*Auburn University*
*Department of Economics*
*Working Paper Series*

# Testing the Predictability of Consumption Growth: Evidence from China

Liping Gao† and Hyeongwoo Kim‡

†Georgia Southern University; ‡Auburn University

***AUWP 2016-09***

# Testing the Predictability of Consumption Growth: Evidence from China

Liping Gao† and Hyeongwoo Kim‡

September 2016

*Abstract*

Using time series macroeconomic data, Chow (1985, 2010, 2011) reported indirect empirical evidence that implies the validity of the permanent income hypothesis in China. We revisit this issue by evaluating direct measures of the predictability of consumption growth in China during the post-economic reform regime (1978-2009). We also implement and report similar analysis for the postwar US data for comparison. Our in-sample analysis provides strong evidence against the PIH for both countries. Out-of-sample forecast exercises show that consumption changes are highly predictable, which sharply contrasts the implications of empirical findings by Chow (1985, 2010, 2011).

***Keywords***: Permanent Income Hypothesis; Consumption; Out-of-Sample Forecast; Diebold-Mariano-West Statistic

***JEL Classification***: E21; E27

† Department of Finance and Economics, Georgia Southern University, GA 30458, USA. Tel: 1-912-478-0184. Email: sophiagao1031@gmail.com.

‡ Department of Economics, Auburn University, AL 36849, USA. Tel: 1-334-844-2928. Fax: 1-334-844-4615. Email: gmmkim@gmail.com.

## 1 Introduction

In his early study, Chow (1985) reported strong evidence in favor of the Permanent Income Hypothesis (PIH) using annual observations in China from 1953 to 1982. Later, Chow (2010) re-evaluated his model for the *post*-economic reform regime, 1978 to 2006, then provided the same conclusion.[1] Using further updated data, Chow (2011) reported weaker but similar supporting evidence for the PIH in China.

We believe these findings are not convincing both theoretically and empirically. Under the PIH, optimizing consumers choose a stable path of consumption over their lifetime.[2] This implies that those consumers must borrow whenever their realized current income falls below their permanent income. However, in the absence of perfect capital markets, poor consumers may not be able to access to credit markets (Shoji, Aoyagi, Kasahara, Sawada, and Ueyama, 2012), which implies that liquidity constrained consumers may fail to obey the PIH.

One may claim that these consumers may still resort to informal credit markets as Guirkinger (2008) points out. We view these possibilities highly improbable in the case of China. For instance, using rural household data, Yuan and Xu (2015) show that the poor in China are severely excluded from not only formal but also informal credit markets. Therefore, it seems difficult to reconcile Chow's proposed indirect evidence of the PIH with such institutional facts about accessibility to credit markets in China.

From an empirical point of view, we note that his empirical analysis focuses on the coefficient of the lagged consumption in an autoregressive (AR) model for consumption. Finding that the confidence band of that coefficient includes 1, he suggested that the PIH is consistent with Chinese annual data. However, his statistical inference may not be valid when consumption obeys an integrated process, because the

---

[1] It is known that Chinese data in the *pre*-economic regime are sometimes unreliable.
[2] For detailed information, see standard macroeconomic textbooks such as Romer (2011).

conventional $t$-test he uses is invalid when consumption obeys an I(1) process.[3] Further, his work does not implement any direct statistical tests for the predictability of consumption growth in China.

This paper fills this gap by directly investigating the in-sample and the out-of-sample predictability of consumption changes in China as well as the postwar US data for comparison. Our findings suggest that consumption changes are highly predictable, which provides a clear contrast with findings shown in Chow (1985, 2010, 2011).

The rest of the paper is organized as follows. Section 2 provides a data description and preliminary test results. In Section 3, we provide our major findings. Section 4 concludes.

## 2 Data and Preliminary Analysis

We obtained per capita disposable income ($Y_t$) and consumption expenditure ($C_t$) data in China from China Statistical Yearbook (2012) following Chow (1985, 2010, 2011). Observations are annual and span from 1978 to 2012. It should be noted that observations prior to 1978 are excluded, because China began their major economic reforms since 1978, and the pre-economic reform regime data are often unreliable. We deflated all observations using the GDP deflator, obtained from the same source. All data are log transformed.

The log real per capita disposable income and the log real consumption expenditures of nondurable goods and services in the US are obtained from the Federal Reserve Economic Data (FRED). Observations are quarterly and cover the period from 1952:Q1 to 2011:Q4.

We first implement the augmented Dickey-Fuller (ADF) test for these variables. Results are reported in Table 1. The ADF test rejects the null of nonstationarity only for

[3] One should not use the normal distribution-based confidence band when the data generating process is an integrated process. That is, Chow's analysis based on such confidence bands is invalid.

differenced series with an exception of $y_t$ in China when time trend is present. Overall, our test results imply that consumption and income are integrated series, which is also consistent with Chow's work.

**Table 1**

## 3 Empirical Findings

### 3.1 In-Sample Analysis

Campbell and Mankiw (1990) test the validity of the PIH in the US by empirically evaluating Hall's (1978) famous claim: Consumption follows a random walk process under the PIH. They assume that a constant fraction of consumers ($\lambda$) does not obey the PIH, because they are liquidity constrained. Since they are not able to access the credit market, they are not capable of smoothing consumption over their lifetime. For these consumers, changes in consumption simply reflect income changes, that is, $\Delta C_t = \Delta Y_t$.

The rest of consumers, on the other hand, are assumed to obey the PIH, which implies $\Delta C_t = \varepsilon_t$, where $\varepsilon_t$ is a white noise process. Aggregating consumptions over a unit mass of agents yields the following estimable equation.

$$\Delta C_t = \lambda \Delta Y_t + u_t, \tag{1}$$

where $u_t = (1-\lambda)\varepsilon_t$. Naturally, finding a non-zero estimate for $\lambda$ provides evidence against the PIH.[4]

We employ their specification for Chinese data with an improved econometric method.[5] To deal with the endogeneity bias, we use the iterative efficient Generalized

[4] Campbell and Mankiw (1990) report significantly positive $\lambda$ estimates using the US data from 1953 to 1986, which implies strong evidence against the PIH. Similarly strong evidence is also reported by Flavin (1981).
[5] They use the two-stage instrumental variable estimator.

Method of Moments (GMM) estimator (Hansen, 1982) and report results in Table 2. Results from the US data are reported in Table 3.

The first column provides instrumental variables used in each regression. The second column reports $\lambda$ estimates along with their robust standard errors. The third column reports specification test ($J$-test) results, when applicable, along with corresponding $p$-values.

All $\lambda$ estimates are positive and statistically significant at the 5% level, which provides strong empirical evidence against the PIH. Our model specification seems reasonable as the $p$-value of the $J$ test statistics is greater than 0.05 in all regressions.

We also note that the $\lambda$ estimates are substantially greater with Chinese data than those from the US. Estimates for $\lambda$ vary from 0.835 to 0.872 in China, while it ranges from 0.287 to 0.769 in the US. This seems plausible because $\lambda$ is a fraction of liquidity constrained consumers, and households in China are more likely to be such consumers relative to US counterparts.

**Tables 2 and 3**

### 3.2 Out-of-Sample Predictability

We next implement a more direct test for the PIH via out-of-sample predictability tests proposed by Diebold and Mariano (1995) and West (1996). We evaluate the predictability of lagged variables for consumption changes relative to that of the benchmark random walk model, which is consistent with the PIH (Hall, 1978). That is,

$$C_{t+1}^{R} = C_t + u_t,$$

where $u_t$ is a white noise process. It is straightforward to see that this random walk model of $C_t$ implies the following.

$$C_{t+1|t}^{R} = C_t, \tag{2}$$

where $C^R_{t+1|t}$ is the *1*-step ahead consumption forecast by the random walk model given information set at time $t$. The competing alternative model that uses a vector of lagged variables as the explanatory variables ($X_t$) is described by the following least squares regression equation,

$$\Delta C^A_{t+1} = \beta' \Delta X_t + u_t \tag{3}$$

Note that we use the difference filter for the consistency of the least squares estimator, because $C_t$ and variables in $\Delta X_t$ obey integrated processes. Given the least squares coefficient estimates, $\hat{\beta}$, we construct the following *1*-step ahead forecast for the ***level*** consumption, $C^A_{t+1|t}$, under this alternative forecast model.

$$C^A_{t+1|t} = \widehat{\Delta C_{t+1|t}} + C_t, \tag{4}$$

where $\widehat{\Delta C_{t+1|t}} = \hat{\beta}' \Delta X_t$ is the fitted value from (3), and $C_t$ is the actual realized observation at time $t$.

Next, we define the following loss function differential variable,

$$d_t = L(\varepsilon^R_{t+1|t}) - L(\varepsilon^A_{t+1|t}),$$

where $L(\varepsilon^j_{t+k|t}),\ j = R, A$ is a loss function and forecast errors are,[6]

$$\varepsilon^R_{t+1|t} = C_{t+1} - C^R_{t+1|t}, \qquad \varepsilon^A_{t+1|t} = C_{t+1} - C^A_{t+1|t}$$

The Diebold-Mariano-West statistic (*DMW*) with the null of equal predictive accuracy, $H_0: Ed_t = 0$, is given,

$$DMW = \frac{\bar{d}}{\sqrt{\widehat{Avar}(\bar{d})}} \tag{5}$$

[6] We use the conventional squared error loss function, $(\varepsilon^j_{t+1|t})^2,\ j = R, A$ .

where $\bar{d} = \frac{1}{T-T_0}\sum_{t=T_o+1}^{T} d_t$ , and $\widehat{Avar}(\bar{d})$ is the asymptotic variance of $\bar{d}$ , $\frac{1}{T-T_0}\sum_{j=-q}^{q} k(j,q)\hat{\Gamma}_j$, where $k(\cdot)$ denotes a kernel function where $k(\cdot) = 0$, $j > q$, and $\hat{\Gamma}_j$ is the $j^{th}$ autocovariance function estimate.[7]

We carry out forecasting recursively by sequentially adding one additional observation from $P\%$ initial observations toward the end of observations. We re-estimate coefficients in the prediction model $\Delta C_{t+1} = \beta' \Delta X_t + u_t$ for each recursive sample, then re-formulate the forecasts.

The ratio of the root mean square prediction error (*RRMSPE*) is defined as the root mean square error of the random walk model relative to that of the competing model. Therefore, a greater value of *RRMSPE* than one implies evidence against the PIH, because the explanatory variables have added predictive power. Since the random walk model ($\beta = 0$) is nested by the alternative one, we evaluate the *DMW* statistics with McCracken's (2007) critical values to prevent size distortion problems, because the *DMW* test statistic is severely under-sized if asymptotic critical values are used when competing models are nested.[8]

We report our out-of-sample forecast exercise results in Tables 4 and 5 for China and the US, respectively. All *RRMSPE* values exceed one for both countries. Further, the DMW test rejects the null of equal predictability at any conventional significance level. Our results are quite robust to alternative initial split ratios.[9] Again, we obtain strong empirical evidence against the PIH via more direct out-of-sample forecast analysis.

**Tables 4 and 5**

[7] Following Andrews and Monahan (1992), we use the quadratic spectral kernel with automatic bandwidth selection for our analysis.
[8] Note that the alternative model nests the random walk model when $\beta$ is a null vector.
[9] The forecast performance may depend of the size of initial number of observations used in the estimation relative to the remaining observations for evaluations.

## 4 Concluding Remarks

This paper revisits the empirical inconsistency of the Permanent Income Hypothesis using household data in China along with the postwar US data as a benchmark.

We present strong evidence against the PIH in the sense that consumption growth is highly predictable, which is in contrast to the work by Chow (1985, 2010, 2011) who reported favorable indirect evidence using invalid normal approximation based tests.

Employing the model specification of Campbell and Mankiw (1990), our in-sample analysis provides very weak evidence of the PIH especially for China. That is, the $\lambda$ point estimates varied in the vicinity of 0.85 for China, which implies that majority consumers in China does not obey the PIH, because $\lambda$ is a fraction of consumers who are liquidity constrained. We obtained smaller values for $\lambda$ using the US data, which seems reasonable because consumers are able to access to credit more easily in the US.

Our out-of-sample forecasting exercises directly address the predictability issue from the PIH. We obtain very strong results against the PIH for both countries in the sense that explanatory variables have substantial predictive contents for consumption growth, which is robust to the choice of sample split ratio.

## References


J.Y. Campbell, N.G. Mankiw (1990), "Permanent income, current income, and consumption," *Journal of Business and Economic Statistics* 8(3), 265–279.

G. Chow (1985), "A model of Chinese national income determination," *Journal of Political Economy* 93(4), 782–792.

G. Chow (2010), "Note on a model of Chinese national income determination," *Economics Letters* 106, 195–196.

G. Chow (2011), "Lessons from studying a simple macroeconomic model for China," *Economics Letters* 112, 233–235.

F.X. Diebold, R.S. Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics* 13, 253–263.

M.A. Flavin (1981), "The adjustment of consumption to changing expectations about future income," *Journal of Political Economy* 89(5), 974–1009.

C. Guirkinger (2008), "Understanding the coexistence of formal and informal credit markets in Piura, Peru," *World Development* 36(8), 1436–1452.

R.E. Hall (1978), "Stochastic implications of the life cycle-Permanent Income Hypothesis: theory and evidence," *Journal of Political Economy* 86(6), 971–987.

L.P. Hansen (1982), "Large sample properties of Generalized Method of Moments estimators," *Econometrica* 50(4), 1029–1054.

M.W. McCracken (2007), "Asymptotics for out of sample tests of Granger causality," *Journal of Econometrics* 140(2), 719-752.

D. Romer (2011), *Advanced Macroeconomics*. 4$^{th}$ Edition, McGraw-Hill, NYC, NY.

M. Shoji, K. Aoyagi, R.J. Kasahara, Y. Sawada, M. Ueyama (2012), "Social capital formation and credit access: Evidence from Sri Lanka," *World Development* 40(12), 2522–2536.

K.D. West (1996), “Asymptotic inference about predictive ability,” *Econometrica* 64, 1067-1084.

Y. Yuan, L. Xu (2015), “Are poor able to access the informal credit market? Evidence from rural households in China,” *China Economic Review* 33, 232-246.

**Table 1. ADF Test Results**

*China*

| Variable | $ADF_c$ | $ADF_t$ |
|---|---|---|
| $C_t$ | -0.889 | -2.693 |
| $Y_t$ | 0.044 | -4.641* |
| $\Delta C_t$ | -3.885* | -3.907† |
| $\Delta Y_t$ | -4.263* | -4.163† |

*US*

| Variable | $ADF_c$ | $ADF_t$ |
|---|---|---|
| $C_t$ | -1.839 | -0.688 |
| $Y_t$ | -2.093 | -0.503 |
| $\Delta C_t$ | -4.966* | -5.280* |
| $\Delta Y_t$ | -9.416* | -14.62* |

Note: $ADF_c$ and $ADF_t$ denote the *ADF t*-statistic with an intercept and with an intercept and time trend, respectively. We chose the number of lags by the Akaike Information Criteria with a maximum 8 lags for quarterly US data, while a maximum 2 lags were used for annual Chinese data.

**Table 2. GMM Estimation Results: Chinese Data (Annual)**

| Instruments | $\lambda$ (s.e) | $J$ ($p$-value) |
|---|---|---|
| None (OLS) | 0.872 (0.018) | $n.a$ |
| $\Delta Y_{t-1}$ | 0.864 (0.015) | $n.a$ |
| $\Delta C_{t-1}$ | 0.868 (0.015) | $n.a$ |
| $\Delta Y_{t-1}, \Delta C_{t-1}$ | 0.835 (0.013) | 3.637 (0.057) |
| $\Delta Y_{t-1}, \Delta Y_{t-2}$ | 0.862 (0.013) | 1.619 (0.203) |
| $\Delta C_{t-1}, \Delta C_{t-2}$ | 0.868 (0.014) | 0.081 (0.776) |
| $\Delta Y_{t-1}, \Delta Y_{t-2}, \Delta C_{t-1}, \Delta C_{t-2}$ | 0.844 (0.008) | 4.836 (0.184) |

Note: Annual observations span from 1978 to 2009. This table reports iterative efficient GMM estimates of $\Delta C = \lambda Y_t + u_t$, using an array of instrumental variables. Numbers in parentheses in column 2 are standard errors for the $\lambda$ estimate. $p$-values for the $J$-test statistic are from the chi-square distribution.

**Table 3. GMM Estimation: US Data (Quarterly)**

| Instruments | $\lambda$ (s.e.) | $J$ ($p$-value) |
|---|---|---|
| None (OLS) | 0.287 (0.045) | *n.a* |
| $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}$ | 0.769 (0.282) | 1.240 (0.538) |
| $\Delta Y_{t-2}, \ldots, \Delta Y_{t-6}$ | 0.447 (0.157) | 4.989 (0.288) |
| $\Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 0.628 (0.137) | 5.885 (0.053) |
| $\Delta C_{t-2}, \ldots, \Delta C_{t-6}$ | 0.664 (0.132) | 8.291 (0.082) |
| $\Delta Y_{t-2}, \ldots, \Delta Y_{t-4}, \Delta C_{t-2}, \ldots, \Delta C_{t-4}$ | 0.487 (0.112) | 9.902 (0.078) |
| $\Delta Y_{t-2}, \ldots, \Delta Y_{t-6}, \Delta C_{t-2}, \ldots, \Delta C_{t-6}$ | 0.505 (0.094) | 13.06 (0.160) |

Note: Following Campbell and Mankiw (1990), we use consumption expenditures for nondurable goods and services, and we excluded the first lagged variables from the set of instruments. Observations are quarterly and span from 1952:Q1 to 2011:Q4. This table reports iterative efficient GMM estimates of $\Delta C_t = \lambda Y_t + u_t$, using an array of instrumental variables. Numbers in parentheses in column 2 are standard errors for the $\lambda$ estimate. $p$-values for the $J$-test statistic are from the chi-square distribution.

**Table 4. Out-of-Sample Forecast: China**

| Split Ratio ($P$) | Explanatory Variables | *RRMSPE* | *DMW* |
|---|---|---|---|
| 0.51 | $\Delta Y_t$ | 3.5999 | 6.6756* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}$ | 3.4202 | 7.3507* |
| | $\Delta C_{t-1}, \Delta C_{t-2}$ | 3.3188 | 7.2628* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}, \Delta C_{t-1}, \Delta C_{t-2}$ | 3.3204 | 7.1548* |
| 0.66 | $\Delta Y_t$ | 3.6606 | 6.8638* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}$ | 3.3623 | 7.6711* |
| | $\Delta C_{t-1}, \Delta C_{t-2}$ | 3.7080 | 8.3429* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}, \Delta C_{t-1}, \Delta C_{t-2}$ | 3.6638 | 8.5777* |
| 0.81 | $\Delta Y_t$ | 4.1636 | 5.0213* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}$ | 5.3395 | 6.0435* |
| | $\Delta C_{t-1}, \Delta C_{t-2}$ | 7.0499 | 6.5777* |
| | $\Delta Y_{t-1}, \Delta Y_{t-2}, \Delta C_{t-1}, \Delta C_{t-2}$ | 5.5285 | 6.5049* |

Note: Out-of-sample forecasting was recursively implemented by sequentially adding one additional observation from *P*% initial observations toward the end of observations. Split ratio denotes the number for *P*. *RRMSPE* denotes the ratio of the root mean squared prediction error of the random walk hypothesis to the competing model. *DMW* denotes the test statistics of Diebold and Mariano (1995) and West (1996). * denotes rejection of the null hypothesis of equal predictability at the 1% significance levels.

**Table 5. Out-of-Sample Forecast Analysis: US**

| Split Ratio ($P$) | Explanatory Variables | *RRMSPE* | *DMW* |
|---|---|---|---|
| 0.50 | $\Delta Y_t$ | 1.8411 | 10.322* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}$ | 1.9177 | 10.602* |
| | $\Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 2.0158 | 10.806* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}, \Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 2.0068 | 10.642* |
| 0.65 | $\Delta Y_t$ | 1.7986 | 9.8501* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}$ | 1.8993 | 9.9372* |
| | $\Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 2.0615 | 10.660* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}, \Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 2.0824 | 10.892* |
| 0.81 | $\Delta Y_t$ | 1.4856 | 5.4806* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}$ | 1.5825 | 5.9971* |
| | $\Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 1.7480 | 6.9882* |
| | $\Delta Y_{t-2}, \Delta Y_{t-3}, \Delta Y_{t-4}, \Delta C_{t-2}, \Delta C_{t-3}, \Delta C_{t-4}$ | 1.7572 | 7.1435* |

Note: Out-of-sample forecasting was recursively implemented by sequentially adding one additional observation from $P$% initial observations toward the end of observations. Split ratio denotes the number for $P$. *RRMSPE* denotes the ratio of the root mean squared prediction error of the random walk hypothesis to the competing model. *DMW* denotes the test statistics of Diebold and Mariano (1995) and West (1996). * denotes rejection of the null hypothesis of equal predictability at the 1% significance levels.