption: Cohort 1960-69

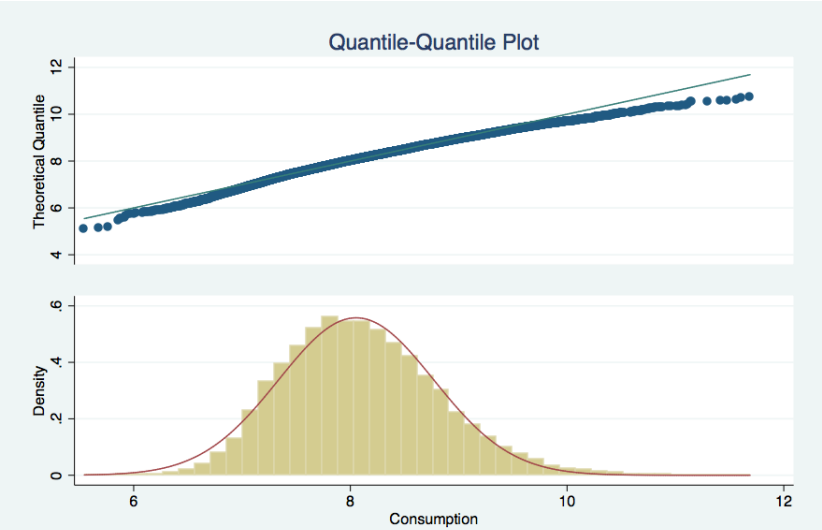
(a) Age 21-25 (Year 1986-1990)



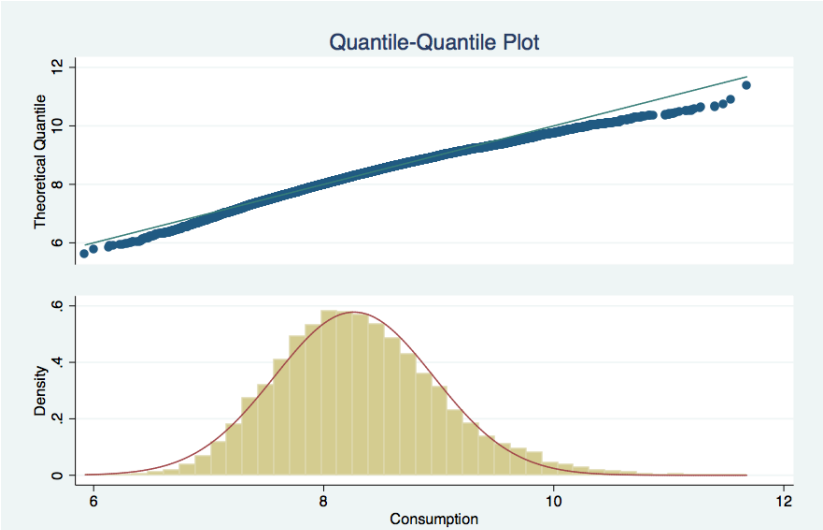
(b) Age 27-31 (Year 1992-1996)



(c) Age 33-37 (Year 1998-2002)



(d) Age 39-41 (Year 2004-2006)



One possible explanation could be that the transitory component in the Thai income-generation function is not great, and therefore it is the reported income itself that represents permanent income, which follows Gibrat's law of proportional random growth. In this case, it will be income that will follow the lognormal distribution. If income does follow the lognormal distribution, then a process of consumption smoothing, which does take place in Thailand as evidenced by the smaller inequality observed in consumption than income over time, will result in the consumption distribution deviating further from the lognormal. Therefore, if we believe that Gibrat's law is the underlying process that explains the distribution of income in Thailand, then one must find an alternative process to explain consumption generation

However, it is more likely the case that Gibrat's law offers a poor explanation as an income-generation process in Thailand. Results from Table 4.8 show that as the cohorts get older, there are no clear patterns of convergence to lognormal as predicted by Gibrat's process. Combining this with the results in the previous subsection, one can conclude that the evidence in favour of Gibrat's law as the definitive income-generation process in Thailand is weak. While the results shown here are in contrast with those observed in the UK and US, they are similar to findings in Italy, which also reject the lognormal fit to consumption and reject Gibrat's law as the income-generation process (Fagiolo et al., 2008).

4.6 Conclusions

In this chapter we attempt to fit the power-law distribution to the income and consumption distribution and study the dynamics of inequality at the top end of both distributions in Thailand. Given Thailand's rapid development from a low-income country to an upper-middle income country during the past 30 years, its income distribution provides an interesting case study of a distribution that could exhibit power-law behaviour with high inequality and heavy concentration of income at the top. Discovery of power law could also provide an indication that some interesting and complex income-generation mechanism may be at work in Thailand.

There are three main findings in this study. First, we have provided additional evidence to the existing empirical econo-physics literature and further established the shortcomings of using the popular but incorrect method based on the linear regression approach to fit power law to any distribution. Using the linear regression approach, we find extremely strong evidence of power-law fit in Thailand, with R^2 exceeding 95% for income and consumption data in all the years. However, on the implementation of a more complete and stringent method of distribution fitting proposed by Clauset et al. (2009), the evidence in support of the power-law fit is much weaker. Only slightly more than half of the data pass the power-law goodness-of-fit test on an absolute basis. Comparing the power-law fit with several competing distributions shows that the lognormal distribution offers an equally good fit to all datasets studied here. Thus, we only find moderate support for the power-law fit for Thailand. We also illustrated the superiority of the CNS approach, which, going forward, should be the default starting point replacing the linear regression method for similar studies of this type.

Second, contrary to the prior hypothesis that there may be high income inequality and concentration at the top in Thailand, our estimates of the power-law exponents for both income and consumption are greater than 3, larger than what is usually observed in the literature ($2 < \alpha < 3$). This means that the tails of both distributions are less thick and indicates that there is less inequality at the top end of the income and consumption distribution in Thailand than anticipated. However, upon further inspection we find that while inequality at the top in Thailand is relatively low, there is a high level of persistent between-group inequality between the top and bottom ends of the distribution. Over the past

20 years, while there has been considerable growth in income and consumption, incomes at the bottom end of the distribution have not caught up to those at the top end at all, keeping relative inequality approximately the same. This helps explain and reconcile the observation of high α (high α signals lower inequality) but persistently high overall inequality in Thailand.

Thirdly, we attempt to investigate the plausibility of one possible underlying income-generating process, Gibrat's law of proportional random growth, and its application to Thailand. In contrast to the result obtained by Battistin et al. (2009) for the UK and the US, the lognormal distribution does not provide a good fit to the consumption distribution in Thailand. Furthermore, employing the KL divergence, we find that consumption, a proxy for permanent income that should follow Gibrat's law, has a distribution that deviates even further from the lognormal than reported income. In addition, we do not observe a pattern of convergence to lognormal as cohorts age and experience more random shocks. Given that the pattern observed in the data does not follow the predictions of Gibrat's law in most instances, there is little support for Gibrat's law of proportional random growth as the income-generating process in Thailand.

While the findings in this chapter have shed some light on the nature and the dynamics of inequality at the top end of the income and consumption distribution in Thailand, there are still some missing pieces. One of the key things missing is an investigation into the accumulated wealth distribution of the Thai population, which is likely to be even more unequal and concentrated than the income or consumption distribution seen here. One major policy implication of observing heavy wealth concentration is the strong case it provides for further progressive taxation. If wealth is really distributed according to the famous Pareto 80-20 rule, then marginally taxing the extremely rich individuals will have a significant impact on increasing the government budget and reducing inequality. The recently proposed tax reforms in Thailand, which include the introduction of property and inheritance tax, demonstrate the increasing importance of this issue. The increasing availability of wealth data going forward, as future rounds of the SES survey start to collect more detailed data on household assets, will provide an opportunity for further research on this issue of inequality in Thailand.

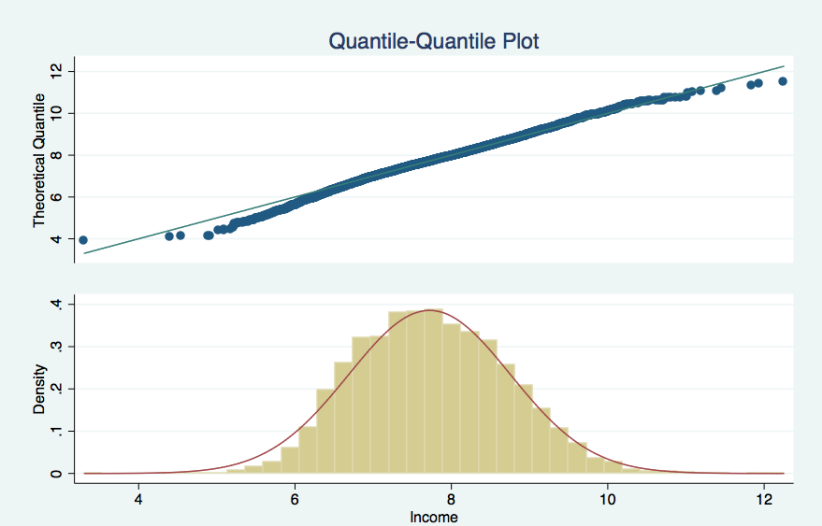
4.A Appendix

4.A.1 Lognormal: Income vs. Consumption Results

This section presents the lognormal fit to the income and consumption distribution of the 1950–59, 1940-49 and 1930-39 cohorts from 1986 to 2006.

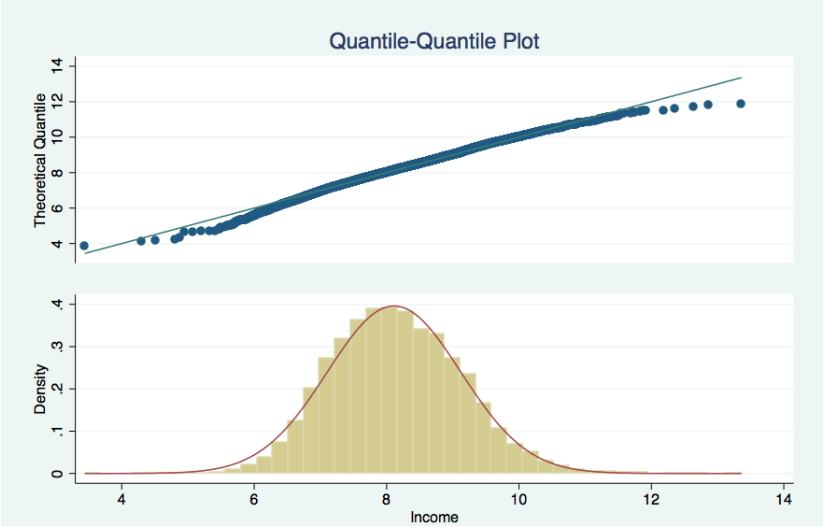
Figure 4.A.1: Log Income: Cohort 1950-69

(a) Age 31-35 (Year 1986-1990)



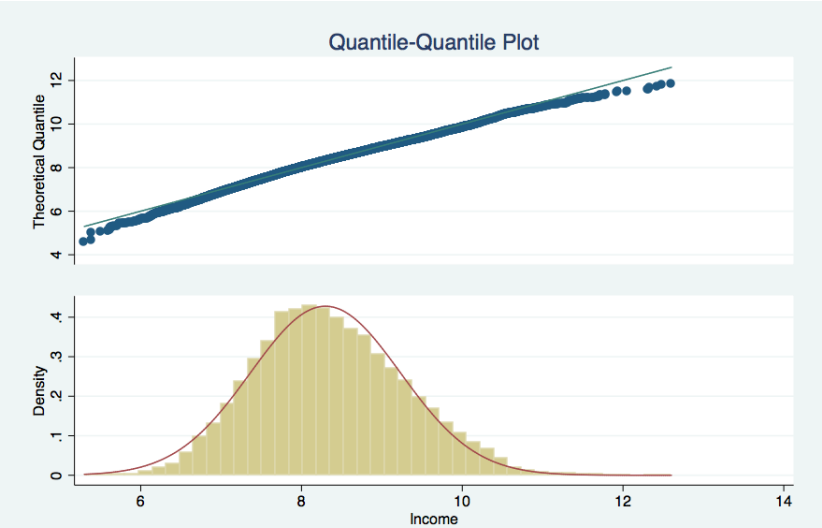
Statistics: Standard Deviation of Log = 1.0343; Skewness = 0.0239; Kurtosis = 0.0877
P-value: Standard Deviation of Log = 0.000; Skewness = 0.096; Kurtosis = 0.000

(b) Age 37-41 (Year 1992-1996)



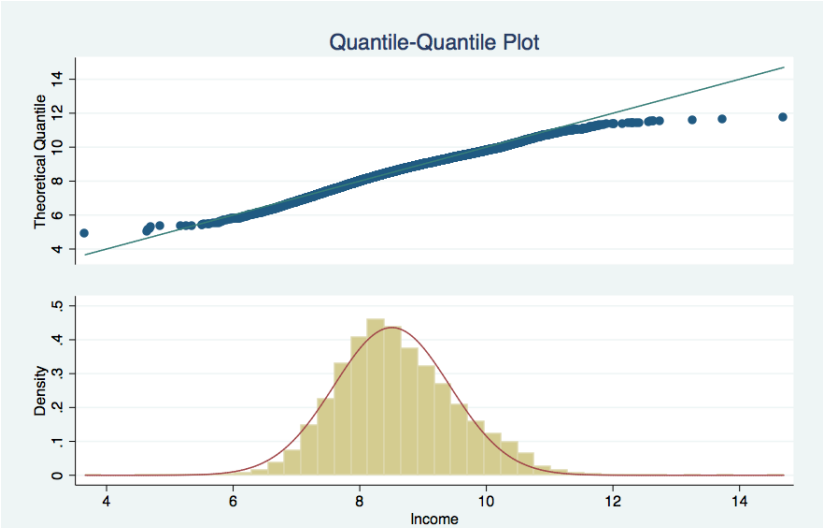
Statistics: Standard Deviation of Log = 1.0070; Skewness = 0.0322; Kurtosis = 0.0715
P-value: Standard Deviation of Log = 0.000; Skewness = 0.004; Kurtosis = 0.000

(c) Age 43-47 (Year 1998-2002)



Statistics: Standard Deviation of Log = 0.9327; Skewness = 0.0640; Kurtosis = 0.0012
P-value: Standard Deviation of Log = 0.000; Skewness = 0.000; Kurtosis = 0.914

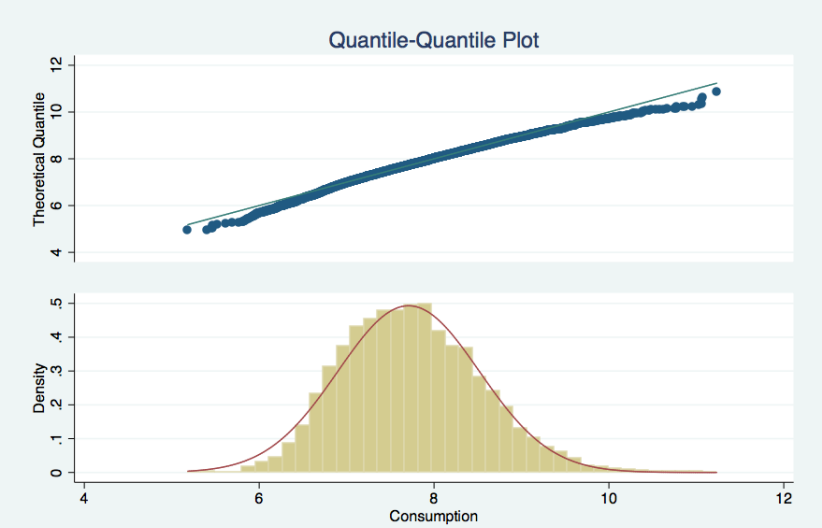
(d) Age 49-51 (Year 2004-2006)



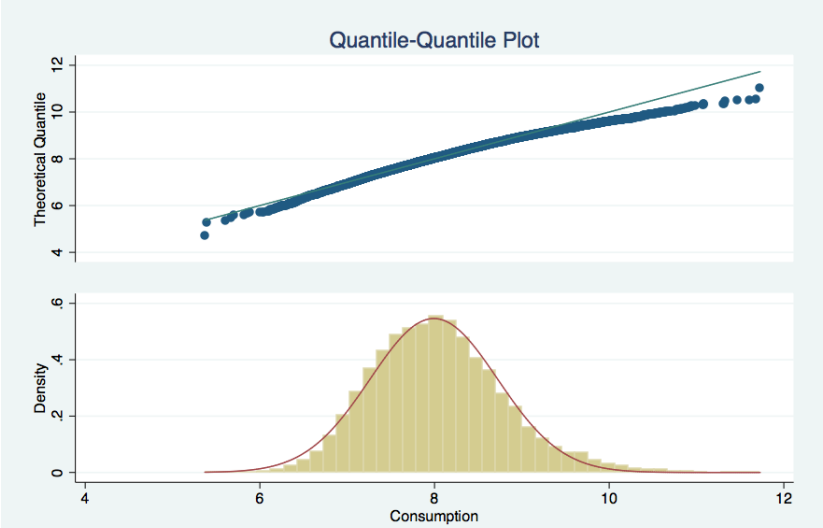
Statistics: Standard Deviation of Log = 0.9148; Skewness = 0.1054; Kurtosis = 0.0297
P-value: Standard Deviation of Log = 0.000; Skewness = 0.000; Kurtosis = 0.024

Figure 4.A.2: Log Consumption: Cohort 1950-69

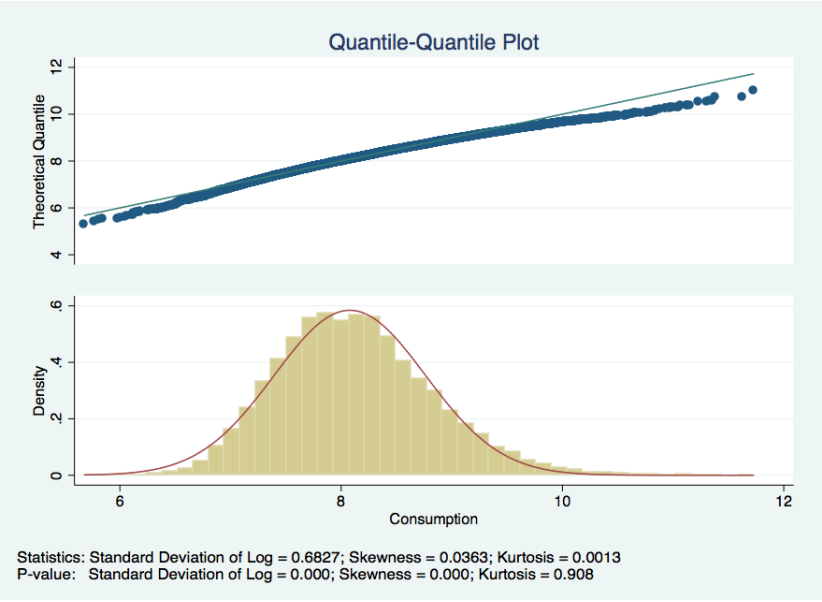
(a) Age 31-35 (Year 1986-1990)



(b) Age 37-41 (Year 1992-1996)



(c) Age 43-47 (Year 1998-2002)



(d) Age 49-51 (Year 2004-2006)

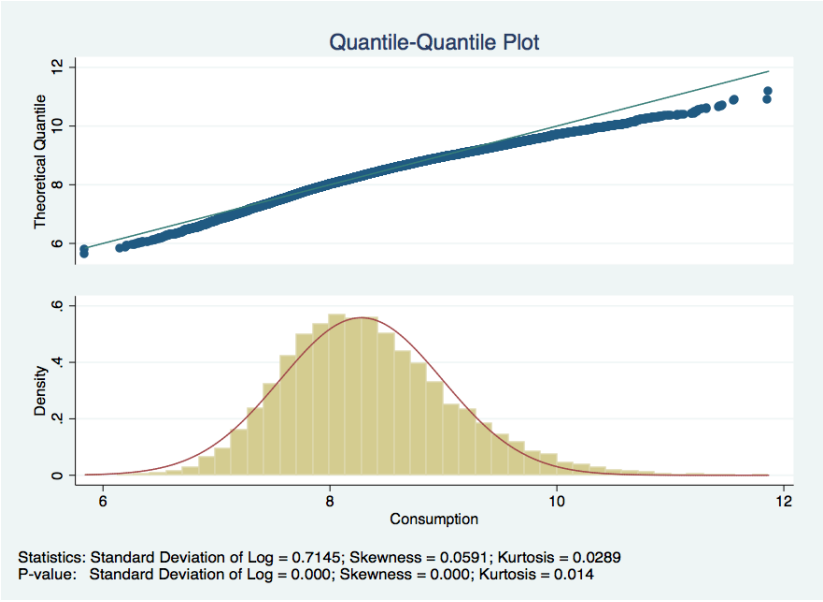
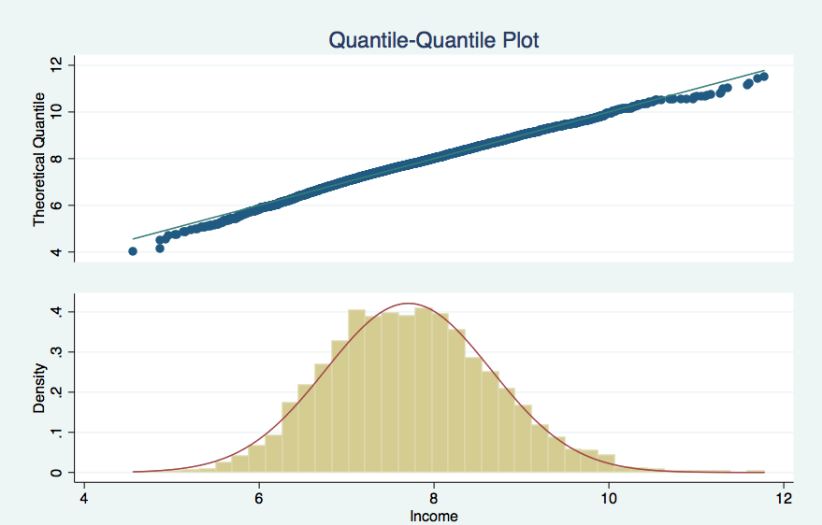
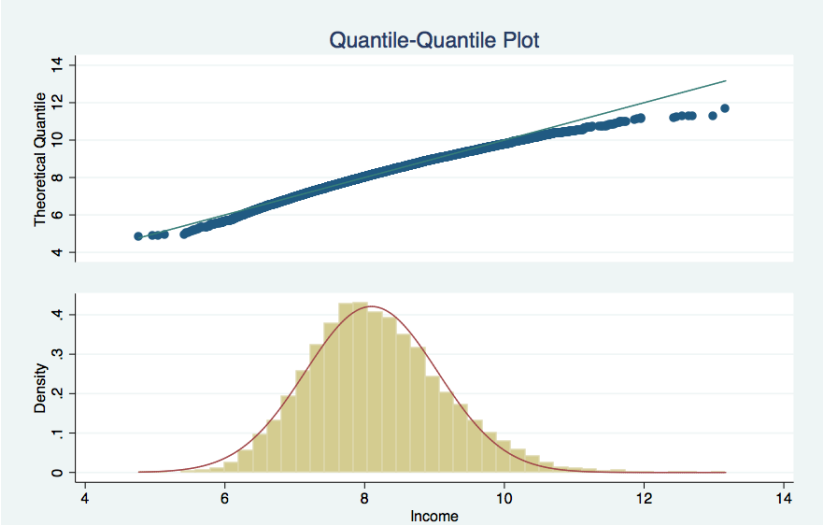


Figure 4.A.3: Log Income: Cohort 1940-49

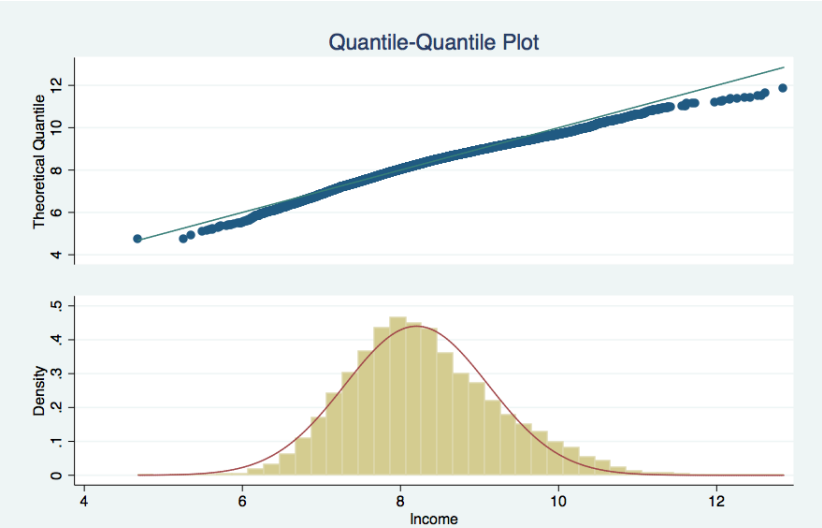
(a) Age 41-45 (Year 1986-1990)



(b) Age 47-51 (Year 1992-1996)



(c) Age 53-57 (Year 1998-2002)



(d) Age 59-61 (Year 2004-2006)

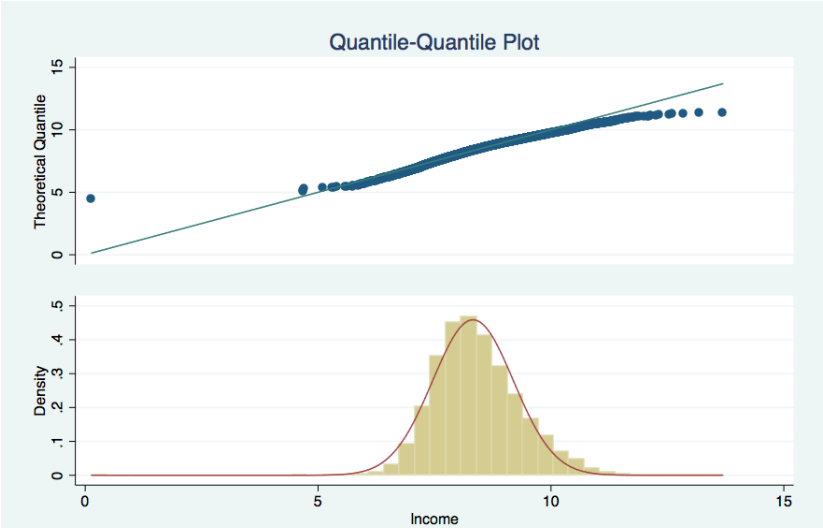
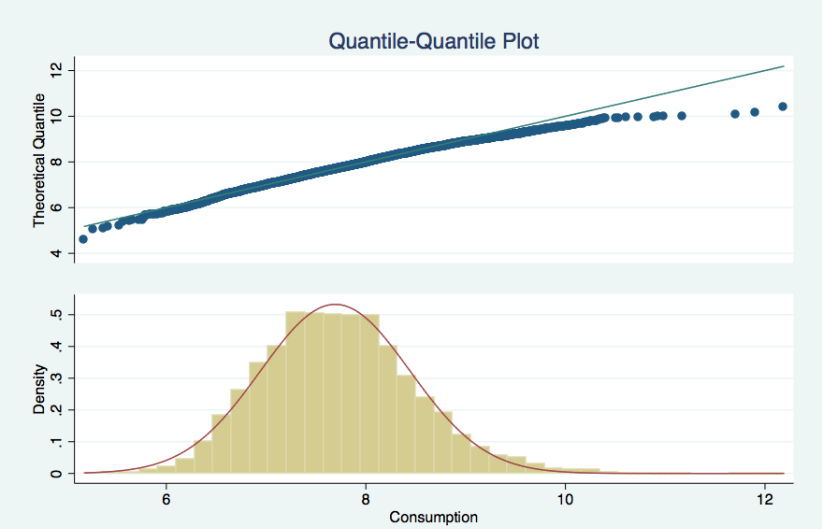
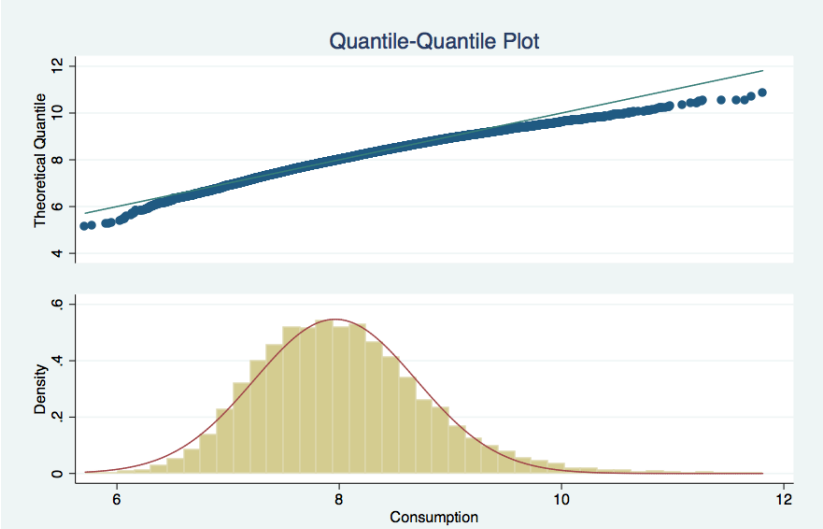


Figure 4.A.4: Log Consumption: Cohort 1940-49

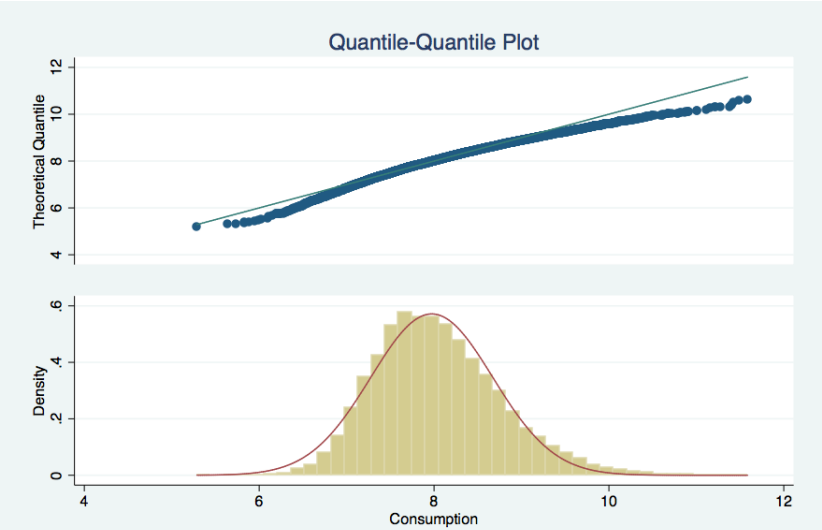
(a) Age 41-45 (Year 1986-1990)



(b) Age 47-51 (Year 1992-1996)



(c) Age 53-57 (Year 1998-2002)



(d) Age 59-61 (Year 2004-2006)

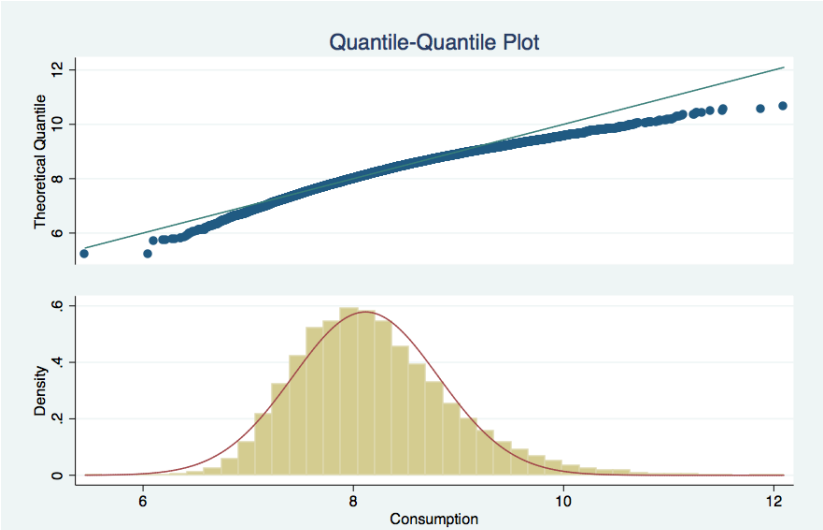
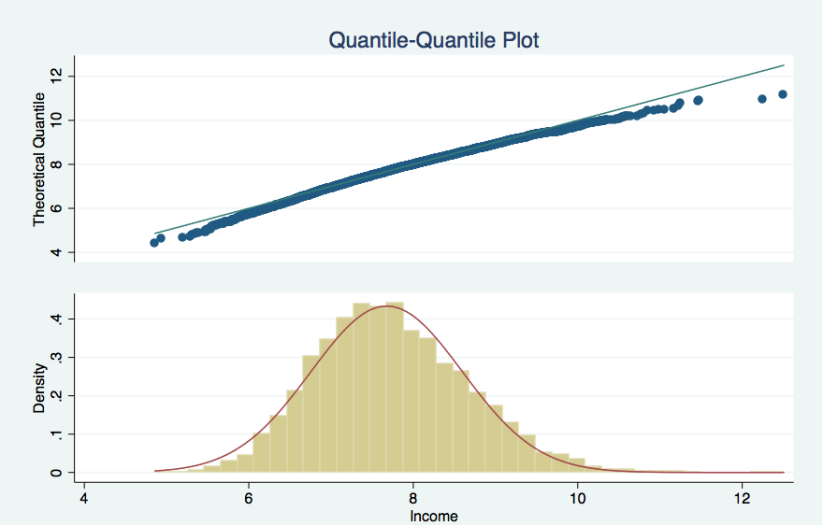
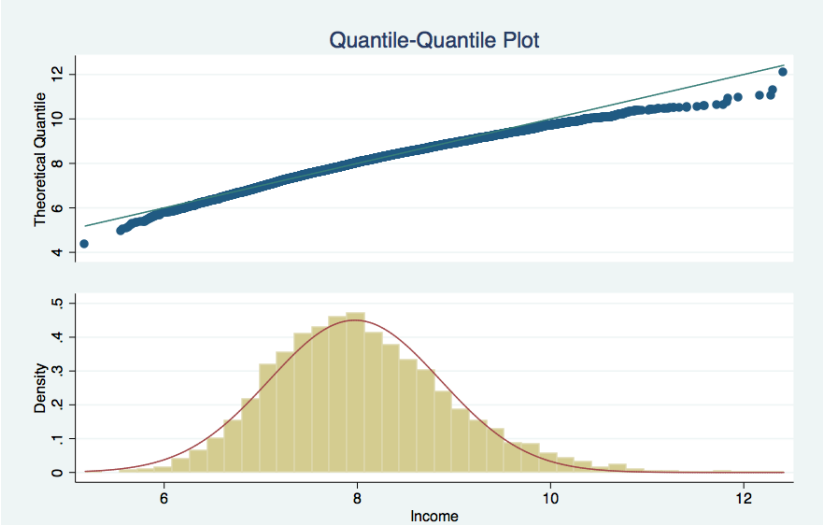


Figure 4.A.5: Log Income: Cohort 1930-39

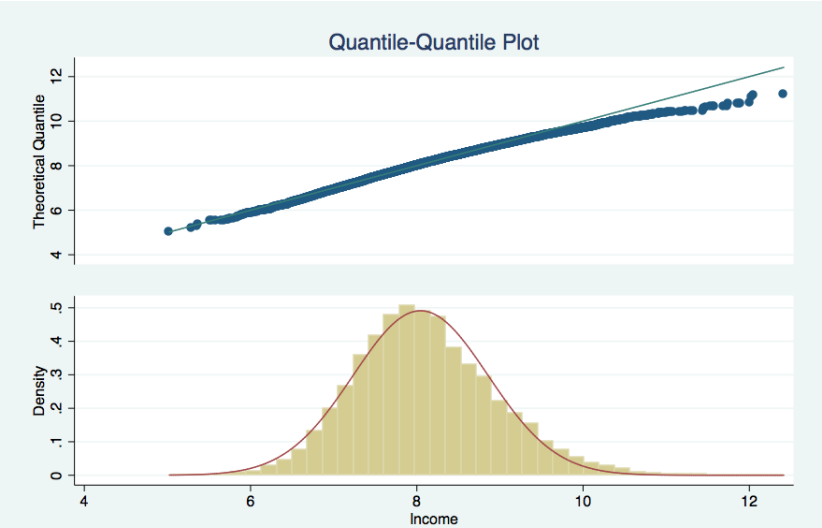
(a) Age 51-55 (Year 1986-1990)



(b) Age 57-61 (Year 1992-1996)



(c) Age 63-67 (Year 1998-2002)



(d) Age 69-71 (Year 2004-2006)

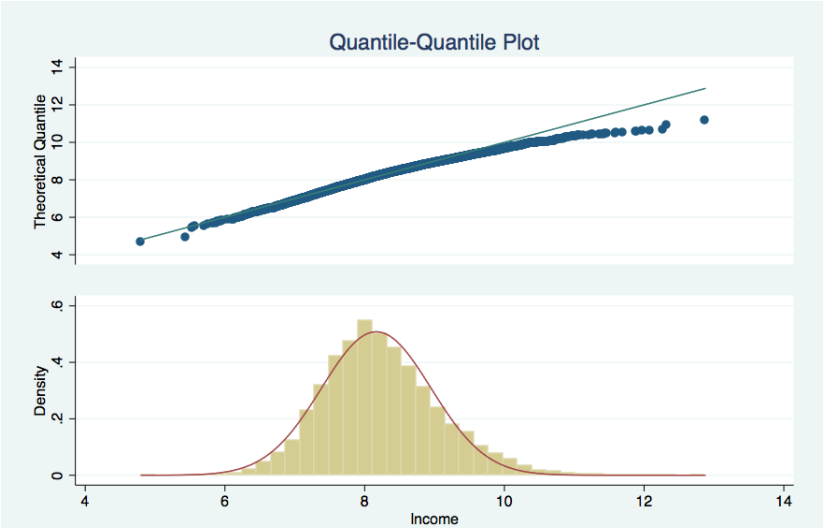
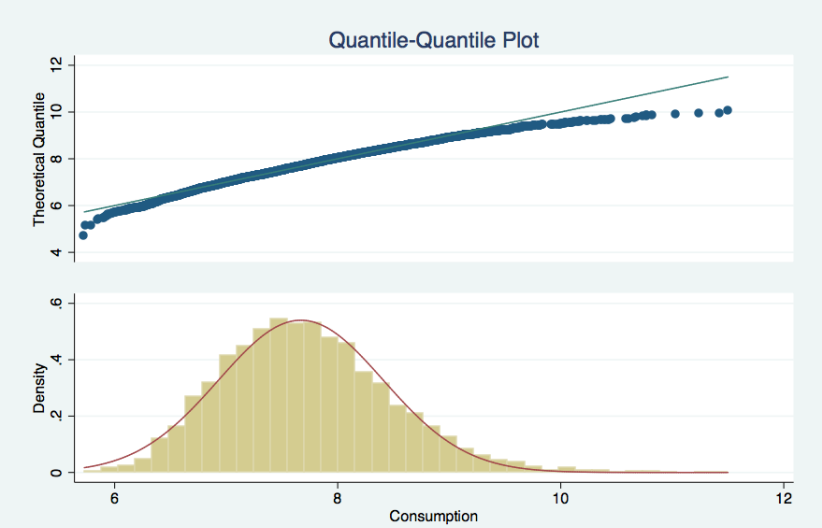
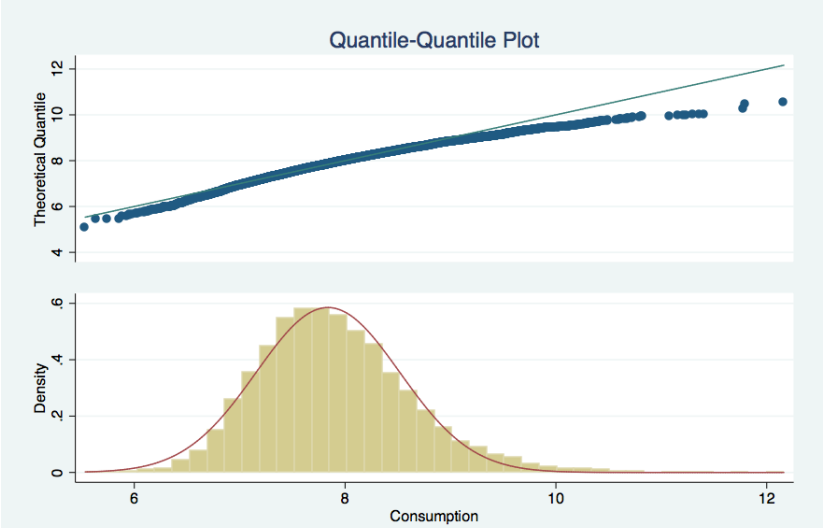


Figure 4.A.6: Log Consumption: Cohort 1930-39

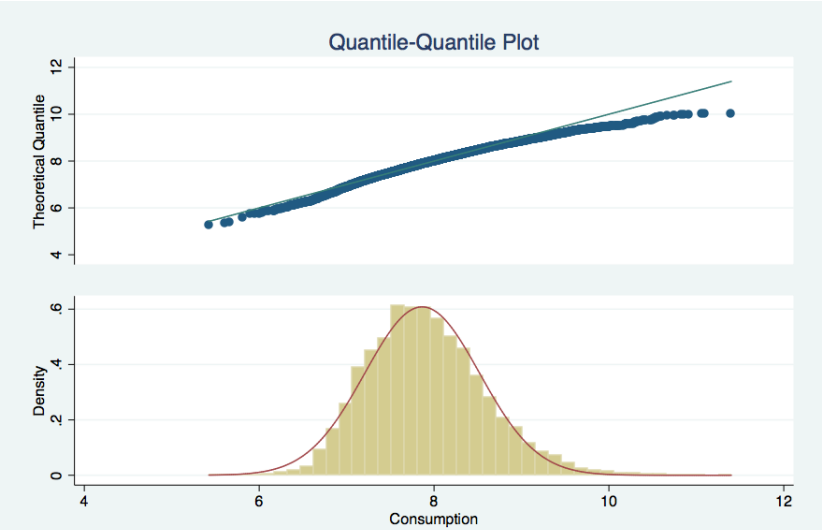
(a) Age 51-55 (Year 1986-1990)



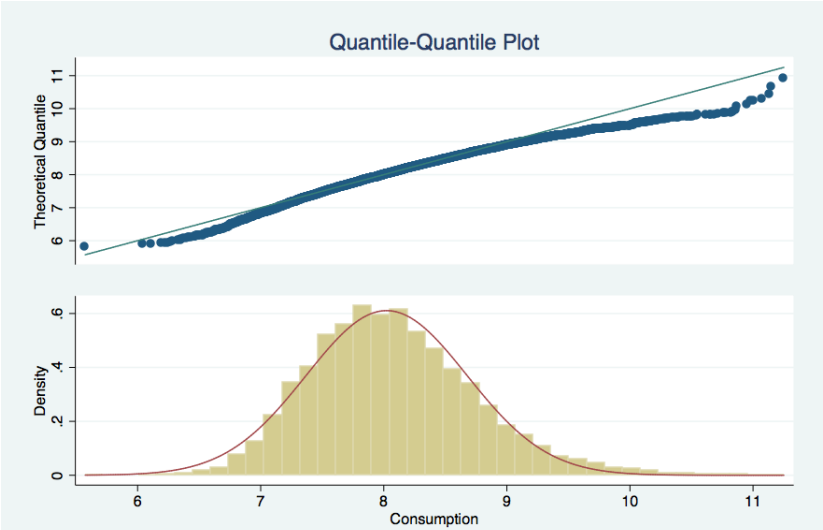
(b) Age 57-61 (Year 1992-1996)



(c) Age 63-67 (Year 1998-2002)



(d) Age 69-71 (Year 2004-2006)



4.A.2 Gini Coefficient and Power-law Exponent

The derivation of the relationship between the Gini coefficient and the power-law exponent α show here follow the setup of Aaberge (2005).

The Lorenz curve is the main tool used in the study of income inequality. For any distribution, the Lorenz curve $L(F)$ can be written in terms of the PDF, f, or CDF, F, as

$$L(F) = \frac{\int_{x_m}^{x(F)} xf(x)dx}{\int_{x_m}^{\infty} xf(x)dx} = \frac{\int_0^F x(F')dF'}{\int_0^1 x(F')dF'}$$

Where $x(F)$ is the inverse of the CDF.

The CDF for the power-law distribution is given by

$$F_X(x) = \begin{cases} 1 - (\frac{x_m}{x})^\alpha & , x \geq x_m \\ 0 & , x < x_m \end{cases}$$

In this case, $x(F)$, the inverse for the power-law distribution is

$$x(F) = \frac{x_m}{(1-F)^{1/\alpha}}$$

Substituting in $x(F)$ for power-law distribution, The Lorenz curve is calculated to be

$$L(F) = \frac{\left[-\frac{1}{1-\frac{1}{\alpha}} x_m (1-F)^{-\frac{1}{\alpha}+1}\right]_0^F}{\left[-\frac{1}{1-\frac{1}{\alpha}} x_m (1-F)^{-\frac{1}{\alpha}+1}\right]_0^1} = \frac{-\frac{\alpha}{\alpha-1} x_m (1-F)^{1-1/\alpha} + \frac{\alpha}{\alpha-1} x_m}{\frac{\alpha}{\alpha-1} x_m} = 1 - (1-F)^{1-1/\alpha}$$

Although the numerator and the denominator of $L(F)$ diverge for $0 \leq \alpha < 1$, their ratio does not, giving $L(F) = 0$ in these cases, which is equivalent to a Gini coefficient of 1.

The Gini coefficient is a measure of the deviation of the Lorenz curve from the equidistribution line ($\alpha = \infty$). Specifically, the Gini coefficient is twice the area between the Lorenz curve and the equidistribution line. The Gini coefficient for the power-law distribution (where $\alpha > 1$) is then calculated to be

$$\begin{aligned} G &= 1 - 2\left(\int_0^1 L(F)dF\right) \\ &= 1 - 2\left[F + \frac{1}{2-1/\alpha}(1-F)^{1-1/\alpha+1}\right]_0^1 \\ &= 1 - 2\left[1 - \frac{\alpha}{2\alpha-1}\right] = 1 - 2\frac{\alpha-1}{2\alpha-1} = 1 - \frac{2\alpha-2}{2\alpha-1} = \frac{1}{2\alpha-1} \end{aligned}$$

It is important to note that α used in the derivation here is the power-law exponent for the CDF of the power-law distribution. In this chapter we use α to represent the scaling exponent of the PDF of the power-law distribution. Therefore, when calculating the Gini coefficient from $\hat{\alpha}$ estimated using the CNS approach such as those in Tables 4.3 and 4.4, we must substitute in $\hat{\alpha} - 1$ instead, $G = \frac{1}{2(\hat{\alpha}-1)-1}$.

Chapter 5

Conclusion

This thesis examined three independent topics in empirical development and labour economics. In the main chapters, various microeconometrics techniques were applied to microdata from developing countries including Bangladesh and Thailand with the aim of achieving a consistent and reliable estimation of the object of interest.

The second chapter attempts to test and compare the impact that the types of transfers received have on various households' outcomes. This chapter makes use of a natural experiment in Bangladesh, where households experience an exogenous change in the nature of the transfer they received, from in-kind to cash, at different periods in time. The FE-IV approach was used in the estimation of the impact of the two types of transfers. This approach allows us to control for the unobserved time-invariant individual-specific effects, such as household preference, via the fixed effect estimator. The IV approach fixes the endogeneity of programme participation and possible measurement errors in the identification of the actual beneficiary households. Using the IV or fixed effect estimator by itself does not provide desirable consistent estimates. The FE-IV approach, which is a combination of the two, is preferred, although it requires the stronger assumption that the instruments, the actual eligibility status, is strongly exogenous in relation to the error term.

The estimation results show that though in-kind food grains transfers did cause households to consume more grain than they would have chosen under equal-value cash transfers, the impact on calorie consumption and children's health status is minimal; only marginally significant in the case of calorie consumption and statistically insignificant in the case of children's health. These households were able to substitute effectively out of similar food-

stuffs such as rice, keeping the overall cereal consumption around the same level in order to ease constraints and minimise distortions. Households that received cash, however, were able to reallocate their funds more effectively, and chose to spend their extra income on the purchase of useful items such as clothing and children's non-food consumption, while at the same time spending no more on undesirable items such as cigarettes. While other justifications for providing transfers in-kind may certainly be valid, the evidence supporting the paternalistic one in this instance is minimal. Given the sizeable cost saving that could result from replacing in-kind transfers with cash transfers, more empirical evidence looking more closely at the cost-benefit analysis of cash versus in-kind transfers is needed.

Chapter 3 investigates the dynamics of living standards in Thailand first through the statistical Galton-Markov model of earnings dynamics before proceeding to apply greater economic structure under the permanent earnings model framework, which seeks to model the covariance structure of the earnings process by decomposing earnings into permanent and transitory components and studying their evolution over the life cycle and over time. Empirical estimations of income and earnings persistence in Thailand were performed on constructed LFS pseudo-panel data and the Townsend Thai panel data.

Simple OLS estimates of the degree of persistence are likely to be biased as a result of measurement error in income, non-random attrition and non-representativeness of the sample. Various methods such as instrumental variable estimation, Arellano-Bond dynamic panel estimation, APC maximum entropy estimation and dynamic pseudo-panel methods were employed to ensure the consistent estimation of the degree of income and earnings persistence in Thailand. Galton-Markov estimates obtained from the Townsend panel data found unconditional and conditional income persistence to be low. Pseudo-panel estimates indicate a high level of unconditional persistence in Thailand. However, the introduction of socioeconomic factors and fixed effects lower the persistence estimates significantly.

Extending the analysis to other parts of the income distribution using the quantile regression technique finds that households at different parts of the distribution experience different degrees of income persistence. In both samples, persistence is found to be low at the bottom of the distribution but high at the top, indicating a divergence in income over time. This finding questions the low degree of persistence observed on average in earlier analysis and clearly illustrates the importance of exploring other parts of the distribution beyond the

conditional mean to obtain a more complete picture of the income and earnings process.

A study of the covariance structure of earnings following the permanent earnings model finds that total variation in the earnings process is predominantly driven by moderately persistent transitory components following the AR(1) process. Permanent components are found to be statistically insignificant in the LFS sample, and economically insignificant for the Townsend sample. More complex models of the covariance structure such as the random growth model are rejected in favour of the simple AR(1) model with fixed permanent elements. These findings from the variance decomposition analysis lend support to the result obtained in earlier analysis and the use of the Galton-Markov process to model earnings dynamics in Thailand.

Overall, these findings of low conditional persistence in both samples suggest that households are able to recover quickly from shocks to their labour earnings. Therefore, the high level of inequality found in Thailand is not due to earnings shocks having long-term effects. However, the fact that unconditional persistence remains high, especially at the top of the income and earnings distribution, suggests that household characteristics and fixed effects are important and that the high level of inequality found in the cross-section is likely to persist over time.

While it was originally intended for the findings of the two datasets to be used to validate one another, the differences between the persistence estimates obtained from the two samples stem mainly from the fact that they are made up of completely different sections of the Thai population. The Townsend Thai sample is only representative of the rural farming households in Thailand, whose income quickly fluctuates from year to year, while the LFS sample, especially after the year 2000, has an over-representation of people in waged employment earning stable salaries. Despite these shortcomings, studying both sets of data has led to greater understanding of the dynamics of living standards in the rural agricultural and wage-employed sectors in Thailand.

Chapter 4 attempts to fit the power-law distribution to the income and consumption distribution and study the dynamics of inequality at the top end of both distributions in Thailand. Given Thailand's rapid development from a low-income country to an upper-middle income country during the past 30 years, its income distribution provides an interesting case study of a distribution that could exhibit power-law behaviour with high inequality and a

heavy concentration of income at the top. Discovery of power law could also provide an indication that some interesting and complex income-generation mechanism may be at work in Thailand.

Fitting the power-law distribution to income and consumption data from Thailand's Household Socio-Economic Survey (SES, 1986 - 2006), we find that using the popular but incorrect method based on the linear regression approach will lead to researchers drawing a wrong conclusion. OLS estimates of the power-law exponent, α, provide strong evidence of power-law fit in Thailand, with R^2 exceeding 95% for income and consumption data in all the years. However, from the implementation of the superior Clauset et al. method, the evidence in support of the power-law fit is much weaker, with only slightly more than half of the data passing the power-law goodness-of-fit test on an absolute basis. Comparing the power-law fit with several competing distributions shows that the lognormal distribution offers an equally good fit to all datasets studied here. Thus, we only find moderate support for the power-law fit for Thailand.

Estimates of the power-law exponents for both income and consumption are greater than 3, larger than what is usually observed in the literature ($2 < \alpha < 3$), suggesting that there is less inequality at the top end of the income and consumption distribution in Thailand. Upon further inspection, we find that while there is less inequality at the top in Thailand, there is a high level of persistent between-group inequality between the top and bottom ends of the distribution. This finding helps explain and reconcile the observation of low top-end inequality but persistently high overall inequality in Thailand.

In addition, following the work of Battistin et al. (2009) to further investigate the plausibility of Gibrat's law of proportional random growth as the underlying income-generating process in Thailand, we find that the patterns observed in the data do not follow Gibrat's predictions in most instances. The lognormal distribution does not provide a good fit to the consumption distribution in Thailand. Furthermore, employing the KL divergence, we find that consumption, a proxy for permanent income that should follow Gibrat's law, has a distribution that deviates even further from the lognormal than reported income. These findings cast serious doubt on the validity of Gibrat's law of proportional random growth as the income-generating process in Thailand.

Bibliography

Abadie, A.: 2003, 'Semiparametric Instrumental Variable Estimation of Treatment Response Models'. *Journal of Econometrics* **113(2)**, 231–263.

Abowd, J. M. and D. Card: 1989, 'On the Covariance Structure of Earnings and Hours Changes'. *Econometrica* **57**(2), 411–45.

Ahmed, A. U.: 2005, 'Comparing food and cash incentives for schooling in Bangladesh'. *International Food Policy Research Institute, Washington, D.C., and United Nations University, Tokyo. Mimeo.*

Ahmed, A. U. and C. del Ninno: 2002, 'The Food For Education program in Bangladesh'. FCND discussion papers 138, International Food Policy Research Institute (IFPRI).

Ahmed, A. U. and M. Khondkar: 2010, 'An Analysis of Institutions and Policy Processes for Selected Antipoverty Interventions in Bangladesh'. *International Food Policy Research Institute, Washington, D.C. Mimeo.*

Aitchison, J. and J. A. C. Brown: 1957, *The Lognormal Distribution.* Cambridge University Press, Cambridge.

Altonji, J. G. and L. M. Segal: 1996, 'Small-Sample Bias in GMM Estimation of Covariance Structures'. *Journal of Business & Economic Statistics* **14**(3), 353–66.

Angelucci, M. and O. Attanasio: 2006, 'Estimating ATT Effects with Non-Experimental Data and Low Compliance'. IZA Discussion Papers 2368, Institute for the Study of Labor (IZA).

Angrist, J. D.: 2001, 'Estimations of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice'. *Journal of Business & Economic Statistics* **19**(1), 2–16.

Angrist, J. D. and J.-S. Pischke: 2009, *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Antman, F. and D. McKenzie: 2007a, 'Poverty Traps and Nonlinear Income Dynamics with Measurement Error and Individual Heterogeneity'. *The Journal of Development Studies* **43**(6), 1057–1083.

Antman, F. and D. J. McKenzie: 2007b, 'Earnings Mobility and Measurement Error: A Pseudo-Panel Approach'. *Economic Development and Cultural Change* **56**, 125–161.

Aoki, S. and M. Nirei: 2014, 'Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S'. UTokyo Price Project Working Paper Series 023, University of Tokyo, Graduate School of Economics.

Arellano, M.: 1987, 'Computing Robust Standard Errors for Within-Groups Estimators'. *Oxford Bulletin of Economics and Statistics* **49**(4), 431–34.

Arellano, M. and S. Bond: 1991, 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations'. *Review of Economic Studies* **58**(2), 277–97.

Atkinson, A. and F. Bourguignon: 2000, *Handbook of Income Distribution*, Vol. 1 of *Handbook of Income Distribution*. Elsevier.

Atkinson, A., F. Bourguignon, and C. Morrisson: 1992, *Empirical Studies of Earnings Mobility*. Harwood Academic Publishers.

Atkinson, A. B., F. Bourguignon, and C. Morrisson: 1988, 'Earnings mobility'. *European Economic Review* **32**(2-3), 619–632.

Atkinson, A. B. and T. Piketty: 2007, *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, No. 9780199286881 in OUP Catalogue. Oxford University Press.

Atkinson, A. B., T. Piketty, and E. Saez: 2011, 'Top Incomes in the Long Run of History'. *Journal of Economic Literature, American Economic Association* **49(1)**, 3–71.

Auerbach, F.: 1913, 'Das Gesetz der BevDlkerungskonzentration'. *Petermanns Geographische Mitteilungen* **59**, 74–76.

Axtell, R. L.: 2001, 'Zipf Distribution of U.S. Firm Sizes'. *Science* **293**, 1818–1820.

Bak, P., C. Tang, and K. Wiesenfeld: 1987, 'Self-organized criticality: An explanation of the 1/f noise'. *Physical Review Letters* **59**, 381–384.

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

Battistin, E., R. Blundell, and A. Lewbel: 2009, 'Why Is Consumption More Log Normal than Income? Gibrat's Law Revisited'. *Journal of Political Economy* **117**(6), 1140–1154.

Baulch, B.: 2010, 'The Medium-Term Impact of the Primary Education Stipend in Rural Bangladesh'. *IFPRI Discussion Paper 00976*.

Baulch, B. and J. Hoddinott: 2000, 'Economic mobility and poverty dynamics in developing countries'. *The Journal of Development Studies* **36**(6), 1–24.

Bearse, P., G. Glomm, and E. Janeba: 2000, 'Why poor countries rely mostly on redistribution in-kind'. *Journal of Public Economics* **75**(3), 463–481.

Bee, M., M. Riccaboni, and S. Schiavo: 2011, 'Pareto versus lognormal: a maximum entropy test'. Department of Economics Working Papers 1102, Department of Economics, University of Trento, Italia.

Bloom, H. S.: 1984, 'Accounting for No-shows in Experimental Evaluation Designs'. *Evaluation Review* **8**, 225–246.

Blundell, R. and S. Bond: 1998, 'Initial conditions and moment restrictions in dynamic panel data models'. *Journal of Econometrics* **87**(1), 115–143.

Blundell, R. and M. C. Dias: 2000, 'Evaluation methods for non-experimental data'. *Fiscal Studies* **21**(4), 427–468.

Bosker, M., S. Brakman, H. Garretsen, and M. Schramm: 2008, 'A century of shocks: the evolution of the German city size distribution 1925-1999'. *Regional Science and Urban Economics* **38 (4)**, 330–347.

Bourguignon, F. and C. ching Goh: 2004, 'Trade and labor vulnerability in Indonesia, Republic of Korea and Thailand'. In: *East Asia Integrates: a trade policy agenda for shared growth*.

Bourguignon, F., C.-c. Goh, and D. I. Kim: 2004, 'Estimating individual vulnerability to poverty with pseudo-panel data'. Policy Research Working Paper Series 3375, The World Bank.

Brakman, S., H. Garretsen, C. van Marrewikj, and M. van de Berg: 1999, 'The return of Zipf: towards a further understanding of the rank-size distribution'. *Journal of Regional Science* **39(1)**, 182–213.

Bratberg, E., O. A. Nilsen, and K. Vaage: 2005, 'Intergenerational Mobility: Trends Across the Earnings Distribution'. IZA Discussion Papers 1517, Institute for the Study of Labor (IZA).

Browning, M., I. Crawford, and M. Knoef: 2012, 'The Age-Period-Cohort Problem: Set Identification and Point Identification'. CeMMAP working papers CWP02/12, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Bruce, N. and M. Waldman: 1991, 'Transfers in Kind: Why They Can Be Efficient and Nonpaternalistic'. *American Economic Review* **81**(5), 1345–51.

Brzezinski, M.: 2014, 'Do wealth distributions follow power laws? Evidence from "rich lists"?'. *Physica A: Statistical Mechanics and its Applications* **406**(C), 155–162.

Calonico, S.: 2006, 'Pseudo-Panel Analysis of Earnings Dynamics and Mobility in Latin America'. *Inter-American Development Bank, Research Department. Mimeographed document.*

Cameron, A. C. and P. K. Trivedi: 2005, *Microeconometrics*, No. 9780521848053 in Cambridge Books. Cambridge University Press.

Capehart, K. W.: 2014, 'Is the wealth of the world's billionaires not Paretian?'. *Physica A: Statistical Mechanics and its Applications* **395**(C), 255–260.

Cappellari, L.: 2004, 'The Dynamics and Inequality of Italian Men's Earnings: Long-term Changes or Transitory Fluctuations?'. *Journal of Human Resources* **39**(2).

Chamberlain, G.: 1984, 'Panel data'. In: Z. Griliches and M. D. Intriligator (eds.): *Handbook of Econometrics*, Vol. 2 of *Handbook of Econometrics*. Elsevier, Chapt. 22, pp. 1247–1318.

Champernowne, D.: 1953, 'A model of income distribution'. *The Econonomic Journal* **63**, 318–351.

Clauset, A., C. R. Shalizi, and M. E. J. Newman: 2009, 'Power-Law Distributions in Empirical Data'. *SIAM Review* **51**(4), 661–703.

Clementi, F., T. Di Matteo, and M. Gallegati: 2006, 'The power-law tail exponent of income distributions'. *Physica A: Statistical Mechanics and its Applications* **370**(1), 49–53.

Clementi, F. and M. Gallegati: 2005, 'Power law tails in the Italian personal income distribution'. *Physica A: Statistical Mechanics and its Applications* **350**(2), 427–438.

Coate, S., S. Johnson, and R. Zeckhauser: 1994, 'Pecuniary redistribution through in-kind programs'. *Journal of Public Economics* **55**(1), 19–40.

Cuesta, J., H. Nopo, and G. Pizzolitto: 2011, 'Using Pseudo-Panels To Measure Income Mobility In Latin America'. *Review of Income and Wealth* **57**(2), 224–246.

Cunha, J.: 2010, 'Testing Paternalism: Cash vs. In-kind Transfers in Rural Mexico'.

Currie, J. and F. Gahvari: 2008, 'Transfers in Cash and In-Kind: Theory Meets the Data'. *Journal of Economic Literature* **46**(2), 333–83.

Daly, G. and F. J. Giertz: 1972, 'Welfare Economics and Welfare Reform'. *American Economic Review* **62**(1), 131–38.

Davis, D. and D. Weinstein: 2002, 'Bones, bombs, and break points: the geography of economic activity'. *American Economic Review* **92 (5)**, 1269–1289.

De Wit, G.: 2004, 'Firm Size Distributions'. *SCALES-paper N200418.*

De Wit, G.: 2005, 'Zipf's Law in Economics'. *SCALES-paper N200503.*

Deaton, A.: 1981, *"Theoretical and Empirical Approaches to Consumer Demand Under Rationing" In Essays in the Theory and Measurement of Consumer Behaviour, in honour of Sir Rochard Stone*. Cambridge University Press.

Deaton, A.: 1985, 'Panel data from time series of cross-sections'. *Journal of Econometrics* **30**(1-2), 109–126.

Deaton, A.: 1997, *The Analysis of Household Surveys, A Microeconometric Approach to Development Policy*. World Bank. The Johns Hopkins University Press.

Deaton, A. and J. Muellbauer: 1980, *Economics and Consumer Behavior*, No. 9780521296762 in Cambridge Books. Cambridge University Press.

Deaton, A. and C. Paxson: 1994, 'Intertemporal Choice and Inequality'. *Journal of Political Economy* **102**(3), 437–67.

Del Ninno, C. and P. Dorosh: 2003, 'Impacts of in-kind transfers on household food consumption: Evidence from targeted food programmes in Bangladesh'. *Journal of Development Studies* **40: 1**, 48–78.

Dew-Becker, I. and R. J. Gordon: 2005, 'Where Did Productivity Growth Go? Inflation Dynamics and the Distribution of Income'. *Brookings Papers on Economic Activity* **36**(2), 67–150.

Dew-Becker, I. and R. J. Gordon: 2008, 'Controversies about the Rise in American Inequality: A Survey'. CEPR Discussion Papers 6817, C.E.P.R. Discussion Papers.

Dolores Collado, M.: 1997, 'Estimating dynamic models from time series of independent cross-sections'. *Journal of Econometrics* **82**(1), 37–62.

Dragoset, L. M. and G. S. Fields: 2006, 'U.S. Earnings Mobility: Comparing Survey-Based and Administrative-Based Estimates'. Working Papers 55, ECINEQ, Society for the Study of Economic Inequality.

Dragulescu, A. and V. M. Yakovenko: 2001, 'Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States'. *Physica A: Statistical Mechanics and its Applications* **299**(1), 213–221.

Eeckhout, J.: 2004, 'Gibrat's Law for (All) Cities'. *American Economic Review* **94**(5), 1429–1451.

Eeckhout, J.: 2009, 'Gibrat's Law for (all) cities: A reply'. *American Economic Review* **99**, 1676–1683.

Eide, E. R. and M. H. Showalter: 1999, 'Factors Affecting the Transmission of Earnings across Generations: A Quantile Regression Approach'. *Journal of Human Resources* **34**(2), 253–267.

Fagiolo, G., L. Alessi, M. Barigozzi, and M. Capasso: 2008, 'On the Distributional Properties of Household Consumption Expenditures. The Case of Italy.'. *LEM Working Paper Series*.

Fazio, G. and M. Modica: 2015, 'Pareto or log-normal? A recursive-truncation approach to the distribution of (all) cities'. *Journal of Regional Science*.

Fertig, A.: 2001, 'Trends in Intergenerational Earnings Mobility'.

Fields, G.: 2005a, 'Does Income Mobility Equalize Longer-term Incomes? New Measures of an Old Concept'. *Mimeo, Cornell University*.

Fields, G.: 2005b, 'The Many Facets of Economic Mobility'. *Cornell University, Department of Economics. Mimeographed Document*.

Fields, G., P. Cichello, S. Freije, M. Menendez, and D. Newhouse: 2003, 'For Richer of for Poorer? Evidence from Indonesia, South Africa, Spain and Venezuela'. *Journal of Economics Inequality* pp. 67–99.

Fields, G., R. Duval Hernandez, S. Freije-Rodriguez, and M. L. Sanchez Puerta: 2007, 'Earnings Mobility in Argentina, Mexico, and Venezuela: Testing the Divergence of Earnings and the Symmetry of Mobility Hypotheses'. IZA Discussion Papers 3184, Institute for the Study of Labor (IZA).

Fields, G. S. and E. A. Ok: 1996a, 'The Meaning and Measurement of Income Mobility'. *Journal of Economic Theory* **71**(2), 349–377.

Fields, G. S. and E. A. Ok: 1996b, 'The Measurement of Income Mobility: An Introduction to the Literature'. Working Papers 96-05, C.V. Starr Center for Applied Economics, New York University.

Fields, G. S. and M. L. Sanchez Puerta: 2010, 'Earnings Mobility in Times of Growth and Decline: Argentina from 1996 to 2003'. *World Development* **38**(6), 870–880.

Fitzgerald, J., P. Gottschalk, and R. Moffitt: 1998, 'An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics'. *Journal of Human Resources* **33**(2), 251–299.

Fofack, H. and A. Zeufack: 1999, 'Dynamics of Income Inequality in Thailand: Evidence from Household Pseudo-Panel Data'. *The World Bank*.

Friedman, M.: 1957, *A Theory of the Consumption Function*, No. frie57-1 in NBER Books. National Bureau of Economic Research, Inc.

Friedman, M.: 1992, 'Do Old Fallacies Ever Die?'. *Journal of Economic Literature* **30**(4), 2129–32.

Fujiwaraa, Y., H. Aoyamac, C. Guilmib, W. Souma, and M. Gallegati: 2004, 'Gibrat and ParetoZipf revisited with European firms'. *Physica A: Statistical Mechanics and its Applications* **344**, 112–116.

Gabaix, X.: 1999, 'Zipf's Law For Cities: An Explanation'. *The Quarterly Journal of Economics* **114**(3), 739–767.

Gabaix, X.: 2009, 'Power Laws in Economics and Finance'. *Annual Review of Economics* **1**(1), 255–294.

Gabaix, X.: 2011, 'The Granular Origins of Aggregate Fluctuations'. *Econometrica* **79**(3), 733–772.

Gabaix, X. and A. Landier: 2008, 'Why Has CEO Pay Increased So Much?'. *The Quarterly Journal of Economics* **123**(1), 49–100.

Gan, L., D. Li, and S. Song: 2006, 'Is the Zipf Law spurious in explaining city-size distributions?'. *Economics Letters* **92**, 256–262.

Garfinkel, I.: 1973, '"Is In-kind Redistribution Efficient?"'. *Quarterly Journal of Economics* **87(2)**, 320–30.

Gibrat, R.: 1931, 'Les Inegalites Economiques'. *Librairie du Recueil Sirey, Paris.*

Giesecke, M., T. Bonke, and H. Luthen: 2011, 'The Dynamics of Earnings in Germany: Evidence from Social Security Records'. Annual Conference 2011 (Frankfurt, Main): The Order of the World Economy - Lessons from the Crisis 48692, Verein fÃŒr Socialpolitik / German Economic Association.

Glewwe, P. and P. Nguyen: 2002, 'Economic mobility in Vietnam in the 1990s'. Policy Research Working Paper Series 2838, The World Bank.

Gordon, R. J.: 2009, 'Misperceptions About the Magnitude and Timing of Changes in American Income Inequality'. NBER Working Papers 15351, National Bureau of Economic Research, Inc.

Gottschalk, P.: 1997, 'Inequality, Income Growth, and Mobility: The Basic Facts'. *Journal of Economic Perspectives* **11**(2), 21–40.

Gottschalk, P. and E. Spolaore: 2002, 'On the Evaluation of Economic Mobility'. *Review of Economic Studies* **69**(1), 191–208.

Greene, W. H.: 2011, *Econometric Analysis.* Prentice Hall, 7. edition.

Griliches, Z. and J. A. Hausman: 1986, 'Errors in variables in panel data'. *Journal of Econometrics* **31**(1), 93–118.

Groeneveld, R. A. and G. Meeden: 1984, 'Measuring Skewness and Kurtosis'. *The Statistician* **33**, 391–399.

Guvenen, F.: 2009, 'An Empirical Investigation of Labor Income Processes'. *Review of Economic Dynamics* **12**(1), 58–79.

Hill, B. M.: 1975, 'A Simple Approach to Inference About the Tail of a Distribution'. *Annals of Statistics* **3**, 1163–74.

Hinkley, D. V.: 1975, 'On Power Transformation to Symmetry'. *Biometrika* **62**, 101–111.

Hungerford, T. L.: 1993, 'U.S. Income Mobility in the Seventies and Eighties'. *Review of Income and Wealth* **39**(4), 403–17.

Hungerford, T. L.: 2011, 'How Income Mobility Affects Income Inequality: US Evidence in the 1980s and the 1990s'. *Journal of Income Distribution* **20**(1), 117–126.

Imbens, G. W.: 2010, 'Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)'. *Journal of Economic Literature* **48**(2), 399–423.

Imbens, G. W. and J. M. Wooldridge: 2009, 'Recent Developments in the Econometrics of Program Evaluation'. *Journal of Economic Literature* **47**(1), 5–86.

Ioannides, Y. M. and H. G. Overman: 2003, 'Zipf's law for cities: an empirical examination'. *Regional Science and Urban Economics* **33**(2), 127–137.

Jacoby, H. G.: 2002, 'Is There an Intrahousehold "Flypaper Effect"? Evidence From a School Feeding Programme'. *Economic Journal* **112**(476), 196–221.

Jarvis, S. and S. P. Jenkins: 1998, 'How Much Income Mobility Is There in Britain?'. *Economic Journal* **108**(447), 428–43.

Jayachandran, S., J. Cunha, and G. D. De Giorgi: 2010, 'The Price Effects of Cash Versus In-Kind Transfer'.

Jayadev, A.: 2008, 'A power law tail in India's wealth distribution: Evidence from survey data'. *Physica A: Statistical Mechanics and its Applications* **387**, 270–276.

Jaynes, E. T.: 1957a, 'Information Theory and Statistical Mechanics'. *Physical Review* pp. 106–620.

Jaynes, E. T.: 1957b, 'Information Theory of Statistical Mechanics II'. *Physical Review* pp. 108–171.

Jenkins, S. P.: 2011, *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*, No. 9780199226436 in OUP Catalogue. Oxford University Press.

Jeong, H.: 2008, 'Assessment Of Relationship Between Growth And Inequality: Micro Evidence From Thailand'. *Macroeconomic Dynamics* **12**(S2), 155–197.

Jones, C. I.: 2015, 'Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality'. *Journal of Economic Perspectives* **29**, 29–46.

Jones, C. I. and J. Kim: 2014, 'A Schumpeterian Model of Top Income Inequality'. NBER Working Papers 20637, National Bureau of Economic Research, Inc.

Kalwij, A. S. and R. Alessie: 2007, 'Permanent and transitory wages of British men, 1975-2001: year, age and cohort effects'. *Journal of Applied Econometrics* **22**(6), 1063–1093.

Kesten, H.: 1973, 'Random difference equations and Renewal theory for products of random matrices'. *Acta Mathematica* **131**, 207–248.

Khor, N. and J. Pencavel: 2010, 'Evolution of Income Mobility in the People's Republic of China: 1991-2002'. ADB Economics Working Paper Series 204, Asian Development Bank.

Klass, O. S., O. Biham, M. Levy, O. Malcai, and S. Solomon: 2007, 'The Forbes 400, the Pareto power-law and efficient markets'. *The European Physical Journal B - Condensed Matter and Complex Systems* **55**(2), 143–147.

Kleibergen, F. and R. Paap: 2006, 'Generalized reduced rank tests using the singular value decomposition'. *Journal of Econometrics* **133**(1), 97–126.

Koenker, R. and K. F. Hallock: 2001, 'Quantile Regression'. *Journal of Economic Perspectives* **15**(4), 143–156.

Koenker, R. W. and G. Bassett: 1978, 'Regression Quantiles'. *Econometrica* **46**(1), 33–50.

Krongkaew, M. and N. Kakwani: 2003, 'The Growth-equity Trade-off in Modern Economic Development: The Case of Thailand'. *Journal of Asian Economics* **14**(5), 735–757.

Kuznets, S.: 1955, 'Economic growth and income inequality'. *The American Economic Review* **45**(1), 1–28.

Lambert, P. J. and J. R. Aronson: 1993, 'Inequality Decomposition Analysis and the Gini Coefficient Revisited'. *The Economic Journal* **103**, 1221–1227.

Levy, M.: 2009, 'Gibrat's Law for (All) Cities: Comment'. *American Economic Review* **99**(4), 1672–75.

Levy, M. and H. Levy: 2003, 'Investment Talent and the Pareto Wealth Distribution: Theoretical and Experimental Analysis'. *The Review of Economics and Statistics* **85**(3), 709–725.

Levy, M. and S. Solomon: 1997, 'New evidence for the power-law distribution of wealth'. *Physica A: Statistical Mechanics and its Applications* **242**(1), 90–94.

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

Lillard, L. A. and R. J. Willis: 1978, 'Dynamic Aspects of Earning Mobility'. *Econometrica* **46**(5), 985–1012.

MaCurdy, T. E.: 1982, 'The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis'. *Journal of Econometrics* **18**(1), 83–114.

Malevergne, Y., V. Pisarenko, and D. Sornette: 2011, 'Gibrat's law for cities: uniformly most powerful unbiased test of the Pareto against the lognormal'. *Physical Review E* **83**.

McCulloch, N. and B. Baulch: 2000, 'Simulating the Impact of Policy upon Chronic and Transitory Poverty in Rural Pakistan'. *The Journal of Development Studies* **36**(6), 100–130.

McKenzie, D. J.: 2004, 'Asymptotic Theory for Heterogeneous Dynamic Pseudo-panels'. *Journal of Econometrics* **120**(2), 235–262.

McKenzie, D. J.: 2006, 'Disentangling Age, Cohort and Time Effects in the Additive Model'. *Oxford Bulletin of Economics and Statistics* **68**(4), 473–495.

Meng, X. and J. Ryan: 2010, 'Does a food for education program affect school outcomes? The Bangladesh case'. *Journal of Population Economics* **23**(2), 415–447.

Meyer, B. D. and J. X. Sullivan: 2003, 'Measuring the Well-Being of the Poor Using Income and Consumption'. *Journal of Human Resources*.

Meyer, B. D. and J. X. Sullivan: 2013a, 'Consumption and Income Inequality and the Great Recession'. *American Economic Review* **103**(3), 178–83.

Meyer, B. D. and J. X. Sullivan: 2013b, 'Consumption and Income Inequality in the U.S. Since the 1960s'. *Unpublished*.

Mitzenmacher, M.: 2003, 'A Brief History of Generative Models for Power Law and Lognormal Distributions'. *INTERNET MATHEMATICS* **1**, 226–251.

Moffitt, R.: 1989, 'Estimating the Value of an In-Kind Transfer: The Case of Food Stamps'. *Econometrica* **57**(2), 385–409.

Moffitt, R.: 1993, 'Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-sections'. *Journal of Econometrics* **59**(1-2), 99–123.

Moffitt, R. A. and P. Gottschalk: 1995, 'Trends in the Variances of Permanent and Transitory Earnings in the U.S. and Their Relation to Earnings Mobility'. Boston College Working Papers in Economics 444, Boston College Department of Economics.

Montes-Rojas, G.: 2011, 'Quantile Regression with Classical Additive Measurement Errors'. *Economics Bulletin* **31**(4), 2863–2868.

Morduch, J.: 1998, 'Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh'. Working Papers 198, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.

Navarro, A. I.: 2007, 'Exploring Income Mobility Patterns in Argentina using Quantile Regression'. *Department of Economics, Universidad de San Andres*.

Navarro, A. I.: 2010, 'Estimating Long Term Earnings Mobility in Argentina with Pseudo-Panel Data'. *Revista de Analisis Economico - Economic Analysis Review* **25**(2), 65–90.

Neary, J. P. and K. W. S. Roberts: 1980, 'The theory of household behaviour under rationing'. *European Economic Review* **13**(1), 25–42.

Newman, M. E. J.: 2005, 'Power laws, Pareto distributions and Zipf's law'. *Contemporary Physics*.

Nichols, A. L. and R. J. Zeckhauser: 1982, 'Targeting Transfers through Restrictions on Recipients'. *American Economic Review* **72**(2), 372–77.

Nickell, S. J.: 1981, 'Biases in Dynamic Models with Fixed Effects'. *Econometrica* **49**(6), 1417–26.

Ninno, C. D. and P. Dorosh: 2003, 'Impacts of in-kind transfers on household food consumption: Evidence from targeted food programmes in Bangladesh'. *The Journal of Development Studies* **40**(1), 48–78.

Nirei, M.: 2009, 'Pareto Distributions in Economic Growth Models'. IIR Working Paper 09-05, Institute of Innovation Research, Hitotsubashi University.

Nirei, M. and W. Souma: 2006, 'Income Distribution and Stochastic Multiplicative Process with Reset Event'. *in M. Gallegati, A.P. Kirman and M. Marsili (eds.), The Complex Dynamics of Economic Interaction: Essays in Economics and Econophysics, Springer, New York*.

Ogwang, T.: 2013, 'Is the wealth of the world's billionaires Paretian?'. *Physica A: Statistical Mechanics and its Applications* **392**(4), 757–762.

Okuyama, K., M. Takayasu, and H. Takayasu: 1999, 'Zipfs law in income distribution of companies'. *Phy* **269**, 125–131.

Pareto, V.: 1897, 'Cours d'Economie Politique'. *F. Rouge, Lausanne*.

Pazner, E. A.: 1972, 'Merit Wants and the Theory of Taxation'. *Public Finance = Finances publiques* **27**(4), 460–72.

Persky, J.: 1992, 'Retrospectives: Pareto's Law'. *Journal of Economic Perspectives* **6**(2), 181–192.

Pinto, C. M. A., M. Mendes Lopes, and J. A. Tenreiro Machado: 2012, 'A review of power laws in real life phenomena'. *Communications in Nonlinear Science and Numerical Simulation* **17**, 3558–3578.

Pollak, R. A. and T. J. Wales: 1979, 'Welfare Comparisons and Equivalence Scales'. *American Economic Review* **69**(2), 216–21.

Quah, D.: 1993, 'Galton's Fallacy and Tests of the Convergence Hypothesis'. *Scandinavian Journal of Economics* **95**(4), 427–43.

Rosen, S.: 1981, 'The Economics of Superstars'. *American Economic Review* **71**(5), 845–58.

Rozenfeld, H. D., D. Rybski, X. Gabaix, and H. A. Makse: 2011, 'The Area and Population of Cities: New Insights from a Different Perspective on Cities'. *American Economic Review* **101**(5), 2205–25.

Rubin, D. B.: 1980, 'Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment'. *Journal of the American Statistical Association* **75**(371), pp. 591–593.

Santos, A. L. and A. P. Souza: 2007, 'Earnings Inequality in Brazil: Is it Permanent or Transitory?'. *Brazillian Review of Econometrics* **27**, 259–283.

Shannon, C. E.: 1948, 'A Mathematical Theory of Communication'. *Bell System Technical Journal* **27**, 379–423.

Shorrocks, A.: 1978a, 'Income Inequality and Income Mobility'. *Journal of Economic Theory* **19**(2), 376–393.

Shorrocks, A. F.: 1976, 'Income Mobility and the Markov Assumption'. *Economic Journal* **86**(343), 566–78.

Shorrocks, A. F.: 1978b, 'The Measurement of Mobility'. *Econometrica* **46**, 1013–1024.

Simon, H. A.: 1955, 'On a class of skew distribution functions'. *Biometrica* **42**, 425–440.

Sinha, S.: 2006, 'Evidence for power-law tail of the wealth distribution in India'. *Physica A: Statistical Mechanics and its Applications* **359**(C), 555–562.

Skoufias, E. and T. Gonzalez-Cossio: 2008, 'The Impacts of Cash and In-Kind Transfers on Consumption and Labor Supply: Experimental Evidence from Rural Mexico'. Policy Research Working Paper Series 4778, The World Bank.

Sornette, D. and R. Cont: 1997, 'Convergent Multiplicative Processes Repelled from Zero: Power Laws and Truncated Power Laws'. *Journal de Physique I, France 7* pp. 431–444.

Stewart, M. B.: 2010, 'The Dynamics of Earnings Mobility in the UK'.

Stock, J. and M. Yogo: 2005, *"Testing for Weak Instruments in Linear IV Regression" Ch. 5 in J.H. Stock and D.W.K. Andrews (eds). Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg.* Cambridge University Press.

Thaler, R. H. and C. R. Sunstein: 2003, 'Libertarian Paternalism'. *American Economic Review* **93**(2), 175–179.

Thomas, D., E. Frankenberg, and J. P. Smith: 2001, 'Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey'. *Journal of Human Resources* **36**(3), 556–592.

Tietjen, K.: 2003, 'Bangladesh Primary Education Stipend Project: A Descriptive Analysis'. *Draft report prepared for the World Bank, Washington, D.C.*

Townsend, R. M.: 1994, 'Risk and Insurance in Village India'. *Econometrica* **62**(3), 539–91.

Verbeek, M. and F. Vella: 2005, 'Estimating Dynamic Models from Repeated Cross-sections'. *Journal of Econometrics* **127**(1), 83–102.

Vuong, Q. H.: 1989, 'Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses'. *Econometrica* **57**(2), 307–33.

Warunsiri, S. and R. McNown: 2010, 'The Returns to Education in Thailand: A Pseudo-Panel Approach'. *World Development* **38**(11), 1616–1625.

Whitmore, D.: 2002, 'What Are Food Stamps Worth?'. Working Papers 847, Princeton University, Department of Economics, Industrial Relations Section.

Wooldridge, J. M.: 2002, 'Inverse probability weighted M-estimators for sample selection, attrition and stratification'. CeMMAP working papers CWP11/02, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Working, H.: 1943, 'Statistical Laws of Family Expenditure'. *Journal of American Statistical Association* **38**, 43–56.

Yule, G. U.: 1924, 'A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis F.R.S'. *Philosophical Transactions* **213**, 21–84.

Zaidi, A., K. Rake, and J. Falkingham: 2001, 'Income Mobility in Later Life'. *ESRC Research Group: Simultaing Social Policy in an Aging Society (SAGE). London School of Economics.*

Zimmerman, D. J.: 1992, 'Regression toward Mediocrity in Economic Stature'. *American Economic Review* **82**(3), 409–29.