University of Zurich

Department of Economics

Center for Institutions, Policy and Culture in the Development Process

Working Paper Series

Working Paper No. 202

Estimation Income Processes in China

Xiaoyan Fan, Zheng Song and Yikai Wang

June 2010

Estimating Income Processes in China*

Xiaoyan Fan
Fudan University

Zheng Song
Fudan University

Yikai Wang
University of Zurich

June 29, 2010

Abstract

This paper estimates individual income processes in China. We find that in contrast with the results from the U.S., income shocks for Chinese individuals are more volatile but less persistent. We then impose the estimated income processes in a standard life-cycle model. The model implies a considerably larger size of precautionary savings in China. The estimated income processes increases the aggregate saving rate by ten to fifteen percentage points under empirically reasonable parameterizations.

JEL Classification: J31

Keywords: Income Inequality, Income Variance, RIP, HIP

*We would like to thank Fatih Guvenen, Jonathan Heathcote, Kjetil Storesletten, Gianluca Violante for helpful comments. Yikai Wang acknoledges Swiss National Science Foundation for financial support.

1 Introduction

Income process is a key factor influencing individual economic behaviors. Much of the quantitative research replies on estimated properties of income process. An important question that has been repeatedly addressed in the literature, among many others, is to what extent income risks affect savings when markets are incomplete. In a influential paper, for instance, Aiyagari (1994) found precautionary savings to be modest; risks implied by an income process estimated from PSID increase the aggregate saving rate by no more than three percentage points. Nevertheless, precautionary savings can be much larger if incomes are sufficiently risky. China might provide such an example in reality. On the one hand, China has an astonishingly high saving rate; the aggregate saving rates are above 40% in the last decade. On the other hand, the transition towards a market economy removed the strictly regulated wage scheme but introduced few risk-sharing arrangements, resulting in substantially higher income risks. These two observations indicate an important role of uninsured idiosyncratic risks on Chinese household saving decisions.

The purpose of this paper is therefore two folds. Our primary interest is in estimating individual income processes in China and in contrasting the results with the U.S. and some other countries. Then, we impose the estimated income processes in a standard life-cycle model. The goal is to investigate the quantitative importance of individual income risks for agagregate savings. Our quantitative results suggest that compared with those from PSID, income shocks for Chinese individuals are more volatile but less persistent. Moreover, the model implies considerably larger precautionary savings in China. The estimated Chinese individual income processes increases the aggregate saving rate by ten to fifteen percentage points under empirically reasonable parameterizations.

There are two widely adopted strategies for the estimation of income process: the "restricted income profiles" model (RIP) and "heterogeneous income profiles" model (HIP). RIP assumes a constant income growth rate across individuals, while HIP allows income growth rate to be individual- specific. RIP constantly yields higher estimates of the persistence of income shocks than HIP. In the present paper, we use RIP as our benchmark estimation strategy. Since precautionary motives tend to be dampened when income shocks are more persistent, the potentially upward bias of RIP pro- vides a lowerbound for the size of precautionary savings. [more discussion and detailed results here] Nevertheless, HIP is also applied as a robustness check. [results here]

Estimating RIP and HIP require data following income information of the same person over time. In this paper, we use household-level survey data from the China Household Income Project 2002 (CHIP), mainly con- ducted by the National Bureau of Statistics in China. The dataset provides comprehensive individual information including income, age, education and employment. In particular, the respondents were asked to report their cur- rent annual incomes not only for the current year, but for each of the past four years. This builds up a panel of individual incomes covering a period of five years, which can thus be comparable to income information from PSID (Khor and Pencavel, 2006).

A major concern of using retrospective data is the accuracy of reports. The problem is less severe in the present paper, since annual incomes, rather than historical events, are more likely to be kept in mind. Moreover, we focus on the respondents who provide income information in all five years. The exclusion of incomplete responses has a potential of reducing the proportion of misreports in our sample.

We construct a simple quantitative model to evaluate the impact of in- come risks implied by estimated income processes on aggregate saving. The basic framework is a small open economy with overlapping generations. Our model shows that [to be written] This paper contributes to the growing literature on Chinese household saving behaviors. [to be written]

2 Data

Our data source is from Chinese Household Income Project (CHIP). This project started from 1988, and was implemented in 1988, 1995 and 2002, by Chinese Academy of Social Sciences (CASS) and National Bureau of Statistics of China (NBS). It provides detailed information on income and consumption of households and individuals in both urban and rural areas. Our research uses the urban dataset of CHIP 2002, because China can not be considered as a market economy in 1988 or 1995, and we want to focus on urban areas.

CHIP2002 is a cross-section data containing the income and other information of individuals and households in 2002, and it also provides individual income from 1998 to 2001, in the form of the respondents' recall. Thus we can construct an income panel, which has a dimension of 5 years, from 1998 to 2002.

This paper focus on individual income. In many papers such as *Deaton and Paxson, 1994; Storesletten et al, 2004,* it is necessary to adjust individual income to household income, on the same basis of consumption data. However, in this paper, we do not make use of consumption information, therefore, we do not convert income to household level and we work on individual income. We define an individual's income as total income, which includes wage, bonus, allowance/subsidies, and living expenses for the laid-off. This is because the recalling data from 1998 to 2002 only provide total income in this form.

An individual is selected into the panel if the following conditions are met for the whole five years: (1) Total earnings are positive in each year. (2) Total earning growth rates are no larger than 10 and no less than 1/10 in any consecutive years. Using 10 and 1/10 as the critical values has the advantage that some problematic observations caused by the wrong place of decimal point. (3) This individual's age at that year is larger than 22 and smaller than 60. We impose this restriction because 22 is the age that in China college education is over and 60 is the retirement age.

Following is some statistic description of the cleaned data.

Income@Year	Obs	Mean	Std.Dev.	Min	Max
1998	13643	8488.555	7057.299	60	200000
1999	13643	8867.582	6691.866	60	200000
2000	13643	9504.256	7059.866	60	180000
2001	13643	10229.29	8130.877	60	240000
2002	13643	11377.28	8343.138	60	160000

Birthyear	Obs
1940-1944	811
1945-1949	1413
1950-1954	2268
1955-1959	2197
1960-1964	2167
1965-1969	1596
1970-1974	1092
1975-1979	469

3 Parametric Models of Income Process

3.1 A General Framework of Income Process

$$\begin{aligned}\log Income^i_{h,t} &= g\left(\theta_t, X^i_{h,t}\right) + y^i_{h,t}, \\ y^i_{h,t} &= \alpha^i_c + \beta^i h + z^i_{h,t} + \phi_t \varepsilon^i_{h,t}, \\ z^i_{h,t} &= \rho z^i_{h-1,t-1} + \pi_t \eta^i_{h,t},\end{aligned}$$

where $X^i_{h,t}$ represents observable control variables; $y^i_{h,t}$ is individual-specific (unexplainable) life-cycle earnings. Furthermore, $y^i_{h,t}$ is determined by: α^i_c fixed effect; β^i, heterogeneous income growth rate; η^i_h, persistent shock; ε^i_h, transitory shock.α^i_c and β^i are possibly correlated, while other shocks $\epsilon^i_{h,t}$ and $\eta^i_{h,t}$ are assumed to be i.i.d. .

Also, we can see that the variance of the unexplainable part of income $y^i_{h,t}$ potentially depends cohort (through $\mathrm{var}\left(\alpha^i_c\right)$ - cohort is defined as the group of people who are born in the same year), age (through $\beta^i h$, $z^i_{h,t}$, $\epsilon^i_{h,t}$), and year (through ϕ_t and π_t). Supposing that the cohort effects, age effects and year effects are independent of each other, we may write down the following equation:

$$var(y^i_{h,t}) = x(a,t,c) = g_1(h) + g_2(t) + g_3(c),$$

Literatures such as *Deaton and Paxon 1994, Heathcote et.al 2005* point out that since cohort, age and year are not linearly independent, these three effects can only identified with some assumptions on these effects. So depending on the assumptions on the cohort and year effects, together with the assumptions for the parameters β^i, there are several different ways of estimating income process. In this paper, we discuss three of them, as in the following.

3.2 Restricted Income Process with Cohort Effects

Storesletten et.al 2004 assumes that there are no year effects: $\pi_t = 1, \phi_t = 1$ and there is not heterogeneous income growth rate over age: $\beta^i = 0$. If so, the variance of unexplainable income $y^i_{h,t}$ becomes:

$$var\left(y^i_{h,t}\right) = a_c + var\left(y^i_h\right). \tag{1}$$

The first part is cohort effect: each cohort can have cohort-specific variance of fixed effect $var\left(\alpha^i_c\right)$. The second part only depends on age in the following way:

$$\begin{aligned} var\left(y^i_h\right) &= \sigma^2_\alpha + var\left(z^i_h\right) + \sigma^2_\varepsilon, \\ &= \sigma^2_\alpha + \sigma^2_\eta \frac{1-\rho^{2h}}{1-\rho^2} + \sigma^2_\varepsilon \end{aligned} \tag{2}$$

Since for any h, σ^2_α and σ^2_ϵ appear in the same way, to identify σ^2_α and σ^2_ϵ seperately, other moment conditions are needed. In Storesletten et.al 2004, the variance of the sum of three consecutive years' income is used as an additional moment:

$$var(y^i_{c,h}+y^i_{c,h+1}+y^i_{c,h+2}) = 9\sigma^2_\alpha+3\sigma^2_\varepsilon+((1+\rho)^2+1+\frac{(1+\rho+\rho^2)^2\left(1-\rho^{2h}\right)}{1-\rho^2})\sigma^2_\eta+9a_c. \tag{3}$$

3.3 Restricted Income Process with Year Effects

Some researches such as *Heathcote et.al 2005* show that in the U.S., year effects are more important than cohort effects, so another strand of literature allow year effects and estimate the income process. For example, *Heathcote et.al 2007* uses the following assumptions based on the general income process: (1)$var\left(\beta^i\right) = 0$; (2) $var\left(\alpha^i_c\right)$ is constant over cohorts; (3)π_t, ϕ_t vary across time; (4) For $t < 1$, $\pi_t = \pi_1, \phi_t = \phi_1$. Then the moment conditions become:

$$\begin{aligned} var\left(y^i_{h,t}\right) &= \left[\sigma^2_\alpha + \sigma^2_\beta h^2\right] + var\left(z^i_{h,t}\right) + \phi^2_t\sigma^2_\varepsilon, \\ cov\left(y^i_{h,t}, y^i_{h+n,t+n}\right) &= \left[\sigma^2_\alpha + \sigma^2_\beta h\left(h+n\right)\right] + \rho^n var\left(z^i_{h,t}\right), \\ var\left(z^i_{1,t}\right) &= \pi^2_t\sigma^2_\eta,\ h=1 \\ var\left(z^i_{h,1}\right) &= \pi^2_1\sigma^2_\eta\sum_{j=0}^{h-1}\rho^{2j},\ t=1, h>1 \\ var\left(z^i_{h,t}\right) &= \rho^2 var\left(z^i_{h-1,t-1}\right) + \pi^2_t\sigma^2_\eta,\ t>1, h>1 \end{aligned} \tag{4}$$

3.4 Heterogeneous Income Process with Year Effects

Guvenen 2008 drops the assumption that $var\left(\beta^i\right) = 0$. So given the variance and covariance of α^i, β^i - $\sigma^2_\alpha, \sigma^2_\beta, \sigma_{\alpha\beta}$, the moment conditions are different:

$$\begin{aligned} var\left(y^i_{h,t}\right) &= \left[\sigma^2_\alpha + 2\sigma_{\alpha\beta}h + \sigma^2_\beta h^2\right] + var\left(z^i_{h,t}\right) + \phi^2_t\sigma^2_\varepsilon, \\ cov\left(y^i_{h,t}, y^i_{h+n,t+n}\right) &= \left[\sigma^2_\alpha + \sigma_{\alpha\beta}\left(2h+n\right) + \sigma^2_\beta h\left(h+n\right)\right] + \rho^n var\left(z^i_{h,t}\right), \end{aligned}$$

3.5 Our Choice of Parametric Model

There are two important differences in CHIP dataset from the corresponding U.S. household level panel datasets such as PSID: (1) CHIP dataset has a much shorter time span: 5 years; (2) the relative importance of cohort effects and year effects can be different in China. Therefore we have two challenges: (1) the variance-covariance matrix over 5 years can only provide 15 moments conditions, while in a RIP or HIP model with year effects 11 or 13 parameters need to be estimated; (2) the assumption that the year effects before the first year in observation is the same as the first year seems too strong given the short time span in the dataset. For the first difficulty, our strategy is to use variance-covariance matrix over year and cohort to increase the number of moments. For the second, we drop this assumption, and try to estimate different cumulative year effects for different cohorts.

3.5.1 Moment Conditions

In *Heathcote et al. 2004* and *Guvenen 2007* they use $\frac{T(T+1)}{2}$ moment conditions to estimate $2T+1$ or $2T+3$ parameters, where T is the number of years in the data - more than 20 in PSID. However, given that in our data we have only $T = 5$ years, if we follow their approach and use only the moment conditions across year, we have to estimate 11 or 13 parameters with 15 moments, and from our simulation as the following, this may lead to a very imprecise result. Our solution for this problem is to use moment conditions across year and cohort: for example, instead summing up the variance of certain year's income across all cohorts as one moment condition, we use the income variance of each cohort-year pair. Therefore, we have $\frac{T(T+1)}{2}C = 615$ moment conditions, where C is the number of cohorts in our sample. To test if this approach gives more accurate result than only using year moments, especially in a short time span sample of 5 years, we do the following tests, for both RIP and HIP. For example, in the RIP model with year effects, we do the following: (1) simulate a sample of around 12000 observations in T years, using the parameters in the restricted income process estimated by *Heathcote et al. 2004*. (2) Estimate the income process using $\frac{T(T+1)}{2}$ moments. (3) Estimate the income process using $\frac{T(T+1)}{2}$ moments. (4) Repeat simulations and estimations 50 times. After getting 50 estimation results for all the parameters, we can calculate the mean and the standard error for the parameters in the 50 estimations. The following table shows the preciseness of estimating the RIP model.
Table ?: RIP

	True Value	$T = 5$	$T = 10$	$T = 20$
ρ	0.988	0.9795(0.0587)	1.0022 (0.0248)	0.9901 (0.0085)
σ_α^2	0.058	0.0782 (0.0792)	0.0423 (0.0469)	0.0543 (0.0168)
σ_η^2	0.015	0.0120 (0.0051)	0.0115 (0.0048)	0.0146 (0.0025)
σ_ε^2	0.061	0.0609 (0.0018)	0.0610 (0.0023)	0.0610 (0.0024)
		$T\,(=5) * C$	$T\,(=5) * C, \pi_{t<1}$	
		0.9878 (0.0040)	0.9886 (0.0035)	
		0.0574 (0.0055)	0.0584 (0.0046)	
		0.0154 (0.0022)	0.0148 (0.0020)	
		0.0604 (0.0028)	0.0612 (0.0023)	

We can see that given $T = 5$, the estimation is very imprecise, especially for ρ and σ_α^2. As we increase T in the simulation, the estimation becomes more and more precise, and for $T = 20$, the means of all parameters are pretty close to the true value. These simulation says that for our sample is too short to get a precise estimation, but this is not a problem for a longer sample, as in *Heathcote et al. 2004*.

The important finiding is that if we make use of year-cohort group moment conditions, which give richer information, the accuracy of the estimation, in the short-time span sample, is improved a lot. From the $T\,(=5) * C$ column of Table ?? we can see that even with a small T, using the cohort information can increase the preciseness of the estimation to a level comparable or even higher than a long time span ($T = 20$).

Similarly, we also do the simulations and estimations for the heterogeneous income process model. Different from HIP models, we find that it is more difficult to precisely estimate the parameters, even if the time span is longer, especially for $\sigma_\alpha^2, \sigma_{\alpha\beta},$ and σ_β^2. Using the year-cohort moments can help to make the estimated parameters closer to the parameters used in simulation, but the estimated $\sigma_{\alpha\beta}$ is still imprecise. Table 2: HIP

		$T = 5$	$T = 10$	$T = 20$
ρ	0.821	0.6998 (0.2804)	0.8527 (0.0727)	0.8369 (0.0315)
σ_α^2	0.022	-0.0630 (1.5199)	-0.1576 (0.5919)	0.0189 (0.3975)
$\sigma_{\alpha\beta}$	-0.0006	0.0014 (0.0436)	0.0063 (0.0175)	-0.0002 (0.0105)
σ_β^2	0.00038	0.0003 (0.0009)	0.0001 (0.0005)	0.0003 (0.0001)
σ_η^2	0.029	0.0430 (0.0586)	0.0334 (0.0070)	0.0290 (0.0031)
σ_ε^2	0.047	0.0367 (0.0622)	0.0477 (0.0027)	0.0479 (0.0030)
		$T(=5) * C$	$T(=5) * C, \pi_{t<1}$	
		0.8330 (0.0773)	0.9873 (0.0239)	
		0.0273 (0.0084)	0.0477 (0.0051)	
		-0.0019 (0.0024)	-0.0077 (0.0011)	
		0.0004 (0.0001)	0.0002 (0.0004)	
		0.0294 (0.0047)	0.0236 (0.0019)	
		0.0428 (0.0036)	0.0512 (0.0022)	

3.5.2 Year Effects Before the First Period in the Sample

In Heathcote et.al. 2007 and Guvenen 2008, the assumption that the year effects before the first period are the same is not harmful because the time span in the PSID dataset is very long, thus the year effects before the first sample year have effects on a small part of the observations. However, we can not assume this because of the time span and before 1998, which is the first sample year in our dataset, China experienced huge change in the economy and society. So when we estimate the year effects, we allow them to vary before the first sample period. As we can see in Table ??, even though this procddure increases the number of parameters, our approach of using $T * C$ provides sufficient moment condtions to get a precise estimation of RIP.

4 Empirical Results

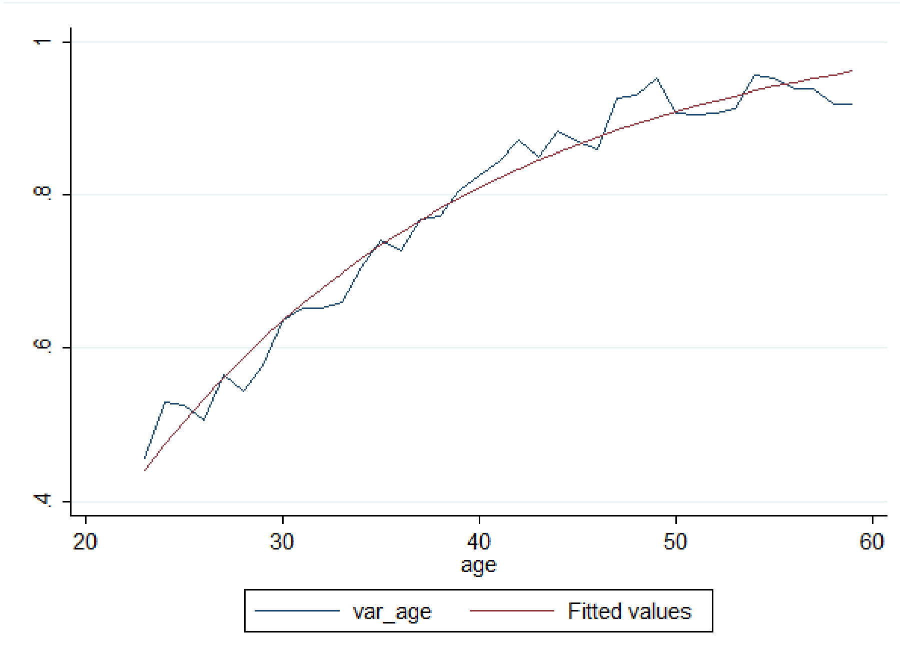
4.1 RIP with Cohort Effects

4.1.1 Within Cohort Age Profile of Income Inequality

Similar to Storesletten et.al 2004 we firstly use a simple and intuitive estimation to show the key patterns of the income process, and then do the formal GMM estimation and report the result. So the first step is to the estimate the cohort effects by seperating them from age effects. We construct cohorts and use the panel to follow cohorts through the time. Cohort is defined as an individual's birth year.

The variance of logarithm of income for each cohort at each year represent income inequality of that cohort-year pair. By simply assuming away year effects, we can decompose the variance of logrithm of income using dummy regressions into cohort and year effects, as in equation (1).

Figure 1 shows the age effects, which are the coefficients of age dummies in the OLS regression.



In this figure we can see that for certain cohort, the income inequality grows as age grows. In the first a few years in the labor market, income inequality grows fast, and as age grows, the speed of income inequality growth slows down, and after 50, the income inequality line becomes flat. This is a different pattern from what Storesletten, et al, 2004 find using USA data. They show that for certain American household cohort, the income inequality grows linearly until the retirement, and this is a sign of high persistency of income shocks. Correspondingly, our concave income inequality profile over life cycle implies a lower persistency of income shocks in China, which we will show formally below.

4.1.2 Implication of the Age Profile of Income Inequality

As equation (2) shows, the initial variance when a cohort enters the market, identifies the sum $\sigma_\alpha^2 + \sigma_\varepsilon^2$. The variance's increasing rate and curvature together can be used to identify the conditional variance of the persistent shocks and the autocorrelation, ρ. Compare our result of age profile of income variance with that in *Storesletten et.al., 2004,* it can be seen that (1) our initial variance of income is higher, so the income inequality of newly born workers are higher in China, in other words, the variance of fixed effects plus the variance of transitory shocks, $\sigma_\alpha^2 + \sigma_\varepsilon^2$, are higher. Income inequality in China grows faster than in USA for the young, but slower for the old. That is to say, conditional variance of the persistent shocks, σ_η^2 is higher in China, while the persistency (the autocorrelation, ρ) is lower than in USA.

We regress $var\left(y_h^i\right)$ on $\sigma_\alpha^2 + \sigma_\varepsilon^2$, σ_η^2, ρ, on the basis of equation (2) using nonlinear OLS estimation. The first column of the following table is the result we get from our data, while the second column shows USA parameters that *Storesletten et.al., 2004* gained from PSID household data, as a comparison. We can see that this result is consistent with what we can read from the age profile of income variance in Figure 1.

	China	USA
$\sigma_\alpha^2 + \sigma_\varepsilon^2$	0.4071	0.2735
σ_η^2	0.0344	0.0166
ρ	0.9726	0.9989

4.1.3 A GMM Estimation

The above approach of estimating parameters are useful in the sense that the mapping between moments and parameters is very transparent. However, this approach has limitations that we are unable to separately identify σ_α^2 and σ_ε^2, and that we may be ignoring other important aspects of the data. Next we will use a more formal GMM-based approach in order to separate σ_α^2 and σ_ε^2, also as a robustness check of the result of nonlinear OLS regression.

The reason why σ_α^2 and σ_ε^2 can not be separately identified is that in $var\left(y_h^i\right)$ for any h, σ_α^2 and σ_ε^2 always exist in the form of $\sigma_\alpha^2 + \sigma_\varepsilon^2$. So, adding other moments in which σ_α^2 and σ_ε^2 appear in other forms would help us to separate them. In order to make our results comparable with the results in Storesletten et.al 2004, we add the same the moments as them, which are the cross-sectional variances of the summation of cosecutive three year's income of individual i, based on equation (3). The following comparison of our estimation with Storesletten et.al. 2004 keeps the same pattern in the simple but intuitive nonlinear OLS regression: larger variance of shocks but lower persistency.

Table 1
Idiosyncratic Income Process: Parameter Estimates:

	China	**USA**
σ_α^2	0.303(0.018)	0.247(0.013)
σ_η^2	0.044 (0.003)	0.024 (0.006)
ρ	**0.969**(0.003)	**0.982**(0.011)
σ_ε^2	**0.032**(0.005)	0.061 (0.022)

4.2 RIP with Year Effects

The estimation produce is based on equations (4) as in Heathcote et.al. 2007 and deviates from it by allowing year effects $\pi_{t<1}$ to vary. This is equivalent to allowing the variance of accumulated persistent shocks at $t = 1$ to vary totally freely across cohort:

$$\begin{matrix} var\left(z^i_{37,1}\right) \\ var\left(z^i_{36,1}\right) \\ ... \\ var\left(z_{2,1}\right) \\ var\left(z_{1,1}\right) \end{matrix} = \begin{matrix} \rho^{36} & \rho^{35} & ... & \rho & 1 \\ 0 & \rho^{35} & ... & \rho & 1 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & \rho & 1 \\ 0 & 0 & ... & 0 & 1 \end{matrix} * \begin{matrix} \pi^2_{-35}\sigma^2_\eta \\ \pi^2_{-34}\sigma^2_\eta \\ ... \\ \pi^2_0\sigma^2_\eta \\ \pi^2_1\sigma^2_\eta \end{matrix}$$

From the above equation we can see that there is a one-on-one mapping between$var\left(z^i_{c,1}\right)$ and $\pi_{t<1}$. In our estimation we use dummies for $var\left(z^i_{c,1}\right)$ to capture the accumulated persistent shock of cohort c in the first sample period. Again, we compare the estimation results with the corresponding research on U.S. households income process in Heathcote et.al. 2007 in the following:

	China	USA
σ^2_α	0.1198	0.0578
σ^2_η	0.0469	0.0497
σ^2_ε	0.030	0.0218
ρ	0.9172	0.9426

$$\begin{aligned} \pi^2_t\sigma^2_\eta &= [0.0434, 0.0519, 0.0454] \\ \phi^2_t\sigma^2_\varepsilon &= [0.0206, 0.0199, 0.0193, 0.0413, 0.0588] \end{aligned}$$

The different pattern in China - a larger variance of persistent shock and a lower persistency - still exists.

4.3 HIP

As shown in the Monte-Carlo simulation, the data with a short time span $T = 5$ can not give a precise estimation, even if we increase the number of moment conditions to $T * C$. If we do the estimation in CHIP data, the standard GMM procedure doesn't converge. So we can not get a precise estimation for the HIP in China.

5 A Quantitative Model to Understand the Effect of the Different Pattern of the Income Process

What does a larger persistent shock variance and a smaller persistency in the income process imply for the economy, for example, the saving behavior? A larger variance of shock implies higher risk therefore other things equal, the economy will have higher precausionary savings. Meanwhile, a lower persistency means that the shocks are more about "luck" therefore increase the savings. We show this by imposing the Chinese income process and the U.S. income process seperately into a revised Aiyagari model, and comparing the different saving rate in the two economies.

5.1 The Model

Following Storesletten et.al. 2004, we incorporate the life cycle pattern in the revised Aiyagari model instead of infinitely living agents in the orginal one, and we add social security as an additional channel for the agent to smooth consumptions.

$V_h(z,a)$ is the value function of households:

$$V_h(z,a) = \max_{c,a'\geq 0} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta\,(1+g)^{1-\gamma}\,\xi_{h+1} E\left[V_{h+1}(\alpha, z', \varepsilon', a')\right] \right\}, \qquad (5)$$

subject to, before retirement,

$$\begin{aligned} c + (1+g)\,a' &\leq a_h R/\xi_h + n_h\,(1-\tau)\,W \\ a' &\geq \underline{a} \end{aligned}$$

after retirement,

$$\begin{aligned} c + (1+g)\,a' &\leq a_h R/\xi_h + bW \\ a' &\geq \underline{a} \end{aligned}$$

where,

$$\begin{aligned} \log n_h &= k_h + z_h, \\ z_h &= \rho z_{h-1} + \eta_h. \end{aligned}$$

The parameters are taken from *Storesletten et.al. (2004)*

age	$22, 65, 100$
ξ_h	US female 1991, growth rate 1.0%
g	1.5%
θ	0.36
δ	0.109
γ	2
β	0.962
income process	RIP with Cohort Effects
$\underline{a}$	0
b	0.5

We take RIP with cohort effects as the income process in the model economy, because we find that it is simple but sufficient to represent the different pattern between the Chinese income process and the U.S. one.

5.2 Different Saving Rate

- In a Partial equilibrium (given the interest rate of US), the different income process can increase the saving rate from 28.3% to 33.8% - a 5.5 percentage point increase.

- In a general equilibrium economy, the increased savings decreases the interest rate, therefore reduces the difference of saving rate to a 2.4 percentage point. However, this difference is still significant as we can see that this is equivalent to decreasing the replacement rate of pension payment from 50% to only 10%.

	ρ	σ_η^2	b	k	saving rate
US	0.982	0.024	0.5	2.2799	28.27%
CN-PE	0.9693	0.0166	0.5	2.7269	33.81%
CN-GE	0.9693	0.0166	0.5	2.4650	30.57%
US -low b	0.982	0.024	0.1	2.4598	30.50%

6 Conclusions

In this paper, we seriously consider the difference between the CHIP dataset and the PSID dataset and use Monte-Carlo test to show that our approach can give a precise estimation for the Ristrictive Income Process using a data with a short time span. Comparing with the U.S. researches, we find that the income process in China has a different pattern, and the most important charactoristics are: a larger variance of persistent shocks and a lower persistency. The different income process can generate a 5.5 percentage point higher saving rate in a partial equilibrium model, and in a general equilibrium model, 2.4 percentage point, which is equivalent to reducing replacement ratio of pension system from 50% to 10%. We find that the higher risk and lower persistency in the Chinese income process can explain a significant portion of the high household saving rate in China.